Abstract: Background: In recent years, evidence has emerged regarding the effectiveness of osteopathic manipulative treatments (OMT). Despite growing evidence in this field, there is need for appropriate research designs that effectively reflect the person-centred system of care promoted in osteopathy and provide data which can inform policy decisions within the healthcare system.

Objective: Identify, appraise and synthesise the evidence from comparative effectiveness and economic evaluation research involving OMT.

Design: Systematic literature review.

Methods: A database search was conducted using CINAHL, PubMed, PEDro, AMED, SCOPUS and OSTMED.DR, from their inception to May 2015. Two separate searches were undertaken to identify original research articles encompassing the economic evaluation and comparative effectiveness of OMT. Identified comparative effectiveness studies were evaluated using the Cochrane risk of bias tool and appraised using the Good Reporting of Comparative Effectiveness (GRACE) principles. Identified economic studies were assessed with the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) guidelines.

Results: Sixteen studies reporting the findings of comparative effectiveness (n=9) and economic evaluation (n=7) research were included. The comparative effectiveness studies reported outcomes for varied health conditions and the majority (n=6) demonstrated a high risk of bias. The economic evaluations included a range of analyses and considerable differences in the quality of reporting were evident.

Conclusion: Despite some positive findings, published comparative effectiveness and health economic studies in OMT are of insufficient quality and quantity to inform policy and practice. High quality, well-designed, research that aligns with international best practice is greatly needed to build a pragmatic evidence base for OMT.
21 June 2016

Dear Dr Ann Moore,

Please find attached the article ‘Osteopathic manipulative treatment: A systematic review and critical appraisal of comparative effectiveness and health economics research’, for exclusive consideration for publication in the Manual Therapy Journal.

This manuscript presents the first review of the comparative effectiveness and cost effectiveness research in osteopathic manipulative treatment. The review offers a critical appraisal of existing research and provides recommendations to strengthen future research in this important field.

All the authors have made substantial contributions to this manuscript. Drs Amie Steel and Tobias Sundberg conceptualised the study and, in collaboration with Ms Rebecca Reid, developed the overall protocol for the review. Ms Reid undertook the search for manuscripts in line with the developed protocol and Dr Sundberg extracted the data from included papers with sample verification of both stages by Dr Steel. Drs Lesley Ward and Felicity Bishop were responsible for the Risk of Bias assessments and Dr Ward collaborated with Dr Holger Cramer to complete the appraisal of included papers according to the GRACE guidelines. Assessment of papers to determine their alignment with comparative effectiveness principles in line with the PRECIS-2 tool was undertaken by Drs Steel and Sundberg. Evaluation of health economic research manuscripts in accordance with the CHEERS statement was completed by Dr Steel and Dr Jon Wardle. Dr Matthew Leach contributed substantially to the interpretation of findings and recommendations for future research. All authors contributed to the drafting, editing and finalisation of the submitted manuscript.
We feel that the significance of our findings will be of particularly interest to the international audience of your journal.

Yours sincerely,

Amie Steel

Australian Research Center in Complementary and Integrative Medicine
Faculty of Health
University of Technology, Sydney
Osteopathic manipulative treatment: A systematic review and critical appraisal of comparative effectiveness and health economics research

Amie Steel¹,², Tobias Sundberg¹,³, Rebecca Reid², Lesley Ward¹,⁴, Felicity L Bishop¹,⁵, Matthew Leach¹,⁶, Holger Cramer¹,⁷, Jon Wardle¹, Jon Adams¹

¹Australian Research Centre in Complementary and Integrative Medicine, Faculty of Health, University of Technology Sydney, 15 Broadway, Ultimo NSW 2007, Australia

²Endeavour College of Natural Health, Level 2, 269 Wickham St Fortitude Valley QLD 4006, Australia

³Research Unit for Studies of Integrative Health Care, Karolinska Institutet (NVS/OMV), Alfred Nobels Alle 23, 141 83 Stockholm, Sweden

⁴Nuffield Department of Orthopaedics, Rheumatology, and Musculoskeletal Sciences (NDORMS), B4495, Oxford OX3 7LD, University of Oxford, Oxford, United Kingdom

⁵Centre for Applications of Health Psychology, Faculty of Social Human and Mathematical Sciences, Building 44 Highfield Campus, University of Southampton, Southampton, SO17 1BJ, United Kingdom

⁶School of Nursing & Midwifery, Health Economics & Social Policy Group, University of South Australia, 101 Currie St, Adelaide, Australia

⁷Department of Internal and Integrative Medicine, Kliniken Essen-Mitte, Faculty of Medicine, University of Duisburg-Essen, 45141 Duisburg, Germany

Corresponding Author
Dr Amie Steel
Email: amie.steel@uts.edu.au
Phone: +61 7 3253 9523
Endeavour College of Natural Health
Level 2, 269 Wickham St
Fortitude Valley QLD 4006
Australia
Osteopathic manipulative treatment: A systematic review and critical appraisal of comparative effectiveness and health economics research

ABSTRACT

Background: In recent years, evidence has emerged regarding the effectiveness of osteopathic manipulative treatments (OMT). Despite growing evidence in this field, there is need for appropriate research designs that effectively reflect the person-centred system of care promoted in osteopathy and provide data which can inform policy decisions within the healthcare system.

Objective: Identify, appraise and synthesise the evidence from comparative effectiveness and economic evaluation research involving OMT.

Design: Systematic literature review.

Methods: A database search was conducted using CINAHL, PubMed, PEDro, AMED, SCOPUS and OSTMED.DR, from their inception to May 2015. Two separate searches were undertaken to identify original research articles encompassing the economic evaluation and comparative effectiveness of OMT. Identified comparative effectiveness studies were evaluated using the Cochrane risk of bias tool and appraised using the Good Reporting of Comparative Effectiveness (GRACE) principles. Identified economic studies were assessed with the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) guidelines.

Results: Sixteen studies reporting the findings of comparative effectiveness (n=9) and economic evaluation (n=7) research were included. The comparative effectiveness studies reported outcomes for varied health conditions and the majority (n=6) demonstrated a high risk of bias. The economic evaluations included a range of analyses and considerable differences in the quality of reporting were evident.
Conclusion: Despite some positive findings, published comparative effectiveness and health economic studies in OMT are of insufficient quality and quantity to inform policy and practice. High quality, well-designed, research that aligns with international best practice is greatly needed to build a pragmatic evidence base for OMT.

Keywords: osteopathy, osteopathic manipulative treatment, comparative effectiveness research, economic evaluation, pragmatic research
Osteopathic manipulative treatment: A systematic review and critical appraisal of comparative effectiveness and health economics research

ABSTRACT

Background: In recent years, evidence has emerged regarding the effectiveness of osteopathic manipulative treatments (OMT). Despite growing evidence in this field, there is need for appropriate research designs that effectively reflect the person-centred system of care promoted in osteopathy and provide data which can inform policy decisions within the healthcare system.

Objective: Identify, appraise and synthesise the evidence from comparative effectiveness and economic evaluation research involving OMT.

Design: Systematic literature review.

Methods: A database search was conducted using CINAHL, PubMed, PEDro, AMED, SCOPUS and OSTMED.DR, from their inception to May 2015. Two separate searches were undertaken to identify original research articles encompassing the economic evaluation and comparative effectiveness of OMT. Identified comparative effectiveness studies were evaluated using the Cochrane risk of bias tool and appraised using the Good Reporting of Comparative Effectiveness (GRACE) principles. Identified economic studies were assessed with the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) guidelines.

Results: Sixteen studies reporting the findings of comparative effectiveness (n=9) and economic evaluation (n=7) research were included. The comparative effectiveness studies reported outcomes for varied health conditions and the majority (n=6) demonstrated a high risk of bias. The economic evaluations included a range of analyses and considerable differences in the quality of reporting were evident.
Conclusion: Despite some positive findings, published comparative effectiveness and health economic studies in OMT are of insufficient quality and quantity to inform policy and practice. High quality, well-designed, research that aligns with international best practice is greatly needed to build a pragmatic evidence base for OMT.

Keywords: osteopathy, osteopathic manipulative treatment, comparative effectiveness research, economic evaluation, pragmatic research
INTRODUCTION

Osteopathic healthcare is a holistic person-centred system of care aligned with the philosophy, principles of practice and application of osteopathic manipulative treatment (OMT) (1). While OMT can be prescribed for the management of various health conditions, it is most commonly indicated for the care of painful disorders such as low back pain (2), headaches (3), and neck pain (4). In addition to OMT, osteopathic practitioners may also prescribe other medical therapies (including pharmaceuticals) depending on the medico-legal and regulatory standards of the country and scope of osteopathic training and practice (1, 5). Osteopathic healthcare has diverse representation across health systems internationally, ranging from full integration within conventional health care systems (i.e. in the US), through to semi-integration as allied and complementary health care therapists (i.e. in many European and Australasian countries) (1). Nonetheless, the primary shared component of all streams of osteopathic practice is OMT applied with an understanding of the relationship between the structure and function of the human body (1).

The prevalence of osteopathic healthcare use has been reported at less than 5% of the general population in Australia (6) and as high as 16% in the United States (7). However, higher rates of use in Australia are found amongst specific populations such as the middle-aged (16%) (8) and pregnant women (6.2%) (9). The use of osteopathic services is also much higher amongst individuals with specific health conditions, for example 13.4% of UK adults with back pain (10). This suggests that there is modest demand for osteopathic healthcare services, at least in Western countries.

Research investigating the effectiveness of osteopathic treatments has intensified over the past decade. Current evidence from randomised controlled trials (RCTs) and systematic reviews of RCTs suggests that osteopathy-related interventions are effective in improving outcomes in patients with back pain (11), neck pain (12), sciatica (13), chronic obstructive pulmonary disease (14), irritable bowel syndrome (15, 16) and various paediatric conditions (17). While the RCT design is considered the gold standard for demonstrating efficacy (18),
concerns have been raised regarding the applicability of the explanatory RCT to everyday clinical practice (REF). These concerns have led to the promotion of comparative effectiveness research (CER) - a pragmatic research design generating evidence that can be more efficiently translated into patient care and health policy (19). CER is defined as “the generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat, and monitor a clinical condition or to improve the delivery of care” (20). Suggested key elements of CER include (a) direct comparisons of active treatments; (b) study patients, clinicians and interventions that are representative of usual practice; and (c) a focus on helping patients, clinicians and policy makers to make informed choices (21). While CER has been used to investigate OMT (22, 23), such studies have not yet been subject to systematic review that specifically focusses on OMT, and so the contribution of CER to the evidence-base in osteopathy is yet to be established.

CER focuses not only on whether an intervention makes an impact under “real-world” conditions, but also on whether an intervention is beneficial in relation to the resources it consumes (18). This latter feature fits within the broad field of health economics through which policy makers attempt to manage the rise of health care expenditure by prioritizing between competing health care interventions based on value for money (24). Accordingly, health economic evaluations are crucial elements in political decision-making regarding the reimbursement and funding of health services (24). Economic evaluation of health interventions, such as osteopathic care, can be undertaken based either on clinical trial data or the modelling of data from a range of data sources (24). Whilst there have been some attempts to understand the cost-effectiveness of health services encompassing osteopathic healthcare, such as spinal manipulation (25) or manual therapy more generally (26), the findings of these studies have not yet provided firm conclusions regarding the cost effectiveness of osteopathic care as a discrete treatment option. Furthermore, cost-effectiveness studies in OMT have not previously been reviewed in relation to CER, despite the natural synergies between these approaches.
With this in mind, this paper presents the first critical systematic review of comparative
effectiveness research and health economic evaluations of OMT. The aim was to review and
critically appraise comparative effectiveness and health economic research on OMT. The
objectives were to elucidate the contribution that these research approaches can make to
the OMT evidence-base, to identify strengths and limitations of existing studies, and to make
recommendations for improving future studies using CER and health economic approaches.

MATERIALS AND METHODS

A comprehensive search of the literature was undertaken to identify published original
research examining the comparative-effectiveness and health economics of OMT. Standard
systematic review techniques were followed in accordance with the PRISMA statement (27)

Search strategy

An initial search was conducted of the following databases, from their inception to May 2015:
CINAHL, PubMed, PEDro, AMED, SCOPUS and OSTMED.DR. Two distinct searches were
undertaken in each database; one focusing on the economic evaluation of OMT, and the
other on the comparative effectiveness of OMT. Shared search terms for both searches
included osteopath*, random*, clinical trial, manipul*, manual therapy, and manual medicine.
The term comparative effectiveness was applied to the search for research articles
examining the comparative effectiveness of OMT. Likewise, cost* was used to identify
papers exploring health economics. To ensure a broad range of articles were identified,
manual searching was also conducted by reviewing references from existing review articles
located through the database search in September 2015.

Inclusion and exclusion criteria

Articles were excluded if they did not present original empirical data, were not written in
English, and did not examine OMT as a system of care. Articles were included if they
evaluated health economic outcomes of OMT or compared the effectiveness of OMT with
another available treatment or technique or standard care (including ‘no care’ where relevant
to the condition). No limits were placed on date of publication. Articles were screened, short-listed and selected for data extraction by R.R. with sample verification of identified references undertaken by A.S. throughout the screening process. Any disagreements were resolved by discussion until consensus was reached. The full literature search processes for comparative effectiveness and health economic studies are outlined in Figures 1 and 2 respectively.

**Data extraction**

Data were extracted by one investigator (T.S.) and verification of extracted data undertaken by another investigator (A.S.). Discussion was used to reach consensus in case of any disagreements. Data were extracted in accordance with the template provided by the Cochrane handbook guidelines (28) and modified for the purposes of this review to include information on methods, participants, intervention and outcomes.

**Critical analysis**

**Risk of Bias**

Comparative effectiveness articles were independently evaluated for risk of bias by two investigators (L.W., F.B.) using criteria outlined in the Cochrane Handbook for Systematic Reviews (29). The Cochrane risk of bias tool was used to assess the domains of selection bias (random sequence generation, allocation concealment), performance bias (blinding of personnel and participants), detection bias (blinding of outcome assessment), attrition bias (drop-outs), reporting bias (selective reporting of outcomes), and any other sources of bias as identified by the reviewers. Ratings were compared (73% agreement, kappa = 0.60) and differences were resolved through discussion in order to reach consensus.

**Appraisal of comparative effectiveness studies**

Articles reporting comparative effectiveness research were scored using the PRagmatic-Explanatory Continuum Indicator Summary-2 (PRECIS-2) tool (30) and independently categorised as employing either an explanatory or observational design by two investigators.
(A.S., T.S.). All papers identified through PRECIS-2 categorisation as reporting observational studies were assessed in accordance with the Good Research for Comparative Effectiveness (GRACE) principles (H.C., L.W) (31).

**Appraisal of health economic studies**

The quality of included health economics articles was assessed in accordance with the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement (32). The CHEERS statement is intended to optimise reporting guidance for economic evaluations via a checklist that is subdivided into six main categories: (1) title and abstract; (2) introduction; (3) methods; (4) results; (5) discussion; and (6) other. The checklist consists of 24 items, however one item was excluded from our analysis (i.e. item 12: “measurement and valuation of preference based outcomes”) as the item is optional and was not applicable to any of the included studies. Each paper was compared against the CHEERS checklist by two authors (A.S., J.W.) and awarded a score out of 23. Any differences between rater scores were discussed and a consensus decision made.

**Data analysis**

Given the nature and broad scope of the review, findings were summarised in narrative form.

**RESULTS**

The outcomes of the literature search for health economics analyses (n=8) (33-40) and comparative effectiveness research (n=8) (41-48) of OMT are presented in Figures 1 and 2, respectively. A total of 16 papers reported on the findings of 15 studies, with one study reporting two smaller yet distinct analyses (39, 40). Research from North America dominated the included studies with 10 papers originating from the United States (35-38, 42, 44-48) and one from Canada (41). Studies originating from the United Kingdom (33, 39, 40, 43) and Italy (34) made up the remaining five papers. The majority of included studies (n=12)
sampled adult participants, with a quarter (n=4) involving children. Sample sizes varied substantially between studies, from 29 (41) to 1556 (35) participants (mean = 276; median =90). However, when retrospective clinical audits (n=2) were excluded, the sample size range narrowed considerably (i.e. 29 to 178 participants; mean = 89; median = 58).

Headache (38, 41), neck (39, 40, 47) and back pain (35, 39, 40, 42, 43, 46) were the most common conditions examined in the included studies; other conditions included otitis media (48), spastic cerebral palsy (44, 45), pancreatitis (37) and preterm birth (34). Studies were undertaken in either community clinics (39-46, 48) or hospital environments (34-38, 47, 49).

Results of the selected studies are displayed in Table 1.

**Comparative effectiveness research**

The comparative effectiveness studies (n=8) identified in this review included both adult and paediatric populations. The characteristics of the interventions used in these studies were either OMT with manipulation or OMT in combination with another intervention (i.e. progressive muscle relaxation, or herbal medicine). OMT was compared with standard care in five studies (42, 44-47), sham OMT in two studies (46, 48), acupuncture (44, 45) in two studies, and pharmaceuticals (47) and physiotherapy (43) in one study each. Where studies included sham OMT, at least one arm of the study included access to standard care or a comparative treatment intervention. Some studies described participants as accessing no care or a wait-list control but the studies specified participants to continue accessing usual care through the study period (46) and as such these studies were considered as using ‘standard care’ as the comparator.

**Outcomes of comparative effectiveness research**

In line with the heterogeneous characteristics of the identified studies, the reported outcomes of OMT were also mixed. Research examining the effectiveness of OMT for the management of low back pain found no significant difference in benefit when compared with standard allopathic treatment (42), group exercise or physiotherapy (43), or sham manipulation (46). The outcomes of ketorolac injection for acute neck pain was comparable
to OMT in a study conducted in an emergency department, whereby both groups reported a similarly significant reduction in pain intensity (47). Significant improvement in mobility measures were reported for OMT treatment of spastic cerebral palsy when compared with wait-list control or acupuncture (45). Similarly, OMT treatment was found to reduce the occurrence of headache-free days for individuals experiencing tension headaches, but with no statistically significant difference in the intensity of the headaches, when compared with progressive muscle relaxation exercises (41).

Critical appraisal of comparative effectiveness research

The majority of the eight studies identified as comparative effectiveness studies of OMT were assessed as having a high risk of bias for blinding of participants (6 studies, 75%) and blinding of outcome assessment (4 studies, 50%), and unclear risk of bias for selective outcome reporting (6 studies, 75%) and allocation concealment (5 studies, 63%). The only domain with an overall low risk of bias across the eight studies was random sequence generation (6 studies, 75%). The two studies with the lowest overall risk of bias were characterised as having robust randomisation, well-reported allocation concealment, blinded outcome assessors, and low rates of attrition (45, 48). The study with the lowest risk of bias was undertaken by Wahl et al (48). The risk of bias assessment for all studies is reported in Table 2.

According to the PRECIS-2 scores (presented in Table 2), all comparative effectiveness studies were identified as ‘observational’ rather than ‘explanatory’ research. McReynolds et al (47) and Andersson et al (42) most closely fit the criteria for an ‘observational’ comparative effectiveness study while Liccardone (46) most closely aligned with the characteristics of an explanatory randomised-control trial study design. The study design element which most consistently supported pragmatic comparative effectiveness research was the setting of the included studies as they were all conducted in a real-life clinical environment. The factors that detracted from these studies aligning with real-life clinical practice to the degree required for pragmatic clinical research included the lack of flexibility
in the delivery of osteopathic treatments and the requirement that clinicians adhere to a
structured treatment protocol (results not shown).

Compliance with the GRACE statement checklist differed across all studies (see Table 3). Three studies (42, 46, 47) complied with 10 of the 12 checklist items and one study complied with 9 of the 12 checklist items. The lowest attributed score was for the study by Wahl et al (48), which complied with 3 of the 12 items. All studies were non-compliant with two items of the GRACE checklist: details of treatment were not adequately recorded; and meaningful analyses were not conducted to test key assumptions on which the primary results were based. Other common areas of non-compliance included failing to restrict the study population to new initiators of treatment, and overlooking important covariates or confounding variables in the study design or analysis.

Economic analysis research

Study characteristics

The included economic papers (n=8) represented a range of economic analyses, including costing studies (33, 35, 38, 40), cost-effectiveness analyses (34, 36, 37) and cost-utility analysis (39). Two papers reported a cost of care (40) and cost utility analysis (39) from the same study. The outcome measure utilised across all three cost-effectiveness studies was length of hospital stay.

Study outcomes

The identified studies reported a reduction in costs for OMT when compared with standard care for the management of neonatal preterm birth recovery (34), lumbar disc herniation-associated sciatica (49), and posterolateral postthoractomy recovery (36), but not for the management of neck or back pain. The ‘cost of treatment’ studies identified either direct cost-savings in the case of lumber disc herniation-associated sciatica (reflecting a savings of £300 per patient) (33), or a reduced cost of care when patients with low back pain (35) or migraine headache (38) were treated by an osteopathic physician as compared to standard
care. Reduction in the length of hospital stay was reported in posterolateral post-thoracotomy recovery (for patients with lung decortication only) (-6.4 days) (36), pancreatitis (-4.5 days) (37), and neonatal preterm birth recovery (-5.9 days) (34). The latter study involving preterm infants was the only study to extrapolate a cash value to this outcome, proposing a net saving of -€2724.91 per infant (34). The only cost-utility analysis identified reported improved pain and quality of life in patients with neck or back pain (39, 40) at a cost of £3760 per quality-adjusted life year (QALY) gained.

**Critical appraisal of economic analysis of osteopathic manipulative techniques**

There were substantial differences in the quality of reporting of the six included papers (from five discrete studies) evaluating the cost-effectiveness of osteopathic manual therapies when assessed against the CHEERS guidelines (see Table 4). Two papers were unable to be assessed against CHEERS guidelines as they failed to report monetary outcomes from their economic analyses (36, 37). The majority of the included papers effectively reported the background and objectives (5/6), target population and subgroups (5/6), estimation of resources and costs (5/6), and the discussion of the findings (5/6). The areas of greatest weakness across the included studies were the identification of study perspective (1/6), discount rate (1/6), and assumptions applied to the economic analysis (1/6). The highest quality reporting was found in the two papers by Williams et al (39, 40) which met 18 and 16 of the 23 CHEERS criteria, respectively; however, these still fell short in some criteria, although different weaknesses were identified in each paper. In contrast, the reporting of the economic evaluations by Burton et al (33) and Schabert and Crow (38) only met 5 and 7 of the required 23 criteria, respectively.

**DISCUSSION**

This paper represents the first systematic review and critical appraisal of comparative effectiveness and health economic research of OMT. The findings point toward an insufficient quantity and inadequate quality of comparative effectiveness and cost-effectiveness research to effectively inform OMT policy and practice. Despite positive
findings across a number of areas, including the cost-effective management of low back pain (39, 40) and preterm neonate recovery (34), the majority of studies provide incomplete data or lack sufficient rigour to be integrated into evidence-based policy decisions (24). Similarly, CER suggests OMT may be as effective as standard care for the management of low back (42) and neck (47) pain; however, studies need to be replicated in different settings and jurisdictions to verify current findings and provide the level of evidence required to inform practice change within the broader health system (18). Given the relatively high use of OMT amongst individuals with conditions such as back pain (10), as well as the use of OMT by pregnant women (9), it is paramount that the abovementioned findings be replicated to ensure that clear guidance can be given to these vulnerable populations accessing OMT.

The high level of heterogeneity across the identified studies significantly limits our ability to draw firm conclusions about the comparative effectiveness and economic value of OMT when compared with other available health services. This highlights the need for clearer guidance on the design, implementation and reporting of osteopathic research. Guidance on comparative effectiveness research, for instance, should be attentive to standard research reporting requirements (including descriptions of treatment), quality outcome measures, appropriate comparator interventions, and suitable blinding and allocation concealment procedures. Guidance on economic evaluations, on the other hand, should focus on clearly defining the study perspective, discount rate and assumptions of the analysis. These recommendations are explored in greater detail in the final section of this discussion.

The majority of included papers reported results from research conducted in the US. This is likely to impact the applicability and generalizability of the findings to other jurisdictions.

Notably, it is argued that general osteopathic practice in the US is substantially different to the rest of the world. This is because osteopathic practitioners in the US are trained as physicians before specialising in OMT; by contrast, European and Australasian osteopathic training focuses on OMT, and osteopathic training does not result in licensure to practice medicine as a physician (1). The reviewed studies focused specifically on the administration
of OMT by osteopaths in clinical settings, and as such there may be potential for
transferability of findings to countries where osteopaths practice as allied health
professionals or complementary therapists (1, 5). However, there remains a clear need to
verify the outcomes of these studies in different professional and health care contexts.

Research gaps and recommendations for future research

The critical appraisal of included papers highlights the need for increased CER and
economic analyses in OMT, as well as the replication of such studies; it also alludes to the
need to improve the quality of future OMT research to ensure findings can inform policy and
practice. We propose a number of key areas which should be considered in the design of
future studies of OMT.

Firstly, researchers need to systematically collect and report the details of OMT used in
comparative effectiveness studies. Whilst the purpose of pragmatic research is to reflect real
life practice as closely as possible, documenting and reporting the specific techniques
utilised in osteopathic research would be highly beneficial for practitioners, educators and
researchers. Not only would such detailed reporting of osteopathic interventions facilitate the
translation of research evidence into practice, education and policy, but it would also allow
inferential statistical analyses to test the potential relationships between specific techniques
and overall effectiveness of OMT. As such, future osteopathy comparative effectiveness
research would benefit from complying with intervention reporting guidelines such as the
Template for Intervention Description and Replication (TIDieR) checklist and guide (50) or
the CONSORT extension for pragmatic trials (51).

Secondly, future OMT studies should incorporate more sensitive and nuanced statistical
analyses. Despite most studies collecting data on important confounding and effect
modifying variables such as body mass index and gender, few analyses controlled for these
variables. Similarly, none of the identified CER studies reported analyses of the primary
results in an attempt to test and verify the key assumptions of the study. For example, if
regression analysis is used to assess the effectiveness of the intervention, then this
assumes treatment does not change throughout the study for any one individual and as such, their responsiveness to treatment is also consistent. Use of time-dependent regression would assist in verifying the validity of the study outcome in this case (52). Similarly, missing data, a common feature of comparative effectiveness studies, was not appropriately managed within many of these studies. Future studies need to evaluate the extent of missing data and its impact on the analysis (52).

The design and reporting of future economic evaluations of OMT can also be improved. The economic perspective of the analysis should be described and justified, i.e. the a priori decisions as to whether the ‘cost’ of the intervention will be restricted to government and third party funders or will be broadened to include the cost to patients, their families and society in general (53). Providing a clearer perspective will enable key stakeholders to make rational decisions about the allocation of scarce healthcare resources, such as the allocation of funds to support the provision of osteopathic services. Likewise, the currency, price date and conversion rate must be included in future economic studies of OMT as this impacts on the transferability of the analysis to other jurisdictions, as well as the relevance of the findings over time. These issues can be overcome by future research groups undertaking studies involving economic evaluations of osteopathy by complying with the CHEERS reporting statement (32) when reporting findings.

**CONCLUSIONS**

Comparative effectiveness and health economic studies offer valuable insights into health services that can inform evidence-based policy and practice. Despite the diverse regional presence of osteopathy and the practice of OMT throughout the world, limited research focusing upon OMT has employed either of these study methodologies to date. There is a need for researchers and the broader osteopathic community to support the advancement of rigorous and robust comparative effectiveness and health economic research that reflects osteopathic practice if this area of health care provision is to advance its role and place within health care systems around the world.
Conflicts of interest

None.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
Figure 1: PRISMA Flow chart for articles reporting comparative effectiveness of osteopathic manipulative treatment
Articles identified through database searching (n = 1,433)
Additional articles identified through other sources (n = 8)

Articles after duplicates are removed (n = 1,249)

Articles screened (n = 1,249)

Articles excluded (n = 1,303)
- Not experimental research (n=208)
- Did not report on osteopathy/osteopathic methods (n=481)
- Unrelated to topic (n=614)

Full-text articles assessed for eligibility (n = 10)

Articles included in review (n = 8)

Full-text articles excluded, with reasons (n = 2)
- Not reporting economic analysis

Figure 2: PRISMA Flow chart for articles reporting health economic analysis of osteopathic manipulative treatment
Table 1: Characteristics of included studies examining the cost or comparative effectiveness of osteopathic manipulative treatment.

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th>Year</th>
<th>Country</th>
<th>Study design and duration</th>
<th>Participants, setting and sample</th>
<th>Condition examined</th>
<th>Interventions (active and control arms)</th>
<th>Outcome measures and time-points</th>
<th>Summary of findings (effectiveness of OMT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andersson et al (42)</td>
<td>1999</td>
<td>United States</td>
<td>Randomised controlled trial (2 arms)</td>
<td>Patients (20-59 yrs) of a health maintenance organisation; n=178</td>
<td>Low back pain (&gt;3 wks but &lt;6 months duration)</td>
<td>Osteopathic manipulation OR standard allopathic treatment</td>
<td>Visual-analogue pain scale (0-100); Roland-Morris; Oswestry</td>
<td>Follow-up 12 wks: No significant difference between groups in any primary outcome measure. Osteopathic group used less medication (analgesics, anti-inflammatory, muscle relaxants) (P&lt;0.001) and less physical therapy (P&lt;0.05).</td>
</tr>
<tr>
<td>Author (Year)</td>
<td>Year</td>
<td>Country</td>
<td>Study design and duration</td>
<td>Participants, setting and sample</td>
<td>Condition examined</td>
<td>Interventions (active and control arms)</td>
<td>Outcome measures and time-points</td>
<td>Summary of findings (effectiveness of OMT)</td>
</tr>
<tr>
<td>---------------</td>
<td>------</td>
<td>---------</td>
<td>----------------------------</td>
<td>---------------------------------</td>
<td>-------------------</td>
<td>---------------------------------------</td>
<td>-----------------------------------</td>
<td>-------------------------------------------</td>
</tr>
<tr>
<td>Licciardone et al (46)</td>
<td>2002</td>
<td>United States</td>
<td>Randomised controlled trial (3 arms)</td>
<td>Patients (18-69 yrs) of a university based clinic; n=91</td>
<td>Nonspecific low back pain (&gt;3 months duration)</td>
<td>Osteopathic manipulative treatment OR sham manipulation, OR a no-intervention control group. All patients were allowed to continue their usual care for low back pain.</td>
<td>SF-36; 10-cm visual analog scale for overall back pain; Roland–Morris; Lost work or school days because of back pain; Satisfaction with back care</td>
<td>Follow-up 6 months: There were no significant benefits with osteopathic manipulative treatment, as compared with sham manipulation. Both active and sham manipulation appear to provide some benefits when used in addition to usual care for the treatment of chronic nonspecific low back pain (e.g. greater improvements in back pain and physical functioning and greater satisfaction of care).</td>
</tr>
<tr>
<td>Author (Year)</td>
<td>Year</td>
<td>Country</td>
<td>Study design and duration</td>
<td>Participants, setting and sample</td>
<td>Condition examined</td>
<td>Interventions (active and control arms)</td>
<td>Outcome measures and time-points</td>
<td>Summary of findings (effectiveness of OMT)</td>
</tr>
<tr>
<td>--------------</td>
<td>------</td>
<td>---------</td>
<td>---------------------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>----------------------------------------</td>
<td>-----------------------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Duncan et al (44)</td>
<td>2004</td>
<td>United States</td>
<td>Randomised controlled trial (3 arms); (and a co-study with 1 arm)</td>
<td>Pediatric patients (11 months to 12 yrs) at a children's clinic for rehabilitation services; n=50 (co-study n=19)</td>
<td>Spastic cerebral palsy</td>
<td>Wait-list control; OR Osteopathic manipulation; OR Acupuncture. (Co-study combined osteopathic manipulation and acupuncture.) (Intervention treatments complemented the patients' standard care.)</td>
<td>Parents were asked: “Did you note any changes in your child as a result of the therapies, and if so, what were the changes?” Parents’ perception of their child’s level of muscle stiffness and their child’s level of happiness on 2 separate visual log scales, 100 millimeters in length.</td>
<td>Follow-up 24 wks: 2 of 17 parents reported positive gains while their child was in a wait-list control period (but all 17 reported gains while in the treatment phase of the study; Twenty-one of the 23 parents of the children in the osteopathic group reported improvement in their child during the course of therapies; All of the 19 parents of children in the acupuncture group reported improvements in their child.</td>
</tr>
<tr>
<td>Author (Year)</td>
<td>Year</td>
<td>Country</td>
<td>Study design and duration</td>
<td>Participants, setting and sample</td>
<td>Condition examined</td>
<td>Interventions (active and control arms)</td>
<td>Outcome measures and time-points</td>
<td>Summary of findings (effectiveness of OMT)</td>
</tr>
<tr>
<td>----------------------</td>
<td>------</td>
<td>---------------</td>
<td>----------------------------</td>
<td>----------------------------------</td>
<td>-------------------------------------</td>
<td>----------------------------------------</td>
<td>------------------------------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td>McReynolds &amp; Sheridan (47)</td>
<td>2005</td>
<td>United States</td>
<td>Randomised controlled trial (2 arms)</td>
<td>Patients (18 to 50 yrs) in emergency department; n=58</td>
<td>Acute neck pain (&lt;3 wks duration)</td>
<td>Intra muscular injection with Ketoralac (30 ml); OR Osteopathic manipulative treatment (5 min)</td>
<td>11-point numerical rating scale for pain</td>
<td>Follow-up one hour post treatment: Both groups had significant reduction in pain intensity but there was no significant difference between the OMT and ketorolac study groups (P=.10)</td>
</tr>
<tr>
<td>Anderson &amp; Seniscal (41)</td>
<td>2006</td>
<td>Canada</td>
<td>Randomised controlled trial (2 arms)</td>
<td>Patients (&gt;16 yrs) recruited from ads and flyers; n=29</td>
<td>Tension type headache (frequent/ chronic/ probable)</td>
<td>Progressive muscular relaxation; OR Progressive muscular relaxation plus 3 osteopathic treatments</td>
<td>Headache diary</td>
<td>Follow-up 4 to 5 weeks post treatment initiation (6 to 7 wks from baseline): Number of headache free days/wk improved with osteopathy (P = .016). There was no significant difference in headache intensity (P = .264).</td>
</tr>
<tr>
<td>Author (Year)</td>
<td>Year</td>
<td>Country</td>
<td>Study design and duration</td>
<td>Participants, setting and sample</td>
<td>Condition examined</td>
<td>Interventions (active and control arms)</td>
<td>Outcome measures and time-points</td>
<td>Summary of findings (effectiveness of OMT)</td>
</tr>
<tr>
<td>--------------</td>
<td>------</td>
<td>-------------</td>
<td>----------------------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>----------------------------------------</td>
<td>----------------------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Chown et al (43)</td>
<td>2008</td>
<td>United Kingdom</td>
<td>Randomised clinical trial (3 arms)</td>
<td>Patients (18-65 yrs) referred to physiotherapy; n=239</td>
<td>Low back pain (&gt;3 months duration)</td>
<td>Group exercise physiotherapy; OR One-to-one physiotherapy; OR One-to-one Osteopathy</td>
<td>Oswestry Disability Index (ODI) was the primary outcome</td>
<td>Follow-up 6 wks: All three treatments indicated comparable reductions in mean (95% confidence intervals) ODI: group exercise, −4.5 (−0.9 to −8.0); physiotherapy, −4.1 (−1.4 to −6.9); and osteopathy, −5.0 (−1.6 to −8.4).</td>
</tr>
<tr>
<td>Author (Year)</td>
<td>Year</td>
<td>Country</td>
<td>Study design and duration</td>
<td>Participants, setting and sample</td>
<td>Condition examined</td>
<td>Interventions (active and control arms)</td>
<td>Outcome measures and time-points</td>
<td>Summary of findings (effectiveness of OMT)</td>
</tr>
<tr>
<td>---------------</td>
<td>------</td>
<td>---------</td>
<td>---------------------------</td>
<td>----------------------------------</td>
<td>--------------------</td>
<td>----------------------------------------</td>
<td>-----------------------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Duncan et al (45)</td>
<td>2008</td>
<td>United States</td>
<td>Pilot randomised controlled trial (3 arms)</td>
<td>Pediatric patients (20 months to 12 yrs) at a children’s clinic for rehabilitation services; n=55</td>
<td>Spastic cerebral palsy</td>
<td>Wait-list control; OR Osteopathic Manipulative Treatment; OR Acupuncture. (Intervention treatments complemented the patients' standard care.)</td>
<td>11 outcome variables</td>
<td>Follow-up 24 wks: statistically significant improvement in two mobility measures for patients who received OMT—the total score of Gross Motor Function Measurement and the mobility domain of Functional Independence Measure for Children (P&lt;.05). No statistically significant improvements were seen among patients in the acupuncture treatment arm.</td>
</tr>
<tr>
<td>Author (Year)</td>
<td>Year</td>
<td>Country</td>
<td>Study design and duration</td>
<td>Participants, setting and sample</td>
<td>Condition examined</td>
<td>Interventions (active and control arms)</td>
<td>Outcome measures and time-points</td>
<td>Summary of findings (effectiveness of OMT)</td>
</tr>
<tr>
<td>---------------</td>
<td>------</td>
<td>---------</td>
<td>---------------------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>----------------------------------------</td>
<td>----------------------------------</td>
<td>---------------------------------------</td>
</tr>
<tr>
<td>Wahl et al (48)</td>
<td>2008</td>
<td>United States</td>
<td>Randomised, placebo-controlled, two-by-two factorial trial (4 arms)</td>
<td>Pediatric patients (12 to 60 months); n=90</td>
<td>Recurrent otitis media (otitis-prone children)</td>
<td>True osteopathic manipulative treatment (OMT) plus placebo Echinacea; OR Echinacea plus sham OMT; OR True echinacea plus true OMT; Double placebo (placebo Echinacea plus Sham OMT)</td>
<td>Prevention of acute otitis media (risk of having at least one episode of acute otitis media during 6-month follow-up compared to placebo/sham)</td>
<td>Follow-up 6 months: No interaction was found between Echinacea purpurea and OMT. E.purpurea was associated with a borderline increased risk of having at least one episode of acute otitis media during 6-month follow-up compared to placebo (65% versus 41%; relative risk, 1.59, 95% CI 1.04, 2.42). OMT did not significantly affect risk compared to sham (44% versus 61%; relative risk, 0.72, 95% CI 0.48, 1.10).</td>
</tr>
</tbody>
</table>

Economic analyses

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th>Year</th>
<th>Country</th>
<th>Study design and duration</th>
<th>Participants, setting and sample</th>
<th>Condition examined</th>
<th>Interventions (active and control arms)</th>
<th>Outcome measures and time-points</th>
<th>Summary of findings (effectiveness of OMT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radjeski, Lumley and Cantieri (37)</td>
<td>1998</td>
<td>United States</td>
<td>Randomised-controlled trial; Hospital in-patients; n=14</td>
<td>Pancreatitis</td>
<td>Standard care</td>
<td>Length of hospital stay</td>
<td>Average length of stay reduced by 4.5 days (p=.039)</td>
<td></td>
</tr>
<tr>
<td>Author (Year)</td>
<td>Year</td>
<td>Country</td>
<td>Study design and duration</td>
<td>Participants, setting and sample</td>
<td>Condition examined</td>
<td>Interventions (active and control arms)</td>
<td>Outcome measures and time-points</td>
<td>Summary of findings (effectiveness of OMT)</td>
</tr>
<tr>
<td>---------------</td>
<td>------</td>
<td>---------</td>
<td>---------------------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>----------------------------------------</td>
<td>-----------------------------------</td>
<td>-----------------------------------------</td>
</tr>
<tr>
<td>Burton, Tillotson and Cleary (49)</td>
<td>2000</td>
<td>United Kingdom</td>
<td>Randomised-controlled trial</td>
<td>Hospital in-patients; n=30</td>
<td>Lumbar disc herniation-associated sciatica</td>
<td>Chemo-nucleolysis</td>
<td>Cost of treatment</td>
<td>Cost savings of £300 per patient with no difference in pain outcomes between groups</td>
</tr>
<tr>
<td>Williams et al (40)</td>
<td>2003</td>
<td>United Kingdom</td>
<td>Randomised-controlled trial</td>
<td>Osteopathy clinic outpatients; n=201</td>
<td>Neck or Back pain (2-12 weeks duration)</td>
<td>Standard GP care</td>
<td>Cost of treatment</td>
<td>Increased health care costs (£65) alongside improved pain and QoL measures</td>
</tr>
<tr>
<td>Williams et al (39)</td>
<td>2004</td>
<td>United Kingdom</td>
<td>Randomised-controlled trial</td>
<td>Osteopathy clinic outpatients; n=201</td>
<td>Neck or Back pain (2-12 weeks duration)</td>
<td>Standard GP care</td>
<td>Cost utility</td>
<td>£3760 per QALY gained as a result of treatment.</td>
</tr>
<tr>
<td>Author (Year)</td>
<td>Year</td>
<td>Country</td>
<td>Study design and duration</td>
<td>Participants, setting and sample</td>
<td>Condition examined</td>
<td>Interventions (active and control arms)</td>
<td>Outcome measures and time-points</td>
<td>Summary of findings (effectiveness of OMT)</td>
</tr>
<tr>
<td>-----------------------</td>
<td>------</td>
<td>-------------</td>
<td>----------------------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>---------------------------------------</td>
<td>-----------------------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Crow and Willis (35)</td>
<td>2009</td>
<td>United States</td>
<td>Retrospective clinic audit</td>
<td>Hospital patients; n=1556</td>
<td>Low back pain (&lt;6 months duration)</td>
<td>Standard care</td>
<td>Cost of treatment</td>
<td>No difference in total cost per episode of care but reduced costs for radiology (&lt;.0001) and prescription medications (&lt;.001).</td>
</tr>
<tr>
<td>Schabert and Crow (38)</td>
<td>2009</td>
<td>United States</td>
<td>Retrospective clinic audit</td>
<td>Hospital clinic outpatients; n=1427</td>
<td>Migraine headache</td>
<td>Standard care OR osteopathic care without manipulative treatment</td>
<td>Cost of treatment</td>
<td>Lower cost per office visit compared with MD but not compared with DO consults which exclude manipulative treatment</td>
</tr>
<tr>
<td>Cerritelli et al (34)</td>
<td>2013</td>
<td>Italy</td>
<td>Randomised-controlled trial</td>
<td>Preterm newborns admitted to Neonatal Intensive Care Unit; n=110</td>
<td>Neonatal preterm birth recovery</td>
<td>Standard paediatric care</td>
<td>Length of hospital stay</td>
<td>Length of stay reduced by 5.9 days with OMT (p&lt;.0001); reduced cost estimates (-€2724.91, p&lt;.0001); no impact on weight gain.</td>
</tr>
<tr>
<td>Author</td>
<td>Year</td>
<td>Country</td>
<td>Study design and duration</td>
<td>Participants, setting and sample</td>
<td>Condition examined</td>
<td>Interventions (active and control arms)</td>
<td>Outcome measures and time-points</td>
<td>Summary of findings (effectiveness of OMT)</td>
</tr>
<tr>
<td>-----------------</td>
<td>------</td>
<td>---------------</td>
<td>---------------------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>---------------------------------------</td>
<td>-------------------------------------</td>
<td>-----------------------------------------</td>
</tr>
<tr>
<td>Fleming et al (36)</td>
<td>2015</td>
<td>United States</td>
<td>Retrospective clinic audit; Hospital inpatients; n=38</td>
<td>Posterolateral posthoroactomy recovery</td>
<td>Standard care</td>
<td>Length of hospital stay</td>
<td>Length of stay reduced for patients having lung decortication (-6.4 days, ( p=0.005 )) but not others</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2: Rating of study bias using Cochrane Collaboration risk of bias tool and PRECIS-2 rating of comparative effectiveness studies**

<table>
<thead>
<tr>
<th>First Author, Year</th>
<th>Random sequence generation: selection bias</th>
<th>Allocation concealment: selection bias</th>
<th>Blinding of participants, personnel: performance bias</th>
<th>Blinding of outcome assessment: detection bias</th>
<th>Incomplete outcome data: attrition bias</th>
<th>Selective outcome reporting: reporting bias</th>
<th>Other sources of bias</th>
<th>Explanatory or Observational study (PRECIS-2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>McReynolds &amp; Sheridan (47)</td>
<td>Low</td>
<td>Unclear</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Unclear</td>
<td>Unclear</td>
<td>33</td>
</tr>
<tr>
<td>Andersson et al (42)</td>
<td>Low</td>
<td>Unclear</td>
<td>High</td>
<td>Unclear</td>
<td>Low</td>
<td>Unclear</td>
<td>Low</td>
<td>30</td>
</tr>
<tr>
<td>Study</td>
<td>Compliance</td>
<td>Compliance</td>
<td>Effectiveness</td>
<td>Effectiveness</td>
<td>Evidence Quality</td>
<td>Evidence Quality</td>
<td>Evidence Quality</td>
<td>Evidence Quality</td>
</tr>
<tr>
<td>------------------------------</td>
<td>------------</td>
<td>------------</td>
<td>---------------</td>
<td>---------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Chown et al (43)</td>
<td>Low</td>
<td>Unclear</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Unclear</td>
<td>Unclear</td>
<td></td>
</tr>
<tr>
<td>Duncan et al (44)</td>
<td>Unclear</td>
<td>Unclear</td>
<td>High</td>
<td>High</td>
<td>Unclear</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Anderson &amp; Seniscal (41)</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Unclear</td>
<td>High</td>
<td>Unclear</td>
<td></td>
</tr>
<tr>
<td>Duncan et al (45)</td>
<td>Low</td>
<td>Low</td>
<td>Unclear</td>
<td>Low</td>
<td>High</td>
<td>Unclear</td>
<td>High</td>
<td></td>
</tr>
<tr>
<td>Wahl et al (48)</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Unclear</td>
<td></td>
</tr>
<tr>
<td>Licciardone et al (46)</td>
<td>Unclear</td>
<td>Unclear</td>
<td>High</td>
<td>Unclear</td>
<td>High</td>
<td>Unclear</td>
<td>Low</td>
<td></td>
</tr>
</tbody>
</table>

*Table 3. Compliance of comparative effectiveness related osteopathic papers with the Good ReseArch for Comparative Effectiveness (GRACE) statement checklist*
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data (D1-D6)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D1. Details of treatment adequately recorded</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>D2. Primary outcomes adequately recorded</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>D3. Primary clinical outcome(s) measured objectively</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>D4. Primary outcomes validated/ adjudicated</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>D5. Primary outcome(s) measured or identified in an equivalent manner</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>n/a</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>D6. Important covariates (known confounders/effect modifiers) available</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Methods (M1-M5)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M1. Study population restricted to new initiators of treatment</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>M2. Concurrent comparators or justification of historical comparisons group</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>n/a</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>M3. Important covariates/confounding and effect modifying variables taken into account in the design and/or analysis</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------------------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>M4. Classification of exposed and unexposed person-time free of “immortal time bias”</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>M5. Meaningful analyses conducted to test key assumptions on which primary results are based</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>10</td>
<td>4</td>
<td>10</td>
<td>8</td>
<td>9</td>
<td>6</td>
<td>3</td>
</tr>
</tbody>
</table>
### Table 4: Compliance of health economic related osteopathic papers with the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement checklist.*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title and abstract</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>-</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>-</td>
</tr>
<tr>
<td>Abstract</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Introduction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Background and objectives</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Methods</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Target population and subgroups</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td>Setting and location</td>
<td>-</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Study perspective</td>
<td>-</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Comparators</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Time horizon</td>
<td>-</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td>Discount rate</td>
<td>-</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Choice of health outcomes</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td>Measurement of effectiveness</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td>Estimating resources and costs</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Currency, price date, and conversion</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td>Choice of model</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

*This table reflects the compliance of specific osteopathic health economic papers with the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement checklist. The symbols (-, X) indicate compliance or non-compliance with the respective standards.
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Assumptions</strong></td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Analytical methods</strong></td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td><strong>Results</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study parameters</strong></td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td><strong>Incremental costs and outcomes</strong></td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Characterising uncertainty</strong></td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Characterising heterogeneity</strong></td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
</tr>
<tr>
<td><strong>Discussion</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study findings, limitations, generalizability, and current knowledge</strong></td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Other</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Source of funding</strong></td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td><strong>Conflicts of interest</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td><strong>TOTAL (out of 23)</strong></td>
<td>5</td>
<td>16</td>
<td>18</td>
<td>13</td>
<td>7</td>
<td>14</td>
</tr>
</tbody>
</table>

*Appraisal using CHEERS guidelines was not undertaken for Radjieski, Lumley and Cantieri (37) and Fleming et al. (36) as they did not provide monetary outcome data in their analyses.*


## Title

<table>
<thead>
<tr>
<th>Section/topic</th>
<th>#</th>
<th>Checklist item</th>
<th>Reported on page #</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title</strong></td>
<td></td>
<td>Title</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Identify the report as a systematic review, meta-analysis, or both.</td>
<td></td>
</tr>
</tbody>
</table>

## Abstract

<table>
<thead>
<tr>
<th>Section/topic</th>
<th>#</th>
<th>Checklist item</th>
<th>Reported on page #</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Abstract</strong></td>
<td></td>
<td>Structured summary</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.</td>
<td></td>
</tr>
</tbody>
</table>

## Introduction

<table>
<thead>
<tr>
<th>Section/topic</th>
<th>#</th>
<th>Checklist item</th>
<th>Reported on page #</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Introduction</strong></td>
<td></td>
<td>Rationale</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Describe the rationale for the review in the context of what is already known.</td>
<td></td>
</tr>
<tr>
<td><strong>Objectives</strong></td>
<td>4</td>
<td>Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).</td>
<td>5</td>
</tr>
</tbody>
</table>

## Methods

<table>
<thead>
<tr>
<th>Section/topic</th>
<th>#</th>
<th>Checklist item</th>
<th>Reported on page #</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Methods</strong></td>
<td></td>
<td>Protocol and registration</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Eligibility criteria</td>
<td>5-6</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Information sources</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Search</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Study selection</td>
<td>5-6</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Data collection process</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Data items</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Risk of bias in individual studies</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Summary measures</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>State the principal summary measures (e.g., risk ratio, difference in means).</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Synthesis of results</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I²) for each meta-analysis.</td>
<td></td>
</tr>
</tbody>
</table>
## PRISMA 2009 Checklist

<table>
<thead>
<tr>
<th>Section/topic</th>
<th>#</th>
<th>Checklist item</th>
<th>Reported on page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk of bias across studies</td>
<td>15</td>
<td>Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).</td>
<td>6</td>
</tr>
<tr>
<td>Additional analyses</td>
<td>16</td>
<td>Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.</td>
<td>7</td>
</tr>
</tbody>
</table>

### RESULTS

<table>
<thead>
<tr>
<th>Section/topic</th>
<th>#</th>
<th>Checklist item</th>
<th>Reported on page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study selection</td>
<td>17</td>
<td>Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.</td>
<td>16-17</td>
</tr>
<tr>
<td>Study characteristics</td>
<td>18</td>
<td>For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.</td>
<td>18 (Table 1)</td>
</tr>
<tr>
<td>Risk of bias within studies</td>
<td>19</td>
<td>Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).</td>
<td>9</td>
</tr>
<tr>
<td>Results of individual studies</td>
<td>20</td>
<td>For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.</td>
<td>N/A</td>
</tr>
<tr>
<td>Synthesis of results</td>
<td>21</td>
<td>Present results of each meta-analysis done, including confidence intervals and measures of consistency.</td>
<td>N/A</td>
</tr>
<tr>
<td>Risk of bias across studies</td>
<td>22</td>
<td>Present results of any assessment of risk of bias across studies (see Item 15).</td>
<td>9</td>
</tr>
<tr>
<td>Additional analysis</td>
<td>23</td>
<td>Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).</td>
<td>9-10</td>
</tr>
</tbody>
</table>

### DISCUSSION

<table>
<thead>
<tr>
<th>Section/topic</th>
<th>#</th>
<th>Checklist item</th>
<th>Reported on page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary of evidence</td>
<td>24</td>
<td>Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).</td>
<td>11-12</td>
</tr>
<tr>
<td>Limitations</td>
<td>25</td>
<td>Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).</td>
<td>14</td>
</tr>
<tr>
<td>Conclusions</td>
<td>26</td>
<td>Provide a general interpretation of the results in the context of other evidence, and implications for future research.</td>
<td>14</td>
</tr>
</tbody>
</table>

### FUNDING

<table>
<thead>
<tr>
<th>Section/topic</th>
<th>#</th>
<th>Checklist item</th>
<th>Reported on page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Funding</td>
<td>27</td>
<td>Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.</td>
<td>15</td>
</tr>
</tbody>
</table>


For more information, visit: [www.prisma-statement.org](http://www.prisma-statement.org).
**Author Checklist**

Authors of all papers reporting clinical research should submit this checklist together with their manuscript and the Reporting Guideline Checklist found on the EQUATOR site (http://www.equator-network.org/).

This checklist identifies recognised guidelines for scientific reporting, which authors should use to prepare their manuscript (required for systematic reviews and original research).

<table>
<thead>
<tr>
<th>Standards of reporting</th>
<th>Guideline referred to</th>
<th>Checklist submitted</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standards of reporting</strong></td>
<td>The editors require that manuscripts adhere to recognised reporting guidelines relevant to the research design used. These identify matters that should be addressed in your paper. Please indicate which guidelines you have referred to. These are not quality assessment frameworks and your study need not meet all the criteria implied in the reporting guideline to be worthy of publication in the MATH. The checklists do identify essential matters that should be considered and reported upon. For example, a controlled trial may or may not be blinded but it is important that the paper identifies whether or not participants, clinicians and outcome assessors were aware of treatment assignments. <strong>You are also required to submit a checklist from the appropriate reporting guideline (available on the EQUATOR website (<a href="http://www.equator-network.org/">http://www.equator-network.org/</a>)) together with your paper as a guide to the editors.</strong> Reporting guidelines endorsed by MATH are listed below:</td>
<td></td>
</tr>
<tr>
<td>Statistical reporting</td>
<td>SAMPL - guidelines for statistical reporting – no checklist exists currently but authors are encouraged to view the guidelines on the EQUATOR website <a href="http://www.equator-network.org/reporting-guidelines/sampl/">http://www.equator-network.org/reporting-guidelines/sampl/</a></td>
<td></td>
</tr>
<tr>
<td>Qualitative researchers might wish to consult the guideline listed below</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Qualitative studies</td>
<td>COREQ: Consolidated criteria for reporting qualitative research (<a href="http://www.equator-network.org/reporting-guidelines/coreq/">http://www.equator-network.org/reporting-guidelines/coreq/</a>)</td>
<td></td>
</tr>
<tr>
<td>Other (please give source)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not applicable (please elaborate)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>