Supplementary Methods and Results

Methods

Study selection

We considered all the main and secondary outcomes as described by the original reviews. If a review presented distinct associations on more than one eligible outcome or type of clinical setting, we considered those separately. From this umbrella review, we excluded narrative reviews, letters to the editor, and SRs or MAs examined pain conditions such as low back pain due to cancer, infection, inflammatory arthropathy, osteoporosis, high-velocity trauma, or fracture, low back pain during pregnancy, rheumatoid arthritis or rheumatic pain. Reviews published in other languages than English were also excluded.

For the qualitative synthesis of data, appraisal of methodological quality, and data extraction information from all eligible SRs and MAs was used, regardless of the included studies' design or multiple publication statuses. However, for the quantitative synthesis, only SRs with quantitative synthesis or MAs of randomised controlled trials (RCTs) and controlled clinical trials (CCTs) with at least two primary studies were considered eligible. For this analysis, studied outcomes in each association were categorised into short-term, medium-term, and long-term as defined by the original authors (4). If a definition was not given, then outcomes measured at post-treatment and closest to three months were considered short-term follow-up, above three months and closest to six months were considered medium-term follow-up, and outcomes measured above six months were considered long-term follow-up. When a long-term follow up outcome was presented in multiple times (e.g., both 12 and 24 months), all the time points were assessed separately.

Results

Description of Meta-analytic Associations

Specifically, 98 (73.1%) associations provided evidence for chronic LBP, 19 (14.2%) for subacute LBP alone, six (4.5%) for chronic SP, four (3.0%) for prevention of LBP, three (2.2%) for FMS (including WSP). Of the remaining four associations, two (1.5%) considered subacute and chronic LBP and two (1.5%) chronic NP alone. None of the associations provided evidence regarding MMRP for WAD. Forty-seven (35.1%) associations provided evidence for short-term outcomes, 31 (23.1%) for medium-term and 56 (41.8%) for long-term outcomes (Supplementary Table VI[1]). The most examined outcomes were pain ($n=40$; 29.9%), work ($n=34$; 25.4%) and disability/functional status ($n=31$; 23.1%).

Summary Effect Sizes

The ES of the largest study in each meta-analysis was more conservative than the random-effects summary estimate in 71 (52.9%) associations (Fig. 3) and in 16 (11.9%) associations the ESs of the largest study were both nominally significant and more conservative than the random-effects summary estimate.

Between-Study Heterogeneity and Excess of Significant Findings

These associations of high heterogeneity were pertained to medium-term anxiety ($n=1$), depression ($n=3$), disability/functional status ($n=19$), pain (n=20), and work (n= 14) in all time periods, short-term fear avoidance ($n=1$), short and long-term quality of life ($n=3$) short-term self-efficacy ($n=1$) (2, 4, 6, 17, 21-23, 54, 55, 59). These associations with an excess of significant findings pertained to short-term depression ($n=1$), disability/functional status ($n=9$) in all time periods, pain ($n=8$) in all time periods, long-term quality of life ($n=1$), short-term self-efficacy ($n=1$) and work ($n=7$) in all time periods (4, 6, 17, 21-23, 55).

Descriptive Analysis of Qualitative SRs

Sixteen (66.7%) reviews included solely RCTs, while 8 (33.3%) included non-RCTs. The median number of participants was 972 (IQR = 252– 2106) and the total number of participants was >1000 in 12 (50.0%) reviews. Half of the reviews ($n=12$) provided evidence for chronic LBP and both subacute and chronic LBP. The median number of assessed outcomes was 5 (IQR = 3–7). More than half of these reviews ($n=14$; 58.3%) considered a grouping of outcomes with respect to all three physical, mental and social health outcomes, but without a clear distinction between primary and secondary. All the 24 reviews (100%) examined the physical functioning (including disability and work outcomes) as an outcome and 22 reviews (83.3%) assessed pain outcomes (Table III). Fifteen (62.5%) reviews provided evidence for short-term outcomes and long-term outcomes.