Osteopathic manipulative treatment: A systematic review and critical appraisal of comparative effectiveness and health economics research

Amie Steel a, b, *, Tobias Sundberg a, c, Rebecca Reid b, Lesley Ward a, d, Felicity L. Bishop a, e, Matthew Leach a, f, Holger Cramer a, g, Jon Wardle a, Jon Adams a

a Australian Research Centre in Complementary and Integrative Medicine, Faculty of Health, University of Technology Sydney, 15 Broadway, Ultimo, NSW, 2007, Australia
b Endeavour College of Natural Health, Level 2, 269 Wickham St, Fortitude Valley, QLD 4006, Australia
c Research Unit for Studies of Integrative Health Care, Karolinska Institutet (NVS/OMV), Alfred Nobels Alle 23, 141 83, Stockholm, Sweden
d Nuffield Department of Orthopaedics, Rheumatology, and Musculoskeletal Sciences (NDORMS), B4495, Oxford, OX3 7LD, University of Oxford, Oxford, United Kingdom
e Centre for Applications of Health Psychology, Faculty of Social Human and Mathematical Sciences, Building 44 Highfield Campus, University of Southampton, Southampton, SO17 1BJ, United Kingdom
f School of Nursing & Midwifery, Health Economics & Social Policy Group, University of South Australia, 101 Carie St, Adelaide, Australia
g Department of Internal and Integrative Medicine, Klinikken Essen-Mitte, Faculty of Medicine, University of Duisburg-Essen, 45141, Duisburg, Germany

ARTICLE INFO
Article history:
Received 22 July 2016
Received in revised form 25 October 2016
Accepted 27 October 2016

Keywords:
Osteopathy
Osteopathic manipulative treatment
Comparative effectiveness research
Economic evaluation
Pragmatic research

ABSTRACT
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. The purpose of this systematic review is to identify, appraise and synthesise the evidence from comparative effectiveness and economic evaluation research involving OMT. 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. 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. 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.

© 2016 Elsevier Ltd. All rights reserved.

1. 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) (Osteopathy International Alliance, 2012). 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 (Murthy et al., 2015; van Dun et al., 2013), headaches (van Dun et al., 2013; Dalla Libera et al., 2014), and neck pain (Frawley et al., 2015). 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 (Osteopathy International Alliance, 2012).
Osteopathic healthcare has diverse representation across health systems internationally, ranging from full integration within conventional health care systems (i.e. in the US where osteopathic practitioners are also trained medical doctors and are titled ‘osteopathic physicians’ (Osteopathy International Alliance, 2012), through to semi-integration as allied and complementary health care therapists (i.e. in many European and Australasian countries) (Osteopathy International Alliance, 2012). 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 (Osteopathy International Alliance, 2012).

The prevalence of osteopathic healthcare use has been reported at less than 5% of the general population in Australia (Xue et al., 2008) and as high as 16% in the United States (Licciardone, 2003). However, higher rates of use in Australia are found amongst specific populations such as the middle-aged (16%) (Sibbritt et al., 2006) and pregnant women (6.2%) (Steel et al., 2012). 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 (Ong et al., 2004). 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 (Licciardone et al., 2005), neck pain (Schwerla et al., 2013), sciatica (Franke et al., 2015), chronic obstructive pulmonary disease (Zanotti et al., 2012), irritable bowel syndrome (Florance et al., 2012; Attali et al., 2013) and various pediatric conditions (Posadzki et al., 2013). While the RCT design is considered the gold standard for demonstrating efficacy (Zaorsky and Showalter, 2012), concerns have been raised regarding the applicability of the explanatory RCT to everyday clinical practice (Maurizio, 2009). 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 (Winter and Colditz, 2014). 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” (Institute of Medicine, 2009). 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 (Sax and Goodman, 2012). While CER has been used to investigate OMT (Standaert et al., 2011; Menke, 2014), 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 (Zaorsky and Showalter, 2012). 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 (Goeree and Diaby, 2013). Accordingly, health economic evaluations are crucial elements in political decision-making regarding the reimbursement and funding of health services (Goeree and Diaby, 2013). 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 (Goeree and Diaby, 2013). Whilst there have been some attempts to understand the cost-effectiveness of health services encompassing osteopathic healthcare, such as spinal manipulation (Michaelf et al., 2012) or manual therapy more generally (Tsertsavdze et al., 2014), 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.

2. 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 (Moher et al., 2009).

2.1. 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, manip*, 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.

2.2. 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 Figs. 1 and 2 respectively.

2.3. 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.
**Fig. 1.** PRISMA Flow chart for articles reporting comparative effectiveness of osteopathic manipulative treatment.

**Fig. 2.** PRISMA Flow chart for articles reporting health economic analysis of osteopathic manipulative treatment.
Data were extracted in accordance with the template provided by the Cochrane handbook guidelines (Higgins and Green, 2008) and modified for the purposes of this review to include information on methods, participants, intervention and outcomes.

2.4. Critical analysis

2.4.1. 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 (Higgins et al., 2011). 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.

2.4.2. Appraisal of comparative effectiveness studies

Articles reporting comparative effectiveness research were scored using the PRagmatic-Explanatory Continuum Indicator Summary-2 (PRECIS-2) tool (Loudon et al., 2015) and independently categorised as employing either an explanatory or observational studies. The quality of included health economics articles was assessed in accordance with the Good Research for Comparative Effectiveness (GRACE) principles (Dreyer et al., 2010).

2.4.3. 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 (Husereau et al., 2013). The CHEERS statement is intended to optimise reporting guidance for economic evaluations via a checklist that is subdivided into six main categories: (Osteopathy International Alliance, 2012) title and abstract; (Murthy et al., 2015) introduction; (van Dun et al., 2013) methods; (Dalla Libera et al., 2014) results; (Frawley et al., 2015) discussion; and (Thomson et al., 2014) 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.

2.5. Data analysis

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

3. Results

The outcomes of the literature search for health economics analyses (n = 8) (Burton et al., 2000a; Cerritelli et al., 2013; Crow and Willis, 2009; Fleming et al., 2015; Radjieski et al., 1998; Schabert and Crow, 2009; Williams et al., 2003, 2004) and comparative effectiveness research (n = 8) (Anderson and Seniscal, 2006; Andersson et al., 1999; Chown et al., 2008; Duncan et al., 2004, 2008; Licciardone et al., 2002; McReynolds and Sheridan, 2005; Wahl et al., 2008) of OMT are presented in Figs. 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 (Williams et al., 2003, 2004). Research from North America dominated the included studies with 10 papers originating from the United States (Crow and Willis, 2009; Fleming et al., 2015; Radjieski et al., 1998; Schabert and Crow, 2009; Andersson et al., 1999; Duncan et al., 2004, 2008; Licciardone et al., 2002; McReynolds and Sheridan, 2005; Wahl et al., 2008) and one from Canada (Anderson and Seniscal, 2006). Studies originating from the United Kingdom (Burton et al., 2000a; Williams et al., 2004; Williams et al., 2003; Chown et al., 2008) and Italy (Cerritelli et al., 2013) 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 (Anderson and Seniscal, 2006) to 1536 (Crow and Willis, 2009) 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).

3.1. Comparative effectiveness research

The comparative effectiveness studies (n = 8) identified in this review included both adult and pediatric 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 (Andersson et al., 1999; Duncan et al., 2004, 2008; Licciardone et al., 2002; McReynolds and Sheridan, 2005), sham OMT in two studies (Licciardone et al., 2002; Wahl et al., 2008), acupuncture (Duncan et al., 2004, 2008) in two studies, and pharmaceuticals (McReynolds and Sheridan, 2005) and physiotherapy (Chown et al., 2008) 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 examining the effectiveness of OMT for the management of low back pain found no significant difference in benefit when compared with standard allopathic treatment (Andersson et al., 1999), group exercise or physiotherapy (Chown et al., 2008), or sham manipulation (Licciardone et al., 2002). The outcomes of ketorolac injection for acute neck pain was comparable to OMT in a study conducted in...
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. (1999)</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 (≥3 wks but &lt;6 months duration)</td>
<td>Osteopathic manipulation OR standard allopathic treatment</td>
<td>Visual-analog 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>Liciardone et al. (2002)</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>Non-specific low back pain (≥3 months duration)</td>
<td>Osteopathic manipulative treatment OR sham manipulation OR 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>Duncan et al. (2004)</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–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 mm 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. Follow-up 1 h post treatment: Both groups had significant reduction in pain intensity but there was no significant difference between the OMT and ketorolac study groups (P = 0.10).</td>
</tr>
<tr>
<td>McReynolds and Sheridan (2005)</td>
<td>2005</td>
<td>United States</td>
<td>Randomised controlled trial (2 arms)</td>
<td>Patients (18–50 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 4–5 weeks post treatment initiation (6–7 wks from baseline): Number of headache free days/wk improved with osteopathy (P = 0.016). There was no significant difference in headache intensity (P = 0.264).</td>
</tr>
<tr>
<td>Anderson and Seniscal (2006)</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 6 wks: All three treatments indicated comparable reductions in mean (95% confidence interval): 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). 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; 0.05). No statistically significant improvements were seen among patients in the acupuncture treatment arm.</td>
</tr>
<tr>
<td>Chow et al. (2008)</td>
<td>2008</td>
<td>United Kingdom</td>
<td>Randomised clinical trial (3 arms)</td>
<td>Patients (18–65 yrs) referred to one physiotherapy; n = 239</td>
<td>Low back pain (≥5 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 interval): 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). 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; 0.05). No statistically significant improvements were seen among patients in the acupuncture treatment arm.</td>
</tr>
<tr>
<td>Duncan et al. (2008)</td>
<td>2008</td>
<td>United States</td>
<td>Pilot randomised controlled trial (3 arms)</td>
<td>Pediatric patients (20 months–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; 0.05). No statistically significant improvements were seen among patients in the acupuncture treatment arm.</td>
</tr>
</tbody>
</table>

(continued on next page)
<table>
<thead>
<tr>
<th>Author et al. (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>Wahl et al. (2008)</td>
<td>2008</td>
<td>United States</td>
<td>Randomised, placebo-controlled, two-by-two factorial trial (4 arms)</td>
<td>Pediatric patients (12–60 months); n = 90</td>
<td>Recurrent otitis media (otitis-prone children)</td>
<td>True osteopathic manipulative treatment (OMT) plus placebo Echinacea; OR Echinacea plus 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>
<tr>
<td>Radjieski et al. (1998)</td>
<td>1998</td>
<td>United States</td>
<td>Randomised-controlled trial; Hospital in-patients; n = 14</td>
<td></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 = 0.039)</td>
</tr>
<tr>
<td>Burton et al. (2000b)</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. (2003)</td>
<td>2003</td>
<td>United Kingdom</td>
<td>Randomised-controlled trial; 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>
<td></td>
</tr>
<tr>
<td>Williams et al. (2004)</td>
<td>2004</td>
<td>United Kingdom</td>
<td>Randomised-controlled trial; 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>
<td></td>
</tr>
<tr>
<td>Crow and Willis (2009)</td>
<td>2009</td>
<td>United States</td>
<td>Retrospective clinic audit; 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;0.0001) and prescription medications (&lt;0.001).</td>
<td></td>
</tr>
<tr>
<td>Schubert and Crow (2009)</td>
<td>2009</td>
<td>United States</td>
<td>Retrospective clinic audit; 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>
<td></td>
</tr>
<tr>
<td>Cerritelli et al. (2013)</td>
<td>2013</td>
<td>Italy</td>
<td>Randomised-controlled trial; Neonatal Intensive Care Unit; n = 110</td>
<td>Neonatal preterm birth recovery</td>
<td>Standard pediatric care</td>
<td>Length of hospital stay</td>
<td>Length of stay reduced by 59 days with OMT (p &lt; 0.0001); reduced cost estimates (&lt;6.27 ± 1.91, p &lt; 0.0001); no impact on weight gain.</td>
<td></td>
</tr>
<tr>
<td>Fleming et al. (2015)</td>
<td>2015</td>
<td>United States</td>
<td>Retrospective clinic audit; Hospital inpatients; n = 38</td>
<td>Posterolateral thoracotomy recovery</td>
<td>Standard care</td>
<td>Length of hospital stay</td>
<td>Length of stay reduced for patients having lung decompensation (&lt;4.4 days, p = 0.005) but not others</td>
<td></td>
</tr>
</tbody>
</table>
an emergency department, whereby both groups reported a similarly significant reduction in pain intensity (McReynolds and Sheridan, 2005). Significant improvement in mobility measures were reported for OMT treatment of spastic cerebral palsy when compared with wait-list control or acupuncture (Duncan et al., 2008). 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 (Anderson and Seniscal, 2006).

3.1.2. 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 binding of participants (6 studies, 75%) and binding 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 (Duncan et al., 2008; Wahl et al., 2008). The study with the lowest risk of bias was undertaken by Wahl et al. (2008). 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. (McReynolds and Sheridan, 2005) and Andersson et al. (1999) most closely fit the criteria for an ‘observational’ comparative effectiveness study while Licciardone (Licciardone et al., 2002) 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 (Andersson et al., 1999; Licciardone et al., 2002; McReynolds and Sheridan, 2005) 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. (2008), 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 initiators of treatment, and overlooking important covariates or confounding variables in the study design or analysis.

3.2. Economic analysis research

3.2.1. Study characteristics

The included economic papers (n = 8) represented a range of economic analyses, including costing studies (Burton et al., 2000a; Crow and Willis, 2009; Schabert and Crow, 2009; Williams et al., 2003), cost-effectiveness analyses (Cerritelli et al., 2013; Fleming et al., 2015; Radijeski et al., 1998) and cost-utility analysis (Williams et al., 2004). Two papers reported a cost of care (Williams et al., 2003) and cost utility analysis (Williams et al., 2004) from the same study. The outcome measure utilised across all three cost-effectiveness studies was length of hospital stay.

3.2.2. 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 (Cerritelli et al., 2013), lumbar disc herniation-associated sciatica (Burton et al., 2000b), and posterolateral posthoractomy recovery (Fleming et al., 2015), but not for the management of neck or back pain. The ‘cost of treatment’ studies identified either direct cost-savings in the case of lumbar disc herniation-associated sciatica (reflecting a savings of £300 per patient) (Burton et al., 2000a), or a reduced cost of care when patients with low back pain (Crow and Willis, 2009) or migraine headache (Schabert and Crow, 2009) were treated by an

Table 2

<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 and Sheridan (2005)</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. (1999)</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>Chown et al. (2008)</td>
<td>Low</td>
<td>Unclear</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Unclear</td>
<td>Unclear</td>
<td>28</td>
</tr>
<tr>
<td>Duncan et al. (2004)</td>
<td>Unclear</td>
<td>Unclear</td>
<td>High</td>
<td>High</td>
<td>Unclear</td>
<td>High</td>
<td>High</td>
<td>28</td>
</tr>
<tr>
<td>Anderson and Seniscal (2006)</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Unclear</td>
<td>High</td>
<td>Unclear</td>
<td>26</td>
</tr>
<tr>
<td>Duncan et al. (2008)</td>
<td>Low</td>
<td>Low</td>
<td>Unclear</td>
<td>Low</td>
<td>High</td>
<td>Unclear</td>
<td>High</td>
<td>25</td>
</tr>
<tr>
<td>Wahl et al. (2008)</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Unclear</td>
<td>25</td>
</tr>
<tr>
<td>Licciardone et al. (2002)</td>
<td>Unclear</td>
<td>Unclear</td>
<td>High</td>
<td>High</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Low</td>
<td>20</td>
</tr>
</tbody>
</table>
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) (Fleming et al., 2015), pancreatitis (-4.5 days) (Radjeski et al., 1998), and neonatal preterm birth recovery (-5.9 days) (Cerritelli et al., 2013). 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 (Cerritelli et al., 2013). The only cost-utility analysis identified reported improved pain and quality of life in patients with neck or back pain (Williams et al., 2003, 2004) at a cost of £3760 per quality-adjusted life year (QALY) gained.

### 3.2.3. 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 (Fleming et al., 2015; Radjeski et al., 1998). 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. (2003, 2004) 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. (2000a) and Schabert and Crow (2009) only met 5 and 7 of the required 23 criteria, respectively.

### 4. 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 (Williams et al., 2003, 2004) and preterm neonate recovery (Cerritelli et al., 2013), the majority of studies provide incomplete data or lack sufficient rigour to be integrated into evidence-based policy decisions (Goeree and Diaby, 2013). Similarly, CER suggests OMT may be as effective as standard care for the management of low back (Andersson et al., 1999) and neck (McReynolds and Sheridan, 2005) 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 (Zaorsky and Showalter, 2012). Given the relatively high use of OMT amongst individuals with conditions such as back pain (Ong et al., 2004), as well as the use of OMT by pregnant women (Steel et al., 2012), it is paramount that the above-mentioned 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 binding 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 (Osteopathy International Alliance, 2012). 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 (Osteopathy International Alliance, 2012; Thomson et al., 2014). However, there remains a clear need to verify the outcomes of these studies in different professional and health care contexts.

4.1. 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 (Zwarenstein et al., 2008) or the CONSORT extension for pragmatic trials (Zwarenstein et al., 2008).

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 (Velentgas et al., 2013). 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 (Velentgas et al., 2013).

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 (Drummond et al., 2008). 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 data 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 researchers undertaking studies involving economic evaluations of osteopathy by complying with the CHEERS reporting statement (Husereau et al., 2013) when reporting findings.

5. 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.

References


