Do post-traumatic pain and post-traumatic stress symptomatology mutually maintain each other?

A systematic review of cross-lagged studies

Sophie Lykkegaard Ravn (MSc)\textsuperscript{1,3}, Jan Hartvigsen (DC, PhD)\textsuperscript{4,5}, Maj Hansen (MSc, PhD)\textsuperscript{1}, Michele Sterling (MPhty, PhD)\textsuperscript{6}, and Tonny Elmose Andersen (MSc, PhD)\textsuperscript{2}

\textsuperscript{1} ThRIVE, Department of Psychology, University of Southern Denmark, Odense, Denmark

\textsuperscript{2} InCoRE, Department of Psychology, University of Southern Denmark, Odense, Denmark

\textsuperscript{3} The Specialized Hospital for Polio and Accident Victims, Roedovre, Denmark

\textsuperscript{4} Department of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark

\textsuperscript{5} Nordic Institute of Chiropractic and Clinical Biomechanics, Odense, Denmark.

\textsuperscript{6} Recover Injury Research Centre, NHMRC Centre of Research Excellence in Recovery Following Road Traffic Injuries, University of Queensland, Australia

Corresponding author: Sophie Lykkegaard Ravn, Department of Psychology, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark; slravn@health.sdu.dk

Copyright © 2018 by the International Association for the Study of Pain. Unauthorized reproduction of this article is prohibited.
Abstract

Following traumatic exposure, individuals are at risk of developing symptoms of both pain and post-traumatic stress disorder (PTSD). Theory and research suggest a complex and potentially mutually maintaining relationship between these symptomatologies. However, findings are inconsistent and the applied methods are not always well suited for testing mutual maintenance. Cross-lagged designs can provide valuable insights into such temporal associations, but there is a need of a systematic review to assist clinicians and researchers in understanding the nature of the relationship. Thus, the aim of this systematic review was to identify, critically appraise, and synthesize results from cross-lagged studies on pain and PTSD symptomatology in order to assess the evidence for longitudinal reciprocity and potential mediators. Systematic searches resulted in seven eligible studies that were deemed of acceptable quality with moderate risk of bias using the cohort study checklist from Scottish Intercollegiate Guidelines Network. Further, synthesis of significant pathways in the cross-lagged models showed inconsistent evidence of both bidirectional and unidirectional interaction patterns between pain and PTSD symptomatology across time, hence not uniformly supporting the theoretical framework of mutual maintenance. Additionally, the synthesis suggested that hyperarousal and intrusion symptoms may be of particular importance in these cross-lagged relationships, while there was inconclusive evidence of catastrophizing as a mediator. In conclusion, the findings suggest an entangled, but not necessarily mutually maintaining relationship between pain and PTSD symptomatology. However, major variations in findings and methodologies complicated synthesis, prompting careful interpretation and heightening the likelihood that future high quality studies will change these conclusions.

Keywords: pain; posttraumatic stress; PTSD; systematic review; autoregressive cross-lagged panel models; SEM
Introduction

In recent years, there has been increased focus on the co-existence of pain and post-traumatic stress disorder (PTSD) symptomatology [40] as well as on the nature of their potentially interacting relationship [6,9]. According to DSM-IV [1], PTSD is a maladaptive reaction to traumatic exposure comprised of three symptom clusters, i.e., intrusion, avoidance, and hyperarousal, whereas PTSD as defined by DSM-5 [2] also includes a fourth cluster of negative alterations in cognitions and mood. Thus, the DSM-5 criteria constitute a more inclusive and heterogeneous condition than the DSM-IV [22]. Although, PTSD is not the only type of post-traumatic response, it is highly prevalent following traumatic exposure [32,33]. According to DSM-5 [2], traumatic exposure can be defined as exposure to actual or threatened death or serious injury, which a motor vehicle crash (MVC) or similar traumatic incidents can be examples of. Additionally, DSM-IV [1] also defined the exposure to incidents involving threat of physical integrity as a potential trauma (Criterion A1), while there was also a demand for a response of intense fear, helplessness, or horror (Criterion A2). Similarly, persistent pain is also common after a variety of traumatic injuries and events [8,14,27,41,47], making both post-traumatic stress symptoms and post-traumatic pain common after traumatic incidents.

Generally, studies report high rates of simultaneous pain and PTSD symptomatology both in pain populations and PTSD populations [40]. Indeed, a recent meta-analysis of studies assessing PTSD symptomatology in pain samples reported a pooled mean prevalence of self-reported PTSD of 20.4% [52], indicating that a significant portion of patients with pain report clinically relevant PTSD symptomatology. PTSD symptomatology is also associated with increased levels of pain, pain-related disability, and psychological distress across pain populations [23,44,46,48,63]. Similarly, early levels of PTSD have been found to predict later pain and disability [12,27,29,34], and peritraumatic pain has also been found to be a risk factor for later PTSD symptoms [24,43]. Indeed, this potential reciprocity of pain and PTSD symptomatology has been suggested in theoretical frameworks. Sharp and Harvey [51]
suggested that the two conditions mutually maintain one another through an array of cognitive, emotional, behavioural, and physiological factors. Similarly, Liedl and Knaevelsrud [35] suggested that symptoms of pain and PTSD affect one another early after trauma and that symptoms of avoidance and hyperarousal directly interact with the pain experience. Asmundson and colleagues [4], on the other hand, suggested that the two conditions might share vulnerability factors such as anxiety sensitivity that could make an individual more prone to develop both conditions post-trauma, while a recent systematic review also concluded that the two conditions share a number of neurobiological pathways that may also explain their interrelations [49].

The view of mutual maintenance between symptoms of pain and PTSD has been widely applied and is also supported in recent non-systematic literature reviews [6,9]. However, the majority of studies that have investigated the relationship between pain and PTSD were not designed to test for associations over time that could be interpreted as indicative of mutual maintenance. Instead, studies were primarily cross-sectional or only partially tested the suggested relationship, e.g., PTSD symptoms as a predictor of later pain or vice versa. Additionally, such associational patterns can in fact not be used to investigate mutual maintenance per se, as this is referring to a more complex set of processes than the mere testing of associations between pain and PTSD symptoms. This is complicated further by potential problems with separating symptoms of pain and PTSD, making assessments of PTSD symptoms within pain populations very likely to be inflated. Taken together, this problematizes the interpretations that can be made of studies claiming to present evidence of mutual maintenance between pain and PTSD symptoms. Despite of all of this, the view of mutual maintenance has often uncritically been accepted and applied throughout the field. Regardless of this critique, bidirectional relationships and temporal precedence between pain and PTSD symptomatology are still important to better understand the true nature of the longitudinal relationship between the two constructs post-trauma. Such associations can be tested using auto-regressive cross-lagged models, which allow for complicated testing over time in a
single model. The “autoregressive” component means that the construct is regressed on earlier measures of itself, hence capturing and controlling for the stability of the construct itself, while the “cross-lagged” component means that a construct is predicted by earlier measures of one or several other constructs [30,50]. Such models are also known as cross-lagged panel analyses or cross-lagged path analyses and have recently been used to a wider extent to assess the reciprocal relationship of pain and PTSD symptomatology. However, no study has reviewed and synthesized the results of these studies, which will help shed light on the nature of the relationship between pain and PTSD symptomatology and potentially aid and affect the work of clinicians and researchers.

The aim of the present systematic review was to systematically identify, critically appraise, and synthesize research investigating the reciprocal associations between post-traumatic pain and PTSD symptomatology using cross-lagged panel models or the equivalent as well as potential mediators in these relationships.

Methods

This work was conducted and reported using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to create a best-evidence synthesis (the PRISMA checklist is available online as supplemental digital content at http://links.lww.com/PAIN/A616). A protocol of the study was registered in the International Prospective Register of Systematic Reviews (PROSPERO) on July 1th, 2017 (no. CRD42017071607).

Search strategy

The databases of PsycINFO, PubMed, Web of Science, EMBASE, Scopus, and PILOTS were searched for eligible studies on February 21, 2018. A thoroughly and broad search string was developed with search words combined using the Boolean Logic operators (AND and OR). During the
preliminary work, pilot searches were performed to ensure inclusiveness and relevance of search terms and to identify relevant subject headings. The same search strategy was performed in all databases.

(PTSD OR PTSS OR PTS OR PSS OR ASD OR post-traumatic stress*
OR post-traumatic stress* OR post traumatic stress* OR acute stress*)
AND
Pain
AND
(longitudinal* OR prospective* OR cohort* OR observation* OR path analys* OR cross-lagged* OR structural equation mod* OR SEM)

All search words were checked for being registered either as subject headings or mesh terms. If this was the case, the search word was both included as this with auto explosion and as a free text word. Reference lists of eligible studies were also screened for additional references to ensure exhaustiveness.

Eligibility criteria

Studies had to fulfil the following eligibility criteria to be included in the present systematic review:

1) peer-reviewed articles,
2) written in English, Danish, Swedish, or Norwegian,
3) using an observational, prospective design with ≥ three measurement points after a traumatic event, which was not required to explicitly state fulfilment of Criterion A,
4) on adult samples ≥16 years of age, and
5) applying autoregressive cross-lagged panel analysis or the equivalent.

Studies were not included if they:

1) were intervention studies, reviews, dissertations, letters, editorials, book chapters, qualitative studies, cross-sectional studies, longitudinal studies not using a relevant design, and conference abstracts,
2) did not include specific measurements of both pain intensity/severity and PTSD symptomatology,
3) included samples where the index trauma was pregnancy and/or birth, or
4) included participants with severe head injuries (<9 on the Glasgow Coma Scale), severe neurological diseases, and/or severe psychiatric comorbidities.

Data screening

First, all references were transferred to EndNote (version X8) and duplicates were removed. Both the removed and remaining references were double-checked manually. The remaining unique studies were then transferred to the online tool of Covidence, where the first and last author screened titles and abstracts independently according to the eligibility criteria. In cases of doubt or disagreement, the studies were carried forward for full-text screening. Here, disagreement between the raters was discussed among the raters themselves and with the inclusion of the second author, if needed. At this step, excluded papers and reasons for exclusion were logged into Covidence.
Data extraction

Two authors (first and last author) independently extracted data from the included studies into a table for publication (Table 1). Again, discrepancies among the raters were discussed and corrected accordingly.

Assessment of methodological quality and risk of bias

The overall quality of the included studies was rated based on assessment of the methodological quality and the risk of bias. For this purpose, the checklist for cohort studies from Scottish Intercollegiate Guidelines Network (SIGN; available at http://www.sign.ac.uk/checklists-and-notes.html) was used with some modifications to fit our purpose. The modifications included deletion of some of the original items that were not applicable for the types of studies included in the present study, which included original items 1.2, 1.3, 1.8, and 1.9, as well as item 1.12 (an exclusion criterion). Additionally, two items regarding statistical assumptions and model fit were added, and the demonstration of validity and reliability of outcome assessments (original item 1.11) were split as two independent items based on the recommendation of Cancelliere et al. [13] . For every item, it was assessed whether the study in question appropriately did what was asked with a statement of “yes”, “no”, “can’t say”, and sometimes “does not apply”. Ratings were made independently by the first and last authors and afterwards compared and concurred. The second author was included in this process to ensure accuracy. Based on completed checklists and the discussions among the raters, each study was then rated as having little or no risk of bias, moderate risk of bias, or high risk of bias.
Data synthesis

First, descriptive characteristics of included studies were tabulated. In-depth synthesis of the findings in the studies was then carried out in two steps. First, studies using PTSD symptomatology as one total severity variable and studies using PTSD symptom clusters were considered separately. Next, findings from the studies investigating mediators of the cross-lagged relationship between pain and PTSD symptomatology were synthesized. Due to differences the specific measurement time-points, studies were described in detail on this count in each synthesis to enhance visibility. For this purpose, significant cross-lagged associations were assessed based on the reported \( p \)-values. When both cross-lagged coefficients are significant (\( p<0.05 \)), this is indicative of a bidirectional (i.e. mutual) maintenance pattern, while only one cross-lagged coefficient significant is indicative of a unidirectional maintenance pattern, and no significant cross-lagged coefficients means no maintenance pattern at all (Figure 1).

Insert Figure 1 here.

Results

Identification of studies

Based on our search terms, a total of 7,164 studies were identified across the databases. 2,838 duplicates were removed, leaving 4,326 unique publications for screening. Of these, 97 were found eligible for full-text screening. A total of seven studies matched our eligible criteria, while the other 90 were excluded primarily based on study design. Two of the seven studies were based on the same study sample, as Carty et al. [15] used a subsample of the full sample used by Liedl et al. [36], which was confirmed by contacting the author group. Screenings of reference lists of the seven included studies did not reveal additional studies (Figure 2).
Descriptive characteristics

The seven eligible studies included a total of 2,773 unique participants, counting patients with severe injury without severe traumatic brain injury [15,28,36], patients with minor injury post MVC [20,45], blast-exposed military personnel [54], and burn victims [58]. Around one third of the participants were females (35.92% across studies (range: 3.35-66.40%)). Mean age ranged from 27.4 to 40.9 years in the six studies that reported this with a cross-study mean of 37.7 years. Descriptive characteristics for each of the seven studies and their samples are presented in Table 1.

Risk of bias assessments

All studies were appraised to be of acceptable methodological quality with moderate risk of bias. Of note, this category contains great variability, as the amount and types of methodological issues vary from study to study, which is visualized in Table 2.
of studies also had potential attrition bias with a significant subset of the sample dropping out over time [15,28,36,54,58] with only Feinberg et al.’s [20] dropout rate not exceeding the recommended 20% and Ravn et al. [45] not reporting dropout rates at all. Four studies applied dropout analyses [15,36,54,58]. All included studies had clearly defined outcomes and used validated assessment tools of both PTSD and pain. Despite this, however, the outcomes were not consistently focused on trauma related symptoms and the studies often failed to refer to other studies assessing the psychometric properties of the scales, introducing potential detection biases. Additionally, there may be a potential bias associated with self-report of traumatic exposure [45,54] compared to studies sampling from hospitals [15,20,28,36,58], especially when there is a significant time gap between trauma and baseline assessment [54], introducing an additional risk of recall bias. Relatedly, none of the studies were explicitly clear about endorsement of criteria A1 and A2, something particularly important in studies when assessing PTSD symptomatology in samples exposed to objectively minor events [20,45]. Further, there may exist potential confounding related concerns. Even though all studies included the risk of confounding to some degree in designing the study and discussing the results, this was often only briefly touched upon. Also, only two studies statistically controlled for demographics such as age and gender [45,54] and one controlled for catastrophizing by including it as a potential mediator [15]. In addition, only Van Loey et al. [58] provided confidence intervals for the path coefficients, and Feinberg et al. [20] failed to report model fit indices, which is an issue in terms of assessing the legitimacy of the model. Finally, there was a general lack of commenting on the statistical assumptions, leaving the reader unable to judge potential biases related to this, with one study violating the assumption of stationarity by using different assessments of PTSD symptomatology at different time points [28].
Synthesis of association patterns

Six studies investigated the cross-lagged relationship between pain and PTSD symptomatology, while the seventh study used fixed paths limited to associations with catastrophizing [58], hence only illuminating the relationship between pain and PTSD symptomatology through catastrophizing and not directly. Hence, only six studies are relevant for this section.

Four of the six studies investigated PTSD symptomatology as a total severity score. Of these, three reported evidence of bidirectional associations between pain and PTSD symptoms from T1 to T2 [15,28,54], while this changed to unidirectional patterns from T2 to T3, either from pain to PTSD symptoms [15] or from PTSD symptoms to pain [28,54]. However, assessment points varied between studies with T2 being respectively three [15] and six months [28] post-injury, while the third study by Stratton et al. [54] also had T2 at six months post-baseline, but instead had a significantly longer mean period between trauma and baseline (mean 552 days). The fourth study found that only PTSD symptoms predicted pain from T1 (<four weeks post-injury) to three months post-injury and again from six to twelve months post-injury, while no relations were found from three to six months post-injury [45].

The remaining two of the six studies investigated PTSD symptom clusters (intrusion, hyperarousal, and avoidance) and both found evidence of bidirectional associations between hyperarousal and pain in the early months post-trauma from T1 (less than six weeks post-trauma) to T2 (either three and six months post-trauma) and bidirectional associations between intrusion and pain in the chronic months post-trauma from T2 (either three or six months post-trauma) to T3 (twelve months post-trauma) [20,36]. Additionally, Liedl et al. [36] also found evidence of bidirectional associations between hyperarousal and pain from T2 to T3. Further, a number of unidirectional effects were found with intrusion on pain [36] and pain on intrusion [20] found in early months post-trauma (from T1 to T2), while pain on avoidance [36] and pain on hyperarousal [20] were found while in the chronic months post-trauma (from T2 to T3).
Synthesis of evidence of mediators of association patterns

Two studies tested catastrophizing as a mediator in the models [15,58]. Carty et al. [15] found no evidence that catastrophizing was a mediator in the cross-lagged relationship between pain and PTSD symptoms, while Loey et al. [58] found that PTSD symptoms at T1 predicted catastrophizing at T2, which then predicted pain at T3, indicating a mediating role of catastrophizing between initial PTSD symptoms and persistent pain at 12 months.

Discussion

The present systematic review identified seven eligible studies, which were appraised to be of acceptable methodological quality with a moderate risk of bias related possible performance, attrition, and detection biases as well as issues related to confounding and statistics. In synthesizing the findings of these studies, the present review found mixed evidence of both bidirectional and unidirectional associations between PTSD symptomatology and pain over time. Further, the synthesis highlighted the importance of hyperarousal and intrusion symptoms in the cross-lagged relationship between pain and PTSD symptomatology, while there was inconclusive evidence of catastrophizing as a mediator between pain and PTSD symptomatology. In addition to the inconsistent findings across studies, the heterogeneity in study methodologies and the moderate risk of bias across all studies complicated synthesis. Hence, future high quality studies may change these conclusions.

As our results did not uniformly confirm bidirectional association patterns between pain and PTSD symptomatology over time, which were used as indicative of potential mutual maintenance, our results only partly support the applied theoretical framework of mutual maintenance [35,51] and the conclusions of existing non-systematic reviews [6,9]. However, the great variability in individual study findings obscures straightforward conclusions, for which there may be several contributing factors. One reason for cross-study discrepancies may be differences in trauma types or injury severities, e.g. minor
versus severe traumas, maybe causing some sample types to display a more interconnected relationship between pain and PTSD symptomatology. However, we were not able to find any indication of such a pattern in the present review. Additionally, differences in findings may be due to the different designs, as comparing findings from for example early post trauma to 3 and 6 months, respectively, creates some important concerns, as both pain and PTSD symptomatology are fluctuating in nature [7,57]. It may also be that there are certain time-determined differences in interactional patterns, while only very tentative patterns of this were found. Together, this may add further to the complexity in testing these cross-lagged relationships. Further, the use of different assessment tools across studies may also capture both PTSD and pain symptomatology differently, thereby indirectly affecting the relationships tested. Indeed, a recent study showed that even very small changes in the wording in PTSD questionnaires changed the level of specific PTSD symptoms in chronic pain patients [26], highlighting that even minor changes may change the interpretation and perhaps taps differently into the pain symptoms of the respondent. Additionally, a part of the explanation may also be that the conditions influence each other indirectly through processes not captured by the present review such as for example elevated levels of (pain-related) distress [59,60]. Finally, in terms of discussing the overall applicability of the theoretical viewpoint of mutual maintenance, it is important to note that the theory is likely to better apply in selected clinical samples with high levels of pain and PTSD, as it is possible that the reciprocity between the two constructs may be diluted when tested in more broad cohorts with varying (and generally very low) symptom levels.

Only two studies in the present review examined mediators in the cross-lagged models, both investigating the role of catastrophizing with divergent results [15,58]. This difference may be due to design and statistical differences among the studies. The studies’ second outcome assessment was at respectively three [15] and six [58] months post-trauma, while their statistical approach also was different, as Loey et al. [58] did not include cross-lagged paths between pain and PTSD symptomatology.
Similarly, only two studies examined the role of the individual PTSD symptom clusters [20,36], highlighting the importance of primarily hyperarousal and intrusion symptoms with both unidirectional and bidirectional effects over all time points. Avoidance symptoms were, on the other hand, only found to be of relevance at a single time point in one study [36], suggesting that avoidance behaviours are not central in the reciprocity of PTSD symptoms and pain. Of note, the DSM-IV avoidance symptom cluster, as used here, also contains numbing symptoms [1]. Overall, studies on the relationships between pain and PTSD symptomatology have highlighted the centrality of especially hyperarousal symptoms [12,17,31,37,39,53] and to a lesser degree intrusion [55]. Additionally, the importance of both clusters is highlighted in the theoretical perspectives of mutual maintenance [35,51]. The importance of hyperarousal could rely on the tendency to catastrophic misperceptions and negative interpretations as well as anticipations of somatic sensations [31,35,51], which would therefore be predictive of pain, while intrusion is suggested to trigger pain and vice versa [51]. It is, however, important to be critical in the interpretation of the findings regarding hyperarousal, as the finding that hyperarousal symptoms have a reciprocal relationship with pain may stem from the fact that hyperarousal symptoms are simply reflecting the pain experience itself. Research of the latent structure of PTSD has suggested that the hyperarousal clusters consist of both so-called anxious arousal and dysphoric arousal [3] with the latter being more related to general distress and potentially pain-related symptomatology.

Several critical issues were identified in the risk of bias assessments with a few meriting further attention. Firstly, one concern is related to the fact that none of the studies assessed pre-injury symptomatology of pain and PTSD, which is likely to affect post-injury ratings and thereby cause skewed results. Secondly, the measurements of pain and PTSD were not consistently focused on a specific trauma exposure. Only Feinberg et al. [20] and Jenewein et al. [28] explicitly stated that the pain assessments were asking for accident-related pain, and only Ravn et al. [45] and Van Loey et al. [58] explicitly stated that the PTSD assessment were concerning MVC-related or burn-related PTSD symptoms, respectively.
This forms a particular issue in the present review, as the relationship and relative influence of pain and PTSD symptomatology on each other may change heavily depending on whether or not the same trauma caused both conditions, hence undermining the interpretations that can be drawn. Thirdly, there exists a potential validity issue of assessing PTSD symptomatology in participants with persistent pain, which stems from the fact that many PTSD symptoms included in the DSM criteria are not unique to this diagnosis [38]. As such, PTSD responses may be inflated by pain-related symptomatology, thereby increasing the risk of false positives. At the same time, a number of other psychological conditions such as depression and anxiety, which are both very common in chronic pain samples [16], can add an additional risk of false positive answers. Specifically for the purpose of assessing PTSD symptomatology, studies using clinician-administered interviews consistently [15,36] must be regarded of higher quality. A related validity concern is that the studies were generally not explicit about endorsement of criteria A1 and A2. While the A2 criterion has been removed from the DSM-5, as it did not add to the predictive nor diagnostic value of PTSD [11,21], the potential lack of fulfilment of Criterion A1 is something potentially very problematic. Particularly, this forms an potential issue in studies assessing PTSD symptomatology in samples experiencing minor injuries [20,45], as it is more likely that such objectively minor injuries and incidents may not fulfil the DSM-IV criteria on threat of death, serious injury, or physical integrity [1]. As the Criterion A1 is an important part of the diagnostic criteria, a lack of endorsement can indeed introduce a higher risk of validity biases in assessing PTSD symptomatology. However, we argue that objectively minor events can indeed be perceived as a threat of death, serious injury, and/or physical integrity, possibly more so in cases with neck traumas as compared to traumas to other parts of the body. Additionally, even if Criterion A1 is not endorsed for all, a recent study reported that the structural relations between PTSD symptoms were similar in patients who fulfill criterion A and patients who report a subthreshold stressor [62], suggesting that assessing PTSD symptomatology in samples with subthreshold stressors is still relevant. Despite of this, however, it is still potentially critical in terms of
interpretation and feeds into the debate of increased risk of false positives, prompting careful interpretations. A further point is that the majority of the studies in the present review [15,28,36,45,54,58] also assessed PTSD symptomatology very early post-trauma, while the DSM-IV diagnostic criterion is symptoms of at least 30 days to preclude normative transient responses [1]. Hence, these assessments are very likely to capture a normative and transitory stress reaction that not necessarily has anything to do with later PTSD symptomatology, hence also challenging the validity of these assessments. Finally, the evaluation of cross-lagged associations relied on $p$-values in all studies. As a $p$-value is merely a measure for the probability of getting the present (or something beyond the present) result if the null-hypothesis is indeed true [5], this is not a good indicator of clinical relevance [18]. Instead, measures of the magnitudes of the associations (e.g., a type of effect size) along with confidence intervals are preferred way to assess the precision and relevance of the different associations. However, the majority of studies only presented (some of) this information for the significant associations and not the non-significant ones. Additionally, these were standardized regression coefficients, which are problematic to compare in multivariable relationships, as they are then controlled for different variables across studies, making pooling of such effect sizes and their interpretation a challenge [42].

The results of the present review have several implications. First of all, despite the findings underlining a close and potentially changing entanglement of the two conditions over time, it is important to not uncritically apply the mutual maintenance theory of PTSD symptomatology and pain. Further, clinicians are encouraged to be aware of this complex relationship and how it may affect treatment outcomes. This implies that clinicians screen for both pain and PTSD symptoms following traumatic exposure and are attentive of any patterns of mutual maintenance, maybe especially between pain and the PTSD symptom clusters of intrusion and hyperarousal. Future research should investigate the nature of the complex relationship between pain and PTSD symptoms with close attention to the methodological limitations addressed in the present review. Specifically, future studies should attempt to eliminate the
risk of pre-injury presence of the outcomes as well as the risk of false positives when assessing PTSD in pain patients. Thus, we hope to encourage awareness of this potential issue and argue that future work on this should ensure endorsement of the A criterion and use clinically administered interviews or focus on the core symptoms of PTSD symptoms when using questionnaires [25,38,61], while neither, however, rule out the risk of false positives. Additionally, future research should report on statistical assumptions, control and discuss the role of confounding, and include effect sizes along with confidence intervals.

Limitations

The present review is subjected to several limitations, which may influence the interpretation of the results. First, the present study has pooled study findings regardless of the large methodological differences between studies, making interpretations of the findings more complicated and uncertain. Due to these large methodological variations across studies, it was not possible to undertake meta-analysis, which would have added significantly to a narrative synthesis [19]. While this would have been a stronger methodology, we do not think it would have had impact on our conclusions. Second, the risk of bias assessments were carried out using a modified tool developed for observational cohort studies and not specifically cross-lagged modelling studies, which may give rise to risk of bias in the evaluation process by systematically evaluating the studies on potentially inadequate parameters. Third, the present review stated some exclusion criteria that the included studies did not report on, making the evaluation of eligibility unclear in some cases. Relatedly, a more explicitly trauma conceptualisation related to the diagnostic demand of the Criterion A in DSM-IV [1] would have allowed for stronger conclusions, as the lack of this poses a potential validity bias in the assessment of PTSD symptomatology. Fourth, autoregressive cross-lagged models have several limitations, which may bias the results. Among others, this technique assumes that all important predictors are in the model, something very hard to satisfy [30]. Also, the technique assumes synchronicity, which holds that the constructs at a given time points is
measured at exactly the same time, something often violated by practicalities in the data collection process [30]. Fifth, the coefficients of the significant cross-lagged paths were generally small of size, indicating relatively weak associations, which was not taken into account in our analysis. Sixth, as PTSD in DSM-5 constitutes a more inclusive and heterogeneous condition compared to DSM-IV [22] comprised of four clusters and changes to Criterion A [2], it may be that the findings in the present review may not be replicated in studies using DSM-5. Similarly, as post-traumatic distress varies across ethnic and cultural settings [e.g., 10,56], the present findings may not be generalizable to other cultural settings. Additionally, other types of post-traumatic distress than PTSD symptoms can be of relevance. Seventh, bidirectional associations were used as indicative of mutual maintenance across the studies and in the present review as well. However, mutual maintenance as a concept is much more holistic, process-oriented, and complex than the mere testing of reciprocal associations over time between two constructs, prompting critical interpretation and careful use of this terminology. Finally, a number of decisions made by the authors of the present study may influence the results. For instance, a decision was made to ignore the baseline measurement in the Feinberg-study [20] due to the fact that PTSD symptomatology was not measured at this time point. Also, there was no attempt to blind the assessors, which could potentially cause bias (especially since one study is carried out by the present author group [45]). Furthermore, the risk of bias assessments in the end, though strongly guided by the used tool, were subjective evaluations, causing all studies to deemed of acceptable level of quality despite major variations across them.

Conclusions and future directions

The findings of the present systematic review suggest an entangled relationship between pain and PTSD symptomatology over time post trauma with a potential importance of specifically hyperarousal and intrusion symptoms and maybe also catastrophizing, however with major variations in the nature of this relationship across studies and time points. Therefore, these findings only partly and
indirectly support the perspective of mutual maintenance between pain and PTSD symptomatology. In addition to difference in results across studies, synthesis was also complicated by large methodological differences between them as well as an increased risk of bias. All in all, these variations across findings as well as methodology are indicative of tentative findings, hence underlining the importance of very critical and careful interpretation. Hence, future high quality studies may change these conclusions. Such future studies ought to minimize the risk of biases and the general limitations identified by the present review and potentially apply different methodologies. This may, among others, be ecological momentary assessments and qualitative approaches that can further clarify the nature and complexities of the relationship between pain and PTSD symptomatology by adding more detailed and process-related insights.

Acknowledgements and declarations of conflicts of interest

The authors would like to thank the help of the research library in giving feedback on the searches and further state that they have no conflicts of interest to declare. No funding for the work for the present systematic review was received.

Supplemental digital content

Supplemental digital content associated with this article can be found online at http://links.lww.com/PAIN/A616.

References


**Figure Legends**

Figure 1: Visual illustrations of different autoregressive cross-lagged models.

The top illustration is a visual example of an autoregressive cross-lagged model with significant cross-lagged pathways in both directions, hence implying mutual maintenance. The second is a visual example of an autoregressive cross-lagged model with only some significant cross-lagged pathways, hence...
implying only unidirectional and not mutual maintenance. The third is a visual example of an autoregressive cross-lagged model with no significant cross-lagged pathways, hence implying no maintenance and only simple autoregressive effects. Single-headed arrows illustrate prediction and double-headed arrows illustrate correlation.

Figure 2: PRISMA 2009 Flow Diagram illustrating the data selection process.

<table>
<thead>
<tr>
<th>Authors (year)</th>
<th>Sample type (n)</th>
<th>Sample location, age, and sex</th>
<th>Trauma types</th>
<th>Assessment times</th>
<th>Time between trauma and T1</th>
<th>PTSD assessment</th>
<th>Pain assessment</th>
<th>Key findings for the present study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carty et al. (2011)</td>
<td>Hospitalized severe injury patients (n=208)</td>
<td>Australia</td>
<td>MVA (67.8%), fall (11.5%), assault (5.8%), workplace injury (4.3%), other accidents (9.1%)</td>
<td>T1: During hospitalization</td>
<td>M = 6.1 days between admission and interview</td>
<td>CAPS-IV; total severity score (excluding the item on amnesia)</td>
<td>Pain intensity across past two weeks using a 11-point VAS (at baseline, pain intensity at the time of hospital interview)</td>
<td>Evidence of both bidirectional (from baseline to 3 months) and unidirectional (from 3 months pain to 12 months PTSD symptoms) maintenance between pain and PTSD symptoms. Baseline catastrophizing predicted PTSD symptoms at 3 months, and catastrophizing at 3 months predicted pain at 12 months, hence implying a potential importance of catastrophizing, while not being a mediator in the cross-lagged relationship between pain and PTSD symptoms.</td>
</tr>
<tr>
<td>Feinberg et al. (2017)</td>
<td>Presenting at ED within 24 hours post MVC, but did not require hospitalization (n=948)</td>
<td>US</td>
<td>MVC (100 %)</td>
<td>(T1: At ED)</td>
<td>Median 1.2 hours</td>
<td>IES-R-IV; severity scores on symptom clusters</td>
<td>Axial accident-related pain severity on verbal NRS (0-10) assessed for each relevant body region (neck, left shoulder, right shoulder, upper back, lower back), then taking the most severe one.</td>
<td>Evidence of mutual maintenance of hyperarousal and pain from 6 weeks to 6 months and mutual maintenance between pain and intrusion from 6 to 12 months as well as unidirectional maintenance from pain 6 weeks to intrusion 6 months and again from pain 6 months to hyperarousal 12 months.</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Study Population</td>
<td>Location</td>
<td>Mean Age (SD)</td>
<td>Gender</td>
<td>Injuries (Frequency)</td>
<td>Assessment Timeline</td>
<td>Pain Assessment Tool</td>
<td>PTSD Assessment Tool</td>
</tr>
<tr>
<td>--------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>----------</td>
<td>---------------</td>
<td>--------</td>
<td>-------------------------------------------------------------------------------------</td>
<td>--------------------</td>
<td>----------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Jenewein et al. (2009)</td>
<td>Injured accident survivors hospitalized for minimum 2 nights post-accident (n=323)</td>
<td>Switzerland</td>
<td>40.9 (19.9)</td>
<td>35.3% female</td>
<td>Accidents involving sports and leisure time (40.6%), MVA (30%), workplace (24.4%), households (5.0%)</td>
<td>T1: During hospitalization</td>
<td>VAS 100-mm anchored line</td>
<td>CAPS-IV and PDEQ (T1), CAPS-IV (T2), DTS-IV (T3); total severity score</td>
</tr>
<tr>
<td>Liedl et al. (2010)</td>
<td>Hospitalized trauma patients (n=824)</td>
<td>Australia</td>
<td>38.9 (13.7)</td>
<td>27.7% female</td>
<td>MVA (65.9%), fall (15.5%), assault (5.8%), non-fall work-injury (5%), other accidents (6.8%)</td>
<td>T1: At trauma centre prior to discharge</td>
<td>VAS 100 mm assessing average pain level across past two weeks (at baseline, average pain since injury)</td>
<td>CAPS-IV; symptom cluster severity scores</td>
</tr>
<tr>
<td>Ravn et al. (2017)</td>
<td>Whiplash injured due to MVC’s (n=253)</td>
<td>Australia</td>
<td>38.06 (13.34)</td>
<td>66.4% female</td>
<td>MVC (100%)</td>
<td>T1: &lt; 4 weeks post-MVC</td>
<td>PDS-IV; total severity score</td>
<td>VAS assessing average pain within last 24 hours</td>
</tr>
<tr>
<td>Study Authors, Year</td>
<td>Type of Participants (n)</td>
<td>Location(s)</td>
<td>Demographics</td>
<td>Type of Exposure</td>
<td>Time Points</td>
<td>Measures</td>
<td>Findings</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>Stratton et al. (2014)</td>
<td>Military Personnel experiencing blast exposure (n=209)</td>
<td>US</td>
<td><em>M</em> = 27.4, <em>SD</em> = 7.6, 3.35% female</td>
<td>Blast exposure (100%)</td>
<td>T1: Max. 2 years post-blast, T2: 6 months post-baseline, T3: 12 months post-baseline</td>
<td>PCL-IV; total severity score</td>
<td>Evidence of both bidirectional (pain and PTSD symptoms from baseline to 6 months) and unidirectional (between PTSD symptoms at 6 months and pain 12 months) maintenance between pain and PTSD symptoms.</td>
<td></td>
</tr>
<tr>
<td>Van Loey et al. (2018)</td>
<td>Burn Victims (n=216)</td>
<td>The Netherlands and Belgium</td>
<td><em>M</em> = 40.7, <em>SD</em> = 15.4, 33% female</td>
<td>Burns (100%) often caused by fire or scalds</td>
<td>T1: At burn centre prior to discharge (&lt;2 weeks post-burn), T2: 6 months post-burn, T3: 12 months post-burn</td>
<td>IES-R-IV; total severity score</td>
<td>No maintenance patterns between pain and PTSD symptoms directly were investigated. Instead, acute PTSD symptoms predicted catastrophizing at 6 months, which then predicted pain at 12 months.</td>
<td></td>
</tr>
</tbody>
</table>
Table 1. Descriptive characteristics of included studies.

Abbreviations: CAPS-IV = Clinician Administered Posttraumatic Stress Scale for DSM-IV; DTS-IV = Davidson Trauma Scale for DSM-IV; ED = Emergency Department; IES-R-IV = Impact of Events Scale – Revised for DSM-IV; M = Mean; MVA = Motor Vehicle Accident; MVC = Motor Vehicle Crash; N = (Sample) Number; NRS = Numerical Rating Scale; PCL-IV = Posttraumatic Stress Disorder Checklist for DSM-IV; PDEQ = Peritraumatic Dissociative Experiences Questionnaire; PDS-IV = Posttraumatic Diagnostic Scale for DSM-IV; PTSD = Posttraumatic Stress Disorder; SD = Standard Deviation; SF-MPQ = the Short-Form McGill Pain Questionnaire; VAS = Visual Analogue Scale; T1 = time1; T2 = time2; T3 = time3; T4 = time4

a = Initially n=1166, but only participants that completed all parts of the study was used.
b = Originally n=301, but only participants that completed all parts of the study was used.
c = As these papers are from the same study, the measurement of pain intensity must be the same. However, the two studies describe the scale (11 point VAS vs. 100 mm VAS) as well as what is measured at baseline (“pain intensity at the time of hospital assessment” vs. “average pain since injury”) differently.
D = Of note, Feinberg et al.’s first measurement of PTSD symptomatology was 6 weeks post-trauma.
<table>
<thead>
<tr>
<th>Study</th>
<th>Item 1.1: Appropriately and clearly focused question</th>
<th>Item 1.2: Likelihood that some eligible subjects might have the outcome before enrolment</th>
<th>Item 1.3: Percentage of individuals dropped out</th>
<th>Item 1.4: Comparison between full participants and those lost to follow-up</th>
<th>Item 1.5: Clearly defined outcomes</th>
<th>Item 1.6: Reliable assessment of exposure</th>
<th>Item 1.7: Reliable outcome assessments</th>
<th>Item 1.8: Valid outcome assessments</th>
<th>Item 1.9: Main potential confounders identified and taken into account</th>
<th>Item 1.10: Confidence intervals provided</th>
<th>Item 1.11: Assumptions of methodology are taken into account</th>
<th>Item 1.12: Model fit indices indicate a fit to data</th>
<th>Final risk of bias assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carty et al. (2011)</td>
<td>✓</td>
<td>-</td>
<td>31.0%</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓*</td>
<td>?</td>
<td>-</td>
<td>?</td>
<td>✓</td>
<td>✓</td>
<td>+</td>
</tr>
<tr>
<td>Feinberg et al. (2017)</td>
<td>✓ Partly (only for pain)</td>
<td>11.0%</td>
<td>-</td>
<td>✓</td>
<td>✓</td>
<td>✓*</td>
<td>✓*</td>
<td>?</td>
<td>-</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>+</td>
</tr>
<tr>
<td>Jenewein et al. (2009)</td>
<td>✓</td>
<td>-</td>
<td>21.7%</td>
<td>-</td>
<td>✓</td>
<td>✓*</td>
<td>✓*</td>
<td>?</td>
<td>-</td>
<td>?</td>
<td>✓</td>
<td>✓</td>
<td>+</td>
</tr>
<tr>
<td>Liedl et al. (2010)</td>
<td>✓</td>
<td>-</td>
<td>29.3%</td>
<td>✓</td>
<td>✓</td>
<td>✓*</td>
<td>✓*</td>
<td>?</td>
<td>-</td>
<td>?</td>
<td>✓</td>
<td>✓</td>
<td>+</td>
</tr>
<tr>
<td>Ravn et al. (2018)</td>
<td>✓ Partly (only for WAD)</td>
<td>-</td>
<td>-</td>
<td>✓</td>
<td>✓</td>
<td>✓*</td>
<td>✓*</td>
<td>?</td>
<td>-</td>
<td>?</td>
<td>✓</td>
<td>✓</td>
<td>+</td>
</tr>
<tr>
<td>Straton et al. (2014)</td>
<td>✓</td>
<td>-</td>
<td>46.4%</td>
<td>✓</td>
<td>✓</td>
<td>✓*</td>
<td>✓*</td>
<td>?</td>
<td>-</td>
<td>?</td>
<td>✓</td>
<td>✓</td>
<td>+</td>
</tr>
<tr>
<td>Van Loey et al. (2018)</td>
<td>✓</td>
<td>-</td>
<td>33%</td>
<td>✓</td>
<td>✓</td>
<td>✓*</td>
<td>✓*</td>
<td>?</td>
<td>✓</td>
<td>?</td>
<td>✓</td>
<td>✓</td>
<td>+</td>
</tr>
</tbody>
</table>
Table 2. Illustration of risk of bias assessment items for each study.
The ✓ illustrates full endorsement of the item and corresponds to the answer “yes” in the checklist.
The – illustrates a non-fulfillment of the specific item and corresponds to the answer “no” in the checklist.
The ? illustrates doubt on whether or not the item is endorsed, often because of lack of information, and corresponds to the answer of “can’t say” in the checklist. For the risk of bias assessment, + illustrates “acceptable level of quality”.
WAD=whiplash-associated disorder
*While scales were indeed validated, other sources were not provided as demonstration of validity/reliability.
Records identified through database searching (n = 7,164)

Additional records identified through other sources (n = 0)

Records after duplicates removed (n = 4,326)

Records screened (n = 4,326)

Records excluded (n = 4,229)

Full-text articles assessed for eligibility (n = 97)

Full-text articles excluded in total (n = 90):
  wrong design (n = 57)
  wrong type (n = 26)
  wrong outcomes (n = 4)
  wrong language (n = 1)
  duplicate (n = 1)
  wrong language (n = 1)

Studies included in qualitative synthesis (n = 7)

Studies included in quantitative synthesis (meta-analysis) (n = 0)