MARCH 2017

EVIDENCE REVIEW 174

ISCRR
Institute for Safety, Compensation and Recovery Research

Identifying Client Needs

Worldwide evidence review

This review aimed to summarise the characteristics of validated tools used to collect information for the purpose of identifying and responding to client needs, and where possible describe the impact of screening on client outcomes.

Dr Ross Iles, Sunita Bayyavarapu, Dr Katharine Gibson, Emily Bell

A joint initiative of

WorkSafe VICTORIA · TAC Transport Accident Commission · MONASH University

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

This report has been prepared for the Transport Accident Commission (TAC). The Institute of Safety, Compensation and Recovery Research (ISCRR) would like to acknowledge TAC and members of the Needs Identification consultation group for their assistance and collaboration throughout the development of this evidence review. The authors also wish to thank Emily Bell, ISCRR staff and Dr Andrea de Silva for the provision of feedback and input into the report.

LIST OF ABBREVIATIONS

Abbreviation	Tool name
BDI	Beck Depression Inventory
CAPSIV	Clinician Administered PTSD (Post Traumatic Stress Disorder) Scale for DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, 4th Edition)
CCI	Causality Chain Inventory
CES-D	Centre for Epidemiological Studies Depression scale
CPG	Chronic Pain Grade
DASS-21	Depression Anxiety and Stress Scale – 21 item version
FABQ	Fear Avoidance Beliefs Questionnaire
FCE	Functional Capacity Evaluation
FIM	Functional Independence Measure
FRI	Function Rating Index
HADS	Hospital Anxiety and Depression Scale
IEQ	Injustice Experience Questionnaire
IES- revised	Impact of Events Scale - revised
IPQ	Illness Perception Questionnaire
McGill Pain Q	McGill Pain Questionnaire
NDI	Neck Disability Index
NRS	(Pain) Numerical Rating Scale
OMPQ	Orebro Musculoskeletal Pain Questionnaire
PAS	Post traumatic Adjustment Scale
PBPI	Pain Beliefs Perception Inventory
PCS	Pain Catastrophizing Scale
PSS-SR	PTSD (Post Traumatic Stress Disorder) Symptom Scale – Self-Report
PTSD-RI	PTSD (Post Traumatic Stress Disorder) Reaction Index
RMDQ	Roland Morris Disability Questionnaire
SBST	Start Back Screening Tool
SF 12	Short Form 12 Health Survey
SF36	Short Form 36 Health Survey
SMFA	Short Musculoskeletal Function Assessment
STEPP	Screening Tool for Early Predictors of PSTD (Post Traumatic Stress Disorder)
STEPP-Aus	Screening Tool for Early Predictors of PSTD (Post Traumatic Stress Disorder) – Australian version
TSQ	Trauma Screening Questionnaire
WRREQ	Work Related Recovery Expectations Questionnaire
Other abbreviations	
MSK	Musculoskeletal
PTSD	Post-Traumatic Stress Disorder
TAC	Transport Accident Commission

EXECUTIVE SUMMARY

Background and Purpose

Challenges persist in the management of clients who have suffered an injury and (or) trauma as a result of a motor vehicle accident or workplace incident. A large body of research has focused on factors that allow identification of and response to client needs in order to provide the most appropriate care at the best possible time (targeted care).

The majority of research in this area has been performed at an exploratory level, whereas confirmatory approaches provide the highest level of evidence to support targeted care approaches. Therefore, the review performed a single search of the literature to meet two aims:

1. Gather information on validated tools that have been used to collect information to identify client needs.

2. Determine the effectiveness of targeted care in providing interventions to people in the compensation setting.

Approach

A framework was developed (Collect, Interpret, Act) a priori to define the key implementation factors essential to a model of targeted care. A broad search of relevant literature was performed to identify tools that had been used to predict outcomes for people recovering from injury (physical or mental) with a focus on the compensation setting. From more than 8,000 potentially relevant studies, detailed information from 32 tools was extracted and summarised in a searchable database. Only two studies were located that allowed evaluation of the effectiveness of providing targeted care.

Findings

The tools identified in the review provide a wide range of options for identifying client needs. However, no single factor emerged to guide the selection of the most appropriate tool. This suggests a pragmatic approach should be taken to selecting a tool to apply in practice, with implementation factors guiding the decision on which tool to use.

Only two studies were located that allow evaluation of the impact of targeted care. There is an opportunity to provide world-leading research by evaluating the effectiveness of a targeted approach to care in the TAC setting.

Implications

This review indicates there is a lack of high level evidence to guide TAC decision making related to identifying and responding to client need. Gaps exist particularly related to the impact of providing targeted care approaches. In the absence of high level evidence, utilising validated tools guided by the demands of implementation factors is an appropriate approach to testing models of targeted care. It is suggested that TAC carefully consider implementation factors when aiming to identify and act upon client needs, and aim to allow full evaluation of the impact of trials of targeted care.

INTRODUCTION

Challenges persist in the management of clients who have suffered an injury and (or) trauma as a result of a motor vehicle accident or workplace incident. The TAC identifies one of their three broad client groups as "Supported Recovery". This group of recovery clients (representing 19% of all clients) are known to have more complex needs that impact TAC's aim to assist clients to get their "Life Back on Track". Utilising early identification or intervention tools post injury may identify clients at risk of poor recovery outcomes, and these clients may benefit from a different approach to their care. Therefore, the TAC is exploring an integrated and customised approach to respond to client needs and ultimately improve client outcomes. For ease of reference, this approach will be referred to as 'targeted care'.

Targeting refers to the use of individual-level information to identify an approach of care that would not have occurred had the initial information not been collected or analysed. Targeting care involves collecting, interpreting and acting on information specific to the individual in need of care. Regardless of whether this approach is referred to as risk identification, screening or prognosis, there are clear stages in the development of evidence behind methods of targeting care (Figure 1).

Ongoing emphasis on early identification in practice and in the literature has led to large body of research in the early, exploratory stages of prognostic research (Figure 1). Few tools have been examined prospectively or in validation studies and little is known about the actual impact of targeting care in practice. Future research needs to focus on how to implement methods of targeting into practice (confirmatory stages), rather than on new tools or methods of screening.

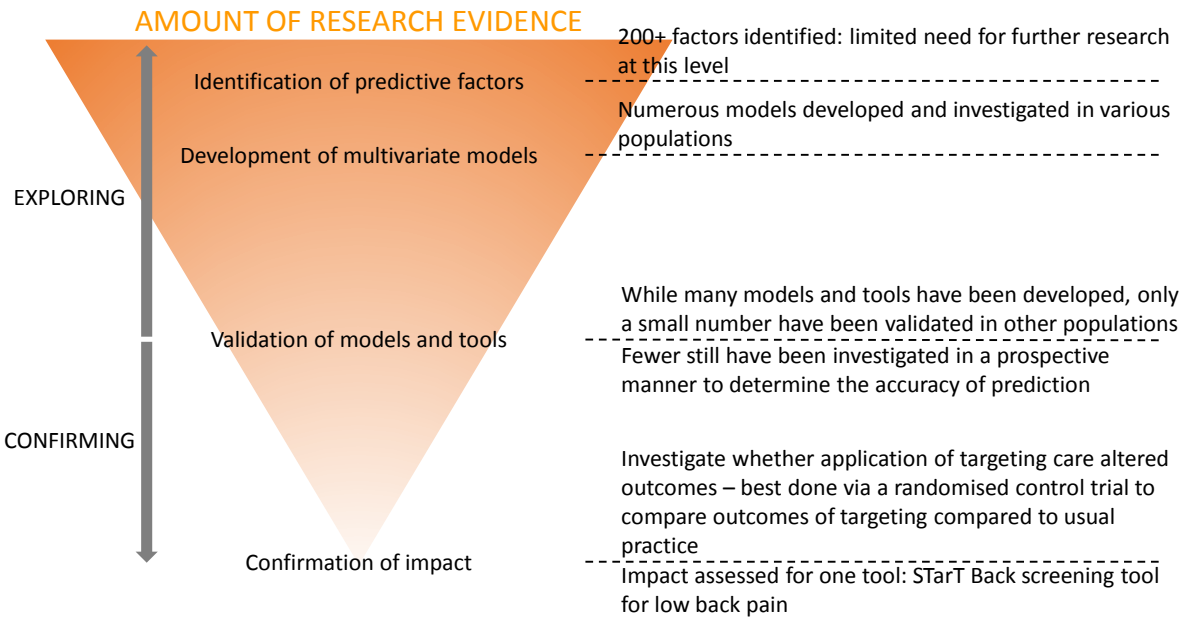

Fig 1. Relative levels of evidence for targeting care research

The result of the focus on exploratory rather than confirmatory research in this area means there is high potential for distraction away from methods that have been investigated to determine their impact. Therefore, the aim of this review was to summarise the characteristics of validated tools used to collect information for the purpose of identifying and responding to client needs, and where possible describe the impact of screening on client outcomes. This aim was achieved using a single

review strategy with two parts, one to gather information related to the tools or methods used in existing research, and the other to determine the impact of applying such tools.

Research questions

Part 1: What validated tools or methods have been used to collect information to identify client needs?

Key sub questions were as follows:

- What are the characteristics of these tools in relation to how information is collected?
- What are the requirements to be able to interpret the information collected by these tools?
- What are the recommended actions after application of the tool?

Part 2: What is the effectiveness of targeted care in providing interventions to people in the compensation setting?

Key sub questions were as follows:

- What resources were required to collect, interpret and act on information collected as part of targeted care?
- What are the key characteristics of a successful targeted care intervention?
- What are the timeframes in which targeted care is most effectively applied?

METHODS

This review adapted the scoping methodology outlined by Arskey and O'Malley (2005) to identify relevant tools in the identification of client needs. This approach consisted of 6 steps: (1) Identifying the research question; (2) identifying relevant studies; (3) study selection; (4) charting the data; (5) collating, summarising and reporting the results; and (6) consulting stakeholders for knowledge translation of findings.

The Collect, Interpret, Act framework (Appendix 1) was developed to chart tools (sort according to key themes) in relation to the nature of the information collected by the tool, how the tool should be interpreted and how the results have been used to initiate action. This framework will also be applied to the discussion sections of this report. This methods section applies to both parts of this review.

Search Strategy

A single comprehensive search was conducted to meet both parts of the review. In October 2016, we undertook a systematic search of the Cochrane, CINAHL, EMBASE, Medline, Scopus and PsycINFO databases for peer reviewed articles between 1995 and 2016 with an English language restriction imposed on all searches. Search strategies were adapted to the various databases as required. Search terms focused on musculoskeletal (MSK) disease, mental illness, brain injury, fractures, screening tools, validity, injury, trauma, and compensation. Examples of key terms included: traffic accidents, musculoskeletal pain, mental disorders (depression, anxiety, and PTSD), or preventive health services (see Appendix 2 for further detail). Where a tool was identified as relevant to the review, a targeted Google search was undertaken to retrieve the complete tool items and scoring descriptions. Author-specific searches were conducted to ensure all tools of interest were captured.

Part 1 inclusion criteria

Title and abstracts of the retrieved articles were screened and 56 references were identified that met the inclusion exclusion criteria developed by the ER team (in collaboration of with the TAC stakeholders) for full review and data extraction for Part 1 of the review.

Population: Studies were included for the reviews if the predictive/screening tools were utilised for adults aged 18 years and above, who have sustained an injury, including those injured in a motor vehicle, workplace, or other accident or incident, and have MSK disorders and or mental health disorders including orthopaedic trauma and mild ABI. The review included all MSK and mental health conditions available in the literature and all terms referring to MSK and mental health disorders. Non injury conditions such as cancer and pregnancy were excluded.

Intervention: Studies of tools were included for review if the tool was used for client examination, screening, or assessment from the time of an injury report; for example, demographic questionnaires; clinical prediction rules; clinical classification systems; self-reported – demographic; health; or economic reports, and risk screening were included. Desired characteristics of a screening tool, such as reliability, validity and accuracy in identification of client needs were incorporated into the search strategy to ensure that we identify only validated tools. The tools that have only been developed using an exploratory approach were excluded.

Outcomes: To be eligible, studies of relevant tools had to include key outcomes of interest to TAC including recovery from pain; improved quality of life; improved physical function; return to work; claim cost; claim duration; progress to legal proceedings.

Part 2 inclusion criteria

In order to determine the true effectiveness of targeted care approaches, the method of targeting must be applied to two separate groups (Hingorani et al., 2013). In one group the method of targeting (collecting and interpreting information) is followed by action (tailored intervention). In the control group, action consists of usual care. This then allows the direct comparison of the outcomes from each sub group identified (for example, two groups identified as high risk) to determine whether targeted and its associated interventions lead to better results between groups and between subgroups.

To be included in the second part of this review, studies had to apply the targeting process to two distinct groups, with one group receiving intervention based on the collection of information and the other group receiving usual care.

Data extraction

The Collect, Interpret, Act framework (Appendix 1) formed the basis of data extraction for each tool identified by the review. Alongside information specific to the study applying the tool (such as study design, the outcome predicted, strength of prediction and authors' conclusion), information describing the nature of the information collected by the tool, the method of interpretation applied in the study and actions resulting from application of the tool were extracted. General characteristics of each tool were summarised in terms of the number of items contained, whether the included studies examined the tool in a compensation setting, a broad categorisation of the constructs measured (physical, mental or mixed), the timeframe when the tool was measured, the timeframe of the outcome that the tool predicted and the TAC complexities measured within the tool.

Where the study design provided evidence of the impact of applying the tool (such as a randomised controlled trial), data related to the outcomes was extracted in addition to the Collect, Interpret, Act framework.

Quality appraisal

Part 1: Validated tools and methods

Since the design of the review aimed to build a catalogue of potentially useful tools, it was possible that the predictive characteristics of the tools displayed in the included studies may not have been an accurate indication of the tool properties across all applications. Secondary searches were conducted to locate systematic reviews or studies of the measurement properties of the included tools to provide a more comprehensive summary of the level of evidence behind each tool (see Appendix 4 for search details). Where systematic reviews were available, tools were rated as having a high level of support (consistently excellent measurement properties), medium level of support (consistently good measurement characteristics) or a low level of support (acceptable or inconsistent measurement properties). The rules applied to determine the predictive strength of the tools, and the decision levels for evidence support appear in Appendix 5.

Part 2: Effectiveness of targeted care

A modified Cochrane risk of bias tool was applied to assess the methodological quality of randomised controlled trails evaluating the impact of targeted care approaches. This tool assesses bias in six domains: sequence generation, allocation concealment, blinding, outcome data, data reporting and other sources of bias. It is used to arrive at an overall assessment of the risk of bias in the study results (high or low risk of bias).

RESULTS

Overview of studies

A flowchart of the number of studies that were identified and screened in this evidence review is presented in Figure 2. In summary, a total of 37 tools were identified during the process of data extraction, of which it was possible to extract detailed information from 32 tools for Part 1 of this review. Details of the references related to each of the tools can be found in Appendix 3.

From the 8,021 unique studies identified in the search, two met the inclusion criteria required to be able to evaluate the impact of the targeted care approach applied.

Study Characteristics

Table 1 describes the study designs and evidence level of the studies included in the review.

Table 1. Study designs included in the review

Study type	Evidence level[1]	Number
Prospective cohort study	III-2	30
Development and validation study	III-3	8
Randomised Controlled Trial	II	4
Longitudinal study	III-2	3
Retrospective study	III-3	3
Observation study	IV	2
Inception cohort study	III-3	2
Population based study	IV	1
Comparative study	III-3	1
Correlation study	IV	1
Pre-post study	IV	1

*1. Full description of NHMRC levels of evidence a*vailable here:
https://www.nhmrc.gov.au/_files_nhmrc/file/guidelines/developers/nhmrc_levels_grades_evidence_120423.pdf

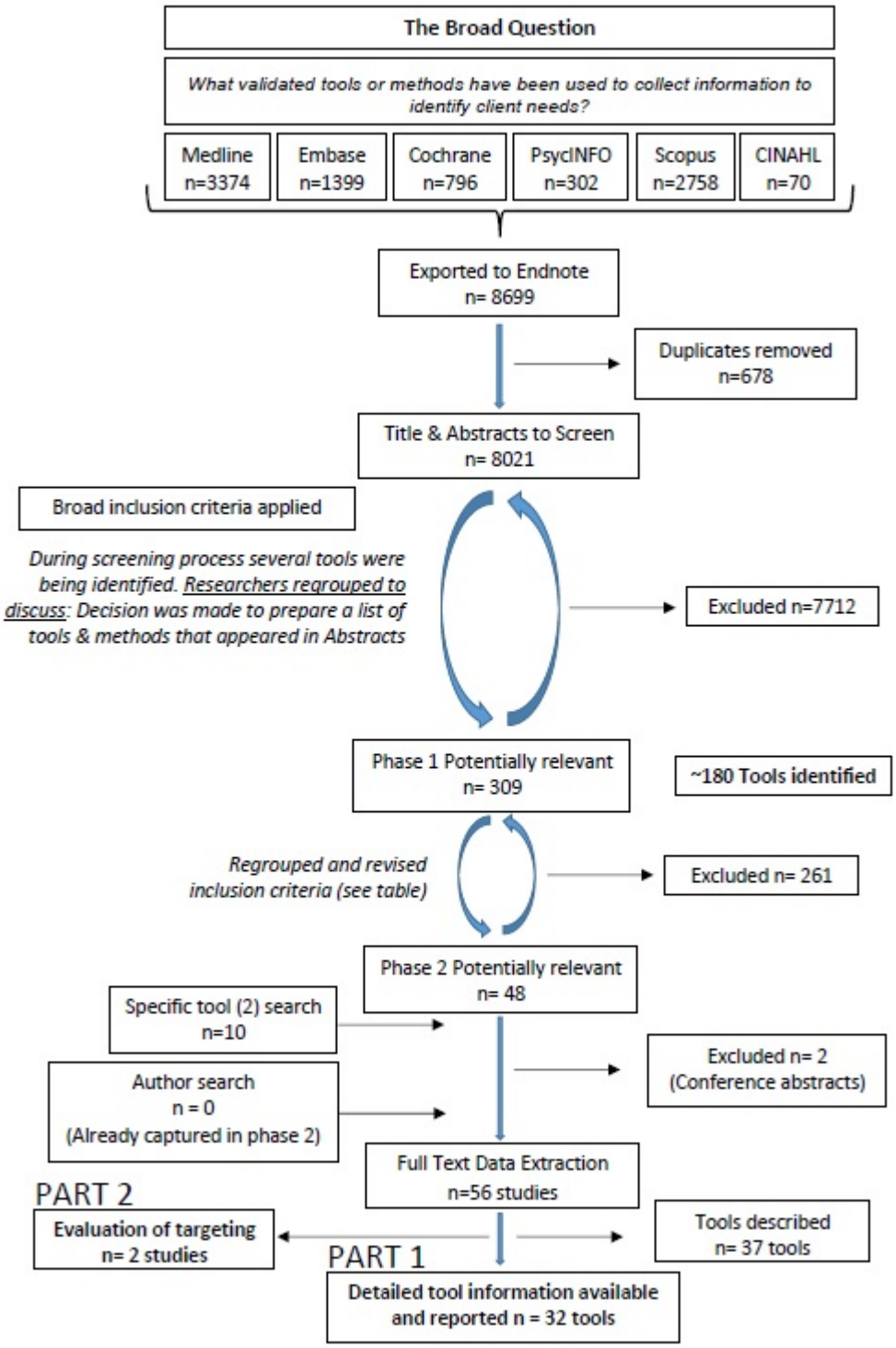

Fig 2. Review flow chart

Part 1: Validated tools used to collect information to identify client needs

Quality appraisal

Using systematic reviews of tool measurement properties, the level of support for each tool was rated as high, medium or low (Table 2). Five tools were rated as having a high level of support, 15 medium, 4 low and for the remaining 8 tools no further information on the measurement properties was located.

Table 2. Research support for included tools and citation of source for evidence of characteristics

Level of support	Tools	Source
High	BDI	(Wang & Gorenstein, 2013)
	FIM	(Glenny & Stolee, 2009)
Consistently excellent measurement properties	FRI	(Feise & Menke, 2010)
	NRS	(Hawker, Mian, Kendzerska, & French, 2011)
	RMDQ	(Smeets, Koke, Lin, Ferreira, & Demoulin, 2011)
Medium	CES-D	(Carleton et al., 2013)
	CPG	(Hawker et al., 2011)
Consistently good measurement properties or mix of good to excellent properties	DASS-21	(Shea, Tennant, & Pallant, 2009)
	FABQ	(Bishop, Thomas, & Foster, 2007)
	HADS	(Norton, Cosco, Doyle, Done, & Sacker, 2013)
	IEQ	(Sullivan et al., 2008a)
	IES	(Sundin & Horowitz, 2002)
	IPQ	(Broadbent et al., 2015)
	McGill Pain Q	(Hawker et al., 2011)
	NDI	(Yao et al., 2015)
	OMPQ	(Sattelmayer, Lorenz, Roder, & Hilfiker, 2012)
	PCS	(Walton, Wideman, & Sullivan, 2013)
	PTSD-RI	(Steinberg et al., 2013)
	SF36	(Hawker et al., 2011)
	STEPP-Aus	(Nixon, Ellis, Nehmy, & Ball, 2010a)
Low	SBST	(Karran et al., 2017)
	SMFA	(Bouffard, Bertrand-Charette, & Roy, 2016)
Lower categories of measurement properties or large inconsistencies	STEPP	(Nixon et al., 2010a)
	TSQ	(Walters, Bisson, & Shepherd, 2007)
Information not available	CAPSIV	
	CCI	
No reviews or studies examining measurement properties of tool located in search	FCE	
	PAS	
	PBPI	
	PSS-SR	
	SF12	
	WRREQ	

Tool characteristics

The Collect, Interpret, Act framework and included study characteristics were combined into a synthesis spreadsheet, provided as a separate companion file. In the companion spreadsheet file the 32 tools for which detailed information could be extracted are represented as both tools (where general characteristics are displayed) and items (allowing investigation of the range of constructs

and response scales applied). This format allows arrangement according to the tool, scoring method, construct measured or any other property extracted from the data. This enables viewing only the items that meet a specific purpose (e.g. items that measure recovery expectation on a scale from 0-10). Each individual item from the 32 tools was examined to determine the nature of the construct assessed by that aspect of the tool. This enabled identification of the range of constructs measured across more than 700 individual questions from the 32 included tools.

As part of data extraction, an indication of the predictive strength of the tool as demonstrated in the included studies is provided. Predictive strength was categorised as excellent (green), good (yellow), moderate (red), and correlation only or not significant (grey). The decision levels for these ratings are provided in Appendix 5. The predictive strength for different outcomes investigated are represented in columns labelled 'Strength of prediction 1', 'Strength of prediction 2' and 'Strength of prediction 3' columns. Details of the outcome predicted (e.g. depression at three months) are described in the corresponding 'Predictive strength details' columns.

The following is an overview of the information appearing in the companion spreadsheet file.

What are the characteristics of these tools in relation to what information is collected and how?

Level of research support

Tools with the highest level of support for their measurement properties did not show strong predictive characteristics, with the exception of the BDI providing an excellent prediction of psychiatric morbidity at 18 months (Table 3). Of the 15 tools with a medium level of support, 6 demonstrated excellent prediction covering measures of mental and physical health, and a mixture of both (Table 4). Of the tools with low support or no information related to these characteristics, one provided clearly excellent prediction of outcomes (SBST). Otherwise, two tools (STEPP, TSQ) suggested good predictive ability for PTSD outcomes (Table 5).

Table 3. High level of research support and predictive strength

Name of tool	Number of items	Strength of prediction	Strength of prediction 2	Strength of prediction 3	SR of meas properties available[1]	Recommendation level from review
BDI	21	Excellent	.	.	Yes	High
FIM	18	Moderate	Moderate	.	Yes	High
FRI	10	Correlation only	.	.	No	High
NRS	1	Correlation only	.	.	Yes	High
RMDQ	24	Moderate	Moderate	Moderate	Yes	High

[1] The meaning of this heading is that whether we were able to identify systematic reviews that reviewed/evaluated the measurement properties of these tools and assessed their validity and reliability.

Table 4. Medium level of research support and predictive strength

Name of tool	Number of items	Strength of prediction	Strength of prediction 2	Strength of prediction 3	SR of meas properties available[1]	Recommendation level from review
CES-D	20	Correlation only	.	.	No	Medium
CPG	7	Not significant	Not significant	.	No	Medium
DASS-21	21	Excellent	Excellent	Excellent	No	Medium
FABQ	16	Excellent	.	.	Yes	Medium
HADS	14	Excellent	Excellent	.	Yes	Medium
IEQ	12	Moderate	Not significant	.	No	Medium
IES-REVISED	22	Correlation only	.	.	No	Medium
IPQ	66	Good	Excellent	.	Yes	Medium
McGill Pain Q	48	Not significant	.	.	Yes	Medium
NDI	10	Excellent	Excellent	.	Yes	Medium
OMPQ	21	Moderate	Excellent	Cut off information	Yes	Medium
PCS	13	Moderate	.	.	No	Medium
PTSD-RI	33	Moderate	.	.	No	Medium
SF36	36	Correlation only	.	.	Yes	Medium
STEPP-Aus	8	Good	Not significant	.	No	Medium

Table 5. Low level of research or missing support and predictive strength

Name of tool	Number of items	Strength of prediction	Strength of prediction 2	Strength of prediction 3	SR of meas properties available[1]	Recommendation level from review
CAPSIV	25	TBD	TBD	TBD	TBD	N/A
CCI	8	Good	.	.	No	N/A
FCE	8	Not significant	.	.	Unknown	N/A
PAS	10	Moderate	Moderate	.	No	N/A
PBPI	16	Moderate	.	.	Yes	N/A
SBST	9	Excellent	Excellent	Excellent	No	Low
SF12	12	TBD	TBD	TBD	TBD	N/A
PSS-SR	17	Good	.	.	No	N/A

Name of tool	Number of items	Strength of prediction	Strength of prediction 2	Strength of prediction 3	SR of meas properties available[1]	Recommendation level from review
SMFA	46	Correlation only	.	.	Yes	Low
STEPP	12	Good	Not significant	.	No	Low
TSQ	10	Good	.	.	No	Low
WRREQ	3	Moderate	Not significant	.	No	N/A

Length of the tool

The average number of questions in each tool was just under 19 questions, with 10 of the 32 tools containing 10 or less questions and 9 of the 32 containing 20 or more (Table 6). The greater the number of items, the more detailed the information captured to address client needs, however this needs to be balanced with the administrative burden of collecting and organizing the information. Longer or more comprehensive tools did not show a consistent pattern of providing stronger prediction or having greater levels of evidence to support use of the tool (Table 6).

Table 6. Predictive strength according to number of items

Name of tool	Number of items	Strength of prediction	Strength of prediction 2	Strength of prediction 3
NRS	1	Correlation only	.	.
WRREQ	3	Moderate	Not significant	.
CPG	7	Not significant	Not significant	.
STEPP-Aus	8	Good	Not significant	.
CCI	8	Good	.	.
FCE	8	Not significant	.	.
SBST	9	Excellent	Excellent	Excellent
NDI	10	Excellent	Excellent	.
PAS	10	Moderate	Moderate	.
FRI	10	Correlation only	.	.
TSQ	10	Good	.	.
IEQ	12	Moderate	Not significant	.
STEPP	12	Good	Not significant	.
SF12	12	TBD	TBD	TBD
PCS	13	Moderate	.	.
HADS	14	Excellent	Excellent	.
FABQ	16	Excellent	.	.
PBPI	16	Moderate	.	.
PSS-SR	17	Good	.	.

Name of tool	Number of items	Strength of prediction	Strength of prediction 2	Strength of prediction 3
FIM	18	Moderate	Moderate	.
CES-D	20	Correlation only	.	.
OMPQ	21	Moderate	Excellent	Cut off information
DASS-21	21	Excellent	Excellent	Excellent
BDI	21	Excellent	.	.
IES-revised	22	Correlation only	.	.
RMDQ	24	Moderate	Moderate	Moderate
CAPSIV	25	TBD	TBD	TBD
PTSD-RI	33	Moderate	.	.
SF36	36	Correlation only	.	.
SMFA	46	Correlation only	.	.
McGIll Pain Q	48	Not significant	.	.
IPQ	66	Good	Excellent	.

Tool setting

Thirteen of the 32 tools were clearly demonstrated in a compensation setting, with 3 of these tools (OMPQ, FABQ and HADS) suggesting excellent predictive ability (Table 7). Of the 16 tools not tested in a compensation setting, 5 demonstrated excellent predictive ability (SBST, DASS, IPQ, NDI, BDI) and 5 good predictive ability (STEPP, STEPP-Aus, CCI, PSS-SR, TSQ) (Table 8). It is not clear whether these predictive characteristics would be maintained if applied in a compensation setting.

Table 7. Predictive strength of tools applied in a compensation setting

Name of tool	Compensation Setting	Strength of prediction	Strength of prediction 2	Strength of prediction 3
CPG	Yes	Not significant	Not significant	.
FABQ	Yes	Excellent	.	.
FCE	Yes	Not significant	.	.
FRI	Yes	Correlation only	.	.
HADS	Yes	Excellent	Excellent	.
IEQ	Yes	Moderate	Not significant	.
McGill Pain Q	Yes	Not significant	.	.
NRS	Yes	Correlation only	.	.
OMPQ	Yes	Moderate	Excellent	Cut off information
PBPI	Yes	Moderate	.	.
RMDQ	Yes	Moderate	Moderate	Moderate
SF36	Yes	Correlation only	.	.
WRREQ	Yes	Moderate	Not significant	.

Table 8. Predictive strength of tools not applied in a compensation setting

Name of tool	Compensation Setting	Strength of prediction	Strength of prediction 2	Strength of prediction 3
BDI	No	Excellent	.	.
CCI	No	Good	.	.
CES-D	No	Correlation only	.	.
DASS-21	No	Excellent	Excellent	Excellent
FIM	No	Moderate	Moderate	.
IES-revised	No	Correlation only	.	.
IPQ	No	Good	Excellent	.
NDI	No	Excellent	Excellent	.
PCS	No	Moderate	.	.
PSS-SR	No	Good	.	.
PTSD-RI	No	Moderate	.	.
SBST	No	Excellent	Excellent	Excellent
SMFA	No	Correlation only	.	.
STEPP	No	Good	Not significant	.
STEPP-Aus	No	Good	Not significant	.
TSQ	No	Good	.	.
PAS	Unclear	Moderate	Moderate	.

Constructs measured

The constructs measured by the tools addressed aspects of mental health (14 of 32 tools), 8 tools measured physical aspects and 10 tools measured a mix of mental and physical aspects of health. In terms of predicting an outcome, tools that measured aspects of mental health to predict a mental-health related outcome performed well (DASS, BDI, HADS, PSS-SR, TSQ) (Table 9). The FABQ measured solely mental health constructs and arrived at an excellent prediction of work outcome (high sensitivity enabling the potential to use the tool to rule out those who will experience a poor work outcome). Other tools using measures of mental health to predict an outcome including physical health reached moderate prediction or less. Tools solely measuring physical aspects did not perform well in a predictive sense, with one exception (NDI) providing excellent prediction of ongoing disability after 6 months (Table 10). Tools that measured a combination of physical and mental constructs displayed a generally stronger pattern of predictive ability for a combination of outcomes (Table 11).

Table 9. Predictive strength of tools measuring aspects of mental health.

Name of tool	Constructs (physical, mental, mixed)	Strength of prediction	Strength of prediction details	Strength of prediction 2	Strength of prediction details 2
BDI	Mental	Excellent	Psychiatric morbidity at 18 months	.	.

Name of tool	Constructs (physical, mental, mixed)	Strength of prediction	Strength of prediction details	Strength of prediction 2	Strength of prediction details 2
CES-D	Mental	Correlation only	.	.	.
DASS-21	Mental	Excellent	Depression at 3 months	Excellent	Depression at 6 months
FABQ	Mental	Excellent	High sensitivity for RTW outcome	.	.
HADS	Mental	Excellent	Diagnosis mood disorder	Excellent	Diagnosis anxiety disorder
IEQ	Mental	Moderate	Work status at 12 months	Moderate	Depression at 12 months
IES-revised	Mental	Correlation only	.	.	.
PAS	Mental	Moderate	PTSD at 12 months	Moderate	Depression at 12 months
PBPI	Mental	Moderate	Pain intensity at 6 months	.	.
PCS	Mental	Moderate	Disability at 8 months	.	.
PSS-SR	Mental	Good	PTSD	.	.
PTSD-RI	Mental	Moderate	PTSD at 12 months	.	.
TSQ	Mental	Good	PTSD at 6 months	.	.
WRREQ	Mental	Moderate	Claim closure and suspension of benefits in low back pain	Not significant	Conditions other than back pain

Table 10. Predictive strength of tools measuring aspects of physical health.

Name of tool	Constructs (physical, mental, mixed)	Strength of prediction	Strength of prediction details	Strength of prediction 2	Strength of prediction details 2
CPG	Physical	Not significant	Improvement at 6 months	Not significant	RTW at 6 months
FCE	Physical	Not significant	Future work capacity	.	.
FIM	Physical	Moderate	Physical disability at 1 month	Moderate	1 month FIM predictive of physical disability at 6 months
FRI	Physical	Correlation only	.	.	.
McGill Pain Q	Physical	Not significant	Therapy outcome after 8 visits	.	.

Name of tool	Constructs (physical, mental, mixed)	Strength of prediction	Strength of prediction details	Strength of prediction 2	Strength of prediction details 2
NDI	Physical	Excellent	Chronic disability at 6 months	Excellent	In combination with other factors
NRS	Physical	Correlation only	.	.	.
RMDQ	Physical	Moderate	Work status at 1 month	Moderate	Work status at 6 months

Table 11. Predictive strength of tools measuring a combination of mental and physical health components.

Name of tool	Constructs (physical, mental, mixed)	Strength of prediction	Strength of prediction details	Strength of prediction 2	Strength of prediction details 2
CAPSIV	Mixed	TBD	TBD	TBD	TBD
CCI	Mixed	Good	Anxiety and depression at 6 months	.	.
IPQ	Mixed	Good	High pain at 12 months	Excellent	3 month IPQ predictive of high pain at 12 months
OMPQ	Mixed	Moderate	Long-term sick leave	Excellent	High specificity for long-term sick leave
SBST	Mixed	Excellent	Poor outcome at 12 months	Excellent	Activity limitation (RMDQ score) at 12 months
SF12	Mixed	TBD	TBD	TBD	TBD
SF36	Mixed	Correlation only	.	.	.
SMFA	Mixed	Correlation only	.	.	.
STEPP	Mixed	Good	PTSD at 3, 6 months	Not significant	Depression at 3, 6 months
STEPP-Aus	Mixed	Good	PTSD at 3, 6 months	Not significant	Depression at 3, 6 months

A deeper analysis of the constructs was conducted to capture the range of information collected across the tools. Each item was examined and labelled according to the construct the question addressed. A total of 74 different constructs were identified (Table 12). Constructs such as anxiety and depression were commonly measured, while others like treatment history appeared only in a single question.

Examination of the questions addressing particular constructs suggests there are items that are likely to be unsuitable to be applied in a case management setting. Items that address sexual function are obvious questions likely to cause discomfort when being asked if the context or level of rapport with the interviewer does not invite that level of examination. Other impacts of asking certain types of

questions must also be considered. For example the potential consequences of probing for information around stressful events (commonly seen in tools addressing aspects of PTSD) require the interviewer to be prepared and even qualified to deal with a client in distress. A common element amongst questionnaires assessing anxiety and depression is the requirement for the respondent to rate the frequency of a range of emotions. This type of question transfers to a case management situation more readily than a question that requires the person to choose a description that matches the person's current emotional state.

Not all constructs thought to be influential in the prediction of outcomes after injury are represented in the tools in this review. Resilience, a construct of interest to TAC, is not addressed across these tools. Health knowledge and understanding of one's condition are represented, but these aspects appear only very briefly across the included tools. This may be due to the focus on tools that have been examined in a confirmatory manner, suggesting these constructs have not progressed to being investigated to this level in the research to date. Similarly, the TAC complexities "prior health", "solicitor involvement", "service environment – external" and "prior finance" are *rarely covered* across the included tools. These aspects suggest that sources outside of this review are needed to assess these constructs deemed to be important factors in Supported Recovery client group.

Table 12. The 74 constructs measured across included tools

Abuse	Communication	Fear avoidance	Improvement	Seriousness
Acceptance	Compensation	Fear for future	Injustice	Sexual abuse
Accident	Concentration	Fear of death	Job satisfaction	Sexual function
Activity limitation	Coping	Finances	Life impact	Sleep
Age	Death	Gender	Locus of control	Stress
Agitation	Depression	General health	Memory	Symptoms
Anger	Despair	Genetics	Neglect	Toileting
Anxiety	Disaster	Get life back	Pain	Transfers
Appetite	Disbelief	Headache	Past experience	Treatment expectation
Attention	Emotional impact	Health knowledge	Perception of others	Treatment history
Avoidance	Extent of injury	Heavy work	PTSD	Understanding of condition
Bereavement	Family support	Helplessness	Recovery expectation	Violence
Catastrophising	Fatigue	Illness perception	Self care	Work ability
Cause of condition	Fault	Impact on others	Sensory disturbance	Work stress
Cognition	Fear	Impaired caregiver	Separation	

Timing of application

Tools were applied to a range of timeframes ranging from an acute hospital admission to beyond 12 months after injury. Different timeframes were summarised into acute (within 4 weeks of injury), sub-acute (between 4 and 12 weeks of injury) and chronic (beyond 12 weeks after injury). Twelve of the 32 tools collected information in the acute period, and four of these demonstrated excellent predictive ability for long term outcomes (Table 13). Those tools collecting information in the sub-acute phase (10 of 32) were less consistent in predicting outcomes; two displayed excellent

predictive characteristics, but five suggested correlations only (Table 14). Only four tools collected information in the chronic phase of injury, and apart from the FABQ suggesting excellent prediction of poor RTW outcome, the remaining three displayed low predictive characteristics (Table 15). Information about when the remaining 6 tools (HADS, CAPSIV, FCE, IEQ, McGill Pain Q and SF-12) were applied was not reported in the studies selected for this review. As a result it was 'unknown' if they were applied in acute, subacute and/or chronic phase.

Table 13. Predictive strength of tools applied in the acute phase.

Name of tool	Timeframe collected	Timeframe window	Strength of prediction	Strength of prediction details	Strength of prediction 2	Strength of prediction details 2
BDI	2 weeks	Acute	Excellent	Psychiatric morbidity at 18 months	.	.
CCI	1 month	Acute	Good	Anxiety and depression at 6 months	.	.
DASS-21	14 days	Acute	Excellent	Depression at 3 months	Excellent	Depression at 6 months
FIM	Acute hospital admission	Acute	Moderate	Physical disability at 1 month	Moderate	1 month FIM predictive of physical disability at y months
IPQ	10 days	Acute	Good	High pain at 12 months	Excellent	3 month IPQ predictive of high pain at 12 months
OMPQ	4, 12 weeks	Acute	Moderate	Long-term sick leave	Excellent	High specificity for long-term sick leave
PAS	2 weeks	Acute	Moderate	PTSD at 12 months	Moderate	Depression at 12 months
PSS-SR	1 month	Acute	Good	PTSD	.	.
PTSD-RI	2 weeks	Acute	Moderate	PTSD at 12 months	.	.
RMDQ	4 weeks	Acute	Moderate	Work status at 1 month	Moderate	Work status at 6 months
STEPP	1 week	Acute	Good	PTSD at 3, 6 months	Not significant	Depression at 3, 6 months
STEPP-Aus	4 weeks	Acute	Good	PTSD at 3, 6 months	Not significant	Depression at 3, 6 months

Table 14. Predictive strength of tools applied in the sub-acute phase.

Name of tool	Timeframe collected	Timeframe window	Strength of prediction	Strength of prediction details	Strength of prediction 2	Strength of prediction details 2
CES-D	2 months	Sub-acute	Correlation only	.	.	.
IES-revised	2 months	Sub-acute	Correlation only	.	.	.
NDI	4,6 weeks	Sub-acute	Excellent	Chronic disability at 6 months	Excellent	In combination with other factors
NRS	25-102 days	Sub-acute	Correlation only	.	.	.
PBPI	6 weeks	Sub-acute	Moderate	Pain intensity at 6 months	.	.
PCS	1,2,3 months	Sub-acute	Moderate	Disability at 8 months	.	.
SBST	Various	Sub-acute	Excellent	Poor outcome at 12 months	Excellent	Activity limitation (RMDQ score) at 12 months
SF36	3 months	Sub-acute	Correlation only	.	.	.
SMFA	1-2 months	Sub-acute	Correlation only	.	.	.
TSQ	8 weeks	Sub-acute	Good	PTSD at 6 months	.	.

Table 15. Predictive strength of tools applied in the chronic phase.

Name of tool	Timeframe collected	Timeframe window	Strength of prediction	Strength of prediction details	Strength of prediction 2	Strength of prediction details 2
CPG	12 months	Chronic	Not significant	Improvement at 6 months	Not significant	RTW at 6 months
FABQ	12, 26 weeks	Chronic	Excellent	High sensitivity for RTW outcome	.	.
FRI	3 months	Chronic	Correlation only	.	.	.
WRREQ	12 months plus	Chronic	Moderate	Claim closure and suspension of benefits in low back pain	Not significant	Conditions other than back pain

What are the requirements to be able to interpret the information collected by these tools?

The embedded spreadsheet provides the response scales for the individual items across the 32 included tools. Response scales ranged from dichotomous yes/no responses, numerical scales and requiring the respondent to select the most appropriate response from a short selection. The majority of tools provided a scoring mechanism based on the responses, with the expectation that a certain score, or cut off, could be interpreted as having a certain chance of developing the outcome of interest. Scoring the majority of tools was algorithm based, making determining the final score relatively straight forward.

However, determining a suitable cut off for action proves difficult across included studies. For example, in the eight studies included in the review that examined the OMPQ, each applied a different cut of score to determine high and low risk groups (i.e. initiate action). This is further emphasised in a systematic review addressing the predictive ability of the OMPQ that does not recommend the use of a single cut off value (Sattelmayer et al., 2012). Other implications of assigning a specific cut off value is that it risks the application of a tool becoming another task to be completed, with a focus on the result rather than the process involved in collecting the information (i.e. listening to the answers). Across the included studies the collection and interpretation of the tool scores was often performed by a member of the research team, so it is not clear how readily these tasks can be transferred to case managers.

What are the recommended actions after application of the tool?

Across all the studies that examined the included tools, only two studies provided information related to the actions to be applied, Hill et al., (2011) and Jull et al., (2013). The effectiveness of these approaches is examined in Part 2 of this review.

Part 2: What is the effectiveness of targeted care in providing interventions to people in the compensation setting?

Of the almost 8,700 references identified in the original search, just two studies allowed analysis of the impact of targeted care, Hill et al (2011) and Jull et al (2013). These two randomised controlled trials are summarised in Appendix 6.

Hill et al (2011) – Start Back Screening Tool

Hill et al., (2011) applied the Start Back Screening Tool (SBST) to people seeking care for low back pain in the UK. Based on the SBST results, participants received a different level of physiotherapy care according to their level of identified risk. Participants were randomly allocated to the intervention or control arm of the study, allowing comparison of each of the low, medium and high risk groups' outcomes.

Quality Appraisal

Quality appraisal was performed using the Cochrane assessment of bias. Hill et al (2011) was deemed to contain a low risk of bias (Table 16).

Table 16. Quality appraisal of Hill et al (2011)

	Sequence generation	Allocation concealment	Blinding of participants	Incomplete outcome data	Selective outcome reporting	Other sources of bias	Overall assessment
Hill et al (2011)	Yes	Yes	Yes	Yes	Yes	Yes	Low risk of bias

What information was collected

Hill et al (2011) stratified according to the SBST and collected information related to demographics, RMDQ disability score, back pain intensity, duration of back pain, pain catastrophizing (PCS), fear avoidance via the Tampa Scale of Kinesiophobia (TSK), and general health (SF12). This information was collected in people who were at least 18 years old with back pain of any duration attending their GP for care for low back pain.

Detailed information related to the SBST is contained in the embedded spreadsheet above. In summary the SBST contains 9 items related to bothersomeness of pain, ability to dress and walk short distances, fear avoidance, anxiety, catastrophizing and depression (i.e. mixed physical and mental health constructs).

How was the information interpreted

There are nine items, with each question is scored 0 or 1. Five items (5-9) are used to calculate a psychological score. The approach employed by Hill et al (2011) using the Start Back Screening Tool (SBST) allocated people seeking primary care for low back pain into three risk categories. Eligible participants were administered the SBST tool by researchers and grouped into low risk (≤3 overall score), medium risk (≥4 overall score +Psych score ≤3) and high risk (≥4 overall score+ ≥4 Psych score) groups (see embedded excel file for SBST items and scoring). Participants were randomly assigned to intervention and control arms in the study, meaning that both intervention and control groups consisted of each risk profile.

What action was initiated

All participants in the intervention arm received a 30 minute assessment and initial treatment according to an agreed protocol, with advice focusing on promotion of appropriate levels of activity, including return to work, and a pamphlet about local exercise venues and self-help groups.

Participants were also shown a 15 min video entitled Get Back Active and given the Back Book to read later. The low risk group received no further therapy. Therapists providing this level of intervention received one full day of training to standardise the approach.

The medium risk group was referred for physiotherapy-led treatment focused on symptoms and function. These physiotherapists were provided with three days of training to standardise the approach applied. The high risk group were referred for psychologically informed physiotherapy focusing on addressing psychosocial obstacles to recovery. These therapists were provided with six full days of training to standardise the approach taken.

The control group participants received usual care, with control group physiotherapists provided with half a day of training to ensure familiarity with the study protocol.

Outcomes

Study outcomes are presented in Tables 17 and 18. Statistically significant results were found in favour of the intervention arm for the medium and high risk groups on the RMDQ at 4 months. No significant difference was found in the low risk group at 4 months. At the 12 month follow up, this difference was maintained in the medium risk group, but not the high risk group. Effect sizes were small to moderate, with the largest effects seen in the high risk group. Outcomes consistently favoured the intervention arms and days off work was significantly different across the groups for both 4 and 12 month follow ups.

Table 17. Primary outcomes: Mean change in scores with the Roland and Morris Disability Questionnaire during 4-month and 12–month follow-ups. Highlighted cells indicate statistically significant outcomes

	All participants	p value	Low-risk participants	p value	Medium-risk participants	p value	High–risk participants	p value
Intervention group	568	..	148	..	263	..	157	..
Control group	283	..	73	..	131	..	79	..
4 months								
Mean change (SD)								
Intervention group	4·7 (5·9)	..	1·6 (4·4)	..	5·3 (6·0)	..	6·8 (6·9)	..
Control group	3·0 (5·9)	..	0·8 (4·3)	..	3·4 (6·1)	..	4·4 (6·1)	..
Mean difference* (95% CI)	1·81 (1·06 to 2·57)	<0·0001	0·66 (−0·41 to 1·74)	0·2211	1·99 (0·75 to 3·22)	0·0012	2·53 (0·90 to 4·16)	0·0024
Sensitivity analysis†	1·54 (0·77 to 2·30)	<0·0001	0·66 (−0·42 to 1·74)	0·2275	1·59 (0·49 to 2·69)	0·0048	2·28 (0·49 to 4·07)	0·0129
Sensitivity analysis‡	1·83 (1·01 to 2·65)	<0·0001	0·74 (−0·50 to 1·98)	0·2447	1·99 (0·73 to 3·25)	0·0021	2·57 (0·81 to 4·32)	0·0045
Effect size§ (95% CI)	0·32 (0·19 to 0·45)	<0·0001	0·19 (−0·12 to 0·51)	0·2211	0·43 (0·16 to 0·70)	0·0012	0·56 (0·20 to 0·92)	0·0024
Participants with good outcome¶								
Intervention group	392 (69%)	..	100 (68%)	..	186 (71%)	..	106 (68%)	..
Control group	158 (56%)	..	42 (58%)	..	75 (57%)	..	40 (51%)	..
Odds ratio- (95% CI)	1·85 (1·36 to 2·51)	<0·0001	1·58 (0·83 to 3·00)	0·1618	2·00 (1·20 to 3·33)	0·0073	1·96 (1·03 to 3·71)	0·0390
Number needed to treat** (95% CI)	7·0 (13·5 to 4·9)	<0·0001	9·5 (−21·8 to 4·4)	0·1618	6·5 (22·8 to 4·1)	0·0073	6·2 (135·0 to 3·5)	0·0390
12 months								
Mean change (SD)								
Intervention group	4·3 (6·4)	..	1·6 (4·5)	..	4·9 (5·9)	..	5·9 (7·2)	..
Control group	3·3 (6·2)	..	1·2 (4·8)	..	3·6 (6·3)	..	4·8 (6·3)	..

	All participants	p value	Low-risk participants	p value	Medium-risk participants	p value	High–risk participants	p value
Mean difference* (95% CI)	1·06 (0·25 to 1·86)	0·0095	0·12 (−1·13 to 1·38)	0·8456	1·33 (0·15 to 2·52)	0·0253	1·22 (−0·47 to 2·91)	0·1547
Sensitivity analysis†	1·06 (0·26 to 1·87)	0·0099	0·29 (−0·86 to 1·44)	0·6177	1·18 (0·01 to 2·34)	0·0482	1·30 (−0·54 to 3·13)	0·1653
Sensitivity analysis‡	1·08 (0·18 to 1·98)	0·0186	0·06 (−1·37 to 1·50)	0·9298	1·35 (0·12 to 2·58)	0·0317	1·56 (−0·72 to 3·84)	0·1710
Effect size§ (95% CI)	0·19 (0·04 to 0·33)	0·0095	0·04 (−0·33 to 0·41)	0·8456	0·29 (0·03 to 0·55)	0·0253	0·27 (−0·10 to 0·65)	0·1547
Participants with good outcome¶								
Intervention group	367 (65%)	..	98 (66%)	..	176 (67%)	..	93 (59%)	..
Control group	160 (57%)	..	44 (60%)	..	77 (59%)	..	39 (49%)	..
Odds ratio- (95% CI)	1·48 (1·02 to 2·15)	0·0344	1·22 (0·60 to 2·47)	0·5754	1·55 (0·93 to 2·57)	0·0874	1·57 (0·82 to 3·00)	0·1723
Number needed to treat** (95% CI)	10·8 (206·0 to 5·8)	0·0344	21·4 (−7·9 to 5·4)	0·5754	10·0 (−56·5 to 5·1)	0·0874	9·0 (−20·2 to 4·0)	0·1723

Data are number (%), unless otherwise indicated. RMDQ=Roland and Morris Disability Questionnaire. *Mean for the intervention group minus mean for the control group (by use of linear regression adjusted for age, sex, baseline RMDQ, and duration of back pain).†Based on observed or available case data.‡Based on imputed datasets with adjustment for clustering by therapist (by use of random-effects linear regression modelling).§Mean difference relative to the pooled SD of baseline scores.¶Defined as at least 30% change in the RMDQ score compared with baseline; the numbers represent average rounded counts of five imputed datasets. Odds of a good outcome in the intervention group relative to the control group (by use of binary logistic regression with adjustment for age, sex, baseline RMDQ, and duration of back pain).**Based on the point estimate for the proportion of patients with good outcomes in the control group and the odds ratio for good outcomes in the intervention group relative to the control group; smaller positive numbers for the 95% CI convey a stronger association than do larger positive numbers and smaller negative numbers indicate a greater advantage towards the control group, and therefore the number needed to treat does not necessarily lie within the 95% CI.

Table 18. Secondary outcome measures at follow-ups of 4 months and 12 months. Highlighted cells indicate statistically significant outcomes.

	All participants				Low–risk participants				Medium–risk participants				High–risk participants			
	Intervention group (n=568)	**Control group (n=283)**	**Point estimate (95% CI)**	**p value**	**Intervention group (n=148)**	**Control group (n=73)**	**Point estimate (95% CI)**	**p value**	**Intervention group (n=263)**	**Control group (n=131)**	**Point estimate (95% CI)**	**p value**	**Intervention group (n=157)**	**Control group (n=79)**	**Point estimate (95% CI)**	**p value**
Back pain intensity																
4 months, mean change (SD)	3·2 (2·5)	2·6(2·4)	0·55[*] (0·23 to 0·86)	0·0005	1·7 (2·2)	1·5(2·1)	0·28[*] (−0·29 to 0·84)	0·3182	3·5 (2·6)	2·8(2·1)	0·58[*] (0·06 to 1·10)	0·0200	4·2 (2·3)	3·4(2·9)	0·73[*] (0·04 to 1·42)	0·0316
12 months, mean change (SD)	3·0 (2·8)	2·8(2·6)	0·11[*] (−0·21 to 0·43)	0·5002	1·7 (2·3)	1·7(2·4)	0·10[*] (−0·51 to 0·70)	0·7461	3·3 (2·6)	3·0(2·8)	0·16[*] (−0·44 to 0·76)	0·5802	3·7 (2·7)	3·6(3·2)	−0·01[*] (−0·72 to 0·70)	0·9755
STarT Back risk group																
4 months	..	..	2·09[†] (1·40 to 3·12)	0·0003	..	..	1·16[†] (0·36 to 3·67)	0·8021	..	..	1·70[†] (0·95 to 3·03)	0·0678	..	..	3·14[†] (1·63 to 6·06)	0·0006
Low risk	439 (77%)	186 (66%)	..	..	134 (91%)	65 (89%)	..	..	201 (76%)	89 (68%)	..	..	104 (66%)	32 (41%)	..	..
Medium risk	103 (18%)	70(25%)	..	..	12 (8%)	7(10%)	..	..	55 (21%)	36(27%)	..	..	36 (23%)	27(34%)	..	..
High risk	26 (5%)	27(10%)	..	..	2 (1%)	1(1%)	..	..	7 (3%)	6(5%)	..	..	17 (11%)	20(25%)	..	..
12 months	..	..	1·51[†] (1·01 to	0·0405	..	..	0·91[†] (0·31 to	0·8562	..	..	1·51[†] (0·85 to	0·1525	..	..	1·62[†] (0·82 to	0·1566

	All participants				Low–risk participants				Medium–risk participants				High–risk participants			
	Intervention group (n=568)	Control group (n=283)	Point estimate (95% CI)	p value	Intervention group (n=148)	Control group (n=73)	Point estimate (95% CI)	p value	Intervention group (n=263)	Control group (n=131)	Point estimate (95% CI)	p value	Intervention group (n=157)	Control group (n=79)	Point estimate (95% CI)	p value
			2·25)				2·65)				2·66)				3·23)	
Low risk	430 (76%)	195 (69%)	..	..	131 (89%)	65 (89%)	..	..	204 (78%)	94 (72%)	..	..	95 (61%)	36 (46%)	..	..
Medium risk	111 (20%)	71 (25%)	..	..	17 (11%)	7(10%)	..	..	56 (21%)	35(27%)	..	..	38 (24%)	29(37%)	..	..
High risk	27 (5%)	17(6%)	..	..	0	1(1%)	..	..	3 (1%)	2(2%)	..	..	24 (15%)	14(18%)	..	..
Global change																
4 months	..	..	1·55† (1·16 to 2·07)	0·0032	..	..	1·45† (0·80 to 2·60)	0·2153	..	..	1·57† (1·04 to 2·36)	0·0304	..	..	1·70† (0·98 to 2·96)	0·0595
Much better	252 (44%)	99(35%)	..	..	70 (47%)	28(38%)	..	..	125 (48%)	50(38%)	..	..	57 (36%)	21(27%)	..	..
Better	151 (27%)	82(29%)	..	..	38 (26%)	23(32%)	..	..	70 (27%)	40(31%)	..	..	43 (27%)	19(24%)	..	..
Not better	165 (29%)	102(36%)	..	..	40 (27%)	22(30%)	..	..	68 (26%)	41(31%)	..	..	57 (36%)	39(49%)	..	..
12 months	..	..	1·23† (0·95 to 1·60)	0·1104	..	..	0·80† (0·41 to 1·54)	0·4812	..	..	1·54† (0·99 to 2·39)	0·0520	..	..	1·26† (0·72 to 2·21)	0·4072
Much better	229 (40%)	99(35%)	..	..	57 (39%)	32(44%)	..	..	116 (44%)	46(35%)	..	..	56 (36%)	21(27%)	..	..
Better	121 (21%)	63(22%)	..	..	36 (24%)	17(23%)	..	..	59 (22%)	31(24%)	..	..	26 (17%)	15(19%)	..	..
Not	218 (38%)	121(43%)	..	..	55 (37%)	24(33%)	..	..	88 (33%)	54(41%)	..	..	75 (48%)	43(54%)	..	..

	All participants				Low–risk participants				Medium–risk participants				High–risk participants			
	Intervention group (n=568)	Control group (n=283)	Point estimate (95% CI)	p value	Intervention group (n=148)	Control group (n=73)	Point estimate (95% CI)	p value	Intervention group (n=263)	Control group (n=131)	Point estimate (95% CI)	p value	Intervention group (n=157)	Control group (n=79)	Point estimate (95% CI)	p value
better																
PCS–catastrophising																
4 months, mean change (SD)	6·4 (10·8)	3·9(10·4)	2·35* (0·91 to 3·79)	0·0010	2·6 (7·3)	1·3(8·0)	1·10* (−0·82 to 3·02)	0·2590	5·9 (10·1)	3·6(10·1)	1·88* (−0·09 to 3·84)	0·0531	10·8 (12·1)	6·9(13·3)	3·91* (0·41 to 7·41)	0·0242
12 months, mean change (SD)	6·1 (10·6)	4·4(11·4)	1·69* (0·37 to 3·01)	0·0113	3·2 (8·6)	1·6(7·7)	1·47* (−0·60 to 3·54)	0·1516	5·7 (9·0)	3·6(10·5)	1·74* (−0·11 to 3·60)	0·0630	9·4 (12·1)	8·1(12·7)	1·35* (−1·70 to 4·40)	0·3816
TSK–fear avoidance																
4 months, mean change (SD)	5·5 (7·0)	3·2(6·0)	2·52* (1·61 to 3·43)	<0·0001	3·5 (7·1)	2·8(7·3)	0·86* (−1·01 to 2·73)	0·3540	5·2 (7·9)	2·8(5·7)	2·72* (1·45 to 3·98)	<0·0001	7·9 (8·0)	4·3(6·0)	3·70* (1·70 to 5·70)	0·0003
12 months, mean change (SD)	5·2 (7·3)	3·3(10·3)	2·18* (0·73 to 3·63)	0·0013	3·5 (7·2)	2·4(9·3)	1·22* (−1·26 to 3·70)	0·3030	4·9 (7·6)	3·2(8·2)	2·14* (0·23 to 4·04)	0·0181	7·3 (7·7)	4·3(7·2)	3·12* (1·18 to 5·06)	0·0016
HADS, anxiety subscale																

	All participants				Low–risk participants				Medium–risk participants				High–risk participants			
	Intervention group (n=568)	Control group (n=283)	Point estimate (95% CI)	p value	Intervention group (n=148)	Control group (n=73)	Point estimate (95% CI)	p value	Intervention group (n=263)	Control group (n=131)	Point estimate (95% CI)	p value	Intervention group (n=157)	Control group (n=79)	Point estimate (95% CI)	p value
4 months, mean change (SD)	1·7 (3·6)	1·2(4·0)	0·63* (0·08 to 1·18)	0·0208	0·6 (3·3)	0·9(3·5)	−0·17* (−0·97 to 0·62)	0·6666	1·7 (3·8)	0·8(3·7)	1·10* (0·33 to 1·86)	0·0039	2·8 (4·3)	2·2(4·5)	0·56* (−0·57 to 1·69)	0·3155
12 months, mean change (SD)	1·3 (3·9)	1·0(4·4)	0·45* (−0·10 to 1·01)	0·1035	0·5 (3·2)	0·8(4·0)	−0·13* (−1·22 to 0·96)	0·8023	1·3 (4·2)	0·6(4·2)	0·88* (0·08 to 1·68)	0·0283	2·1 (4·5)	1·7(5·0)	0·28* (−1·14 to 1·71)	0·6783
HADS, depression subscale																
4 months, mean change (SD)	1·7 (3·7)	1·1(3·3)	0·71* (0·26 to 1·15)	0·0018	0·3 (3·2)	0·2(3·3)	0·09* (−0·69 to 0·88)	0·8086	1·7 (3·6)	1·2(3·5)	0·79* (0·13 to 1·45)	0·0193	3·0 (4·3)	1·9(3·8)	1·13* (0·16 to 2·09)	0·0223
12 months, mean change (SD)	1·4 (4·1)	0·9(4·0)	0·62* (0·07 to 1·17)	0·0227	0·2 (3·3)	0·2(3·5)	−0·03* (−0·92 to 0·86)	0·9494	1·3 (3·7)	1·0(3·8)	0·63* (−0·10 to 1·36)	0·0863	2·7 (4·7)	1·5(4·5)	1·21* (0·02 to 2·40)	0·0413
SF12, physical component																
4 months, mean change (SD)	−7·5 (13·0)	−5·2(13·3)	−2·83* (−4·62 to −1·05)	0·0010	−3·2 (9·6)	−1·8(9·7)	−1·50* (−3·89 to 0·89)	0·2137	−9·1 (11·7)	−6·4(10·7)	−3·27* (−5·51 to −1·03)	0·0037	−8·9 (15·1)	−6·4(15·8)	−3·13* (−7·50 to 1·23)	0·1215

	All participants				Low–risk participants				Medium–risk participants				High–risk participants			
	Intervention group (n=568)	Control group (n=283)	Point estimate (95% CI)	p value	Intervention group (n=148)	Control group (n=73)	Point estimate (95% CI)	p value	Intervention group (n=263)	Control group (n=131)	Point estimate (95% CI)	p value	Intervention group (n=157)	Control group (n=79)	Point estimate (95% CI)	p value
12 months, mean change (SD)	−7·5 (11·3)	−5·2(10·9)	−2·93* (−4·31 to −1·56)	<0·0001	−4·0 (9·7)	−2·4(10·1)	−1·70* (−4·25 to 0·85)	0·1825	−8·8 (11·5)	−5·7(11·7)	−3·74* (−6·02 to −1·46)	0·0012	−8·6 (12·2)	−6·8(13·1)	−2·46* (−5·90 to 0·97)	0·1464
SF12, mental component																
4 months, mean change (SD)	−2·1 (11·3)	−2·1(11·0)	−0·16* (−1·63 to 1·32)	0·8327	0·5 (12·4)	−1·0(10·2)	0·96* (−1·68 to 3·60)	0·4640	−1·5 (10·4)	−1·1(12·0)	−0·80* (−2·92 to 1·31)	0·4497	−5·5 (12·5)	−4·8(14·3)	−0·27* (−3·98 to 3·45)	0·8811
12 months, mean change (SD)	−1·7 (13·0)	−1·2(13·4)	−0·69* (−2·39 to 1·01)	0·4050	1·3 (10·7)	−0·4(9·9)	1·18* (−1·12 to 3·49)	0·3124	−1·2 (12·3)	−0·1(12·7)	−1·53* (−3·62 to 0·55)	0·1468	−5·5 (13·8)	−3·6(13·8)	−1·31* (−5·04 to 2·43)	0·4682
Work loss‡																
Days off work, mean (SD)	4·4 (21·2)	12·2(35·1)	2·76§ (2·52 to 3·01)	<0·0001	0·4 (1·2)	3·0(11·9)	7·06§ (4·28 to 11·6)	<0·0001	4·1 (15·8)	18·4(47·2)	5·17§ (4·52 to 5·92)	<0·0001	9·9 (35·4)	10·6(18·2)	1·46§ (1·24 to 1·72)	<0·0001
Satisfaction with care¶	..	..	1·98† (1·45 to 2·70)	<0·0001	..	..	1·24† (0·68 to 2·25)	0·4784	..	..	2·15† (1·35 to 3·40)	0·0013	..	..	3·18† (1·68 to 6·00)	0·0005
Very	166/416	46/197(2	..	..	30/106	14/56(25	..	..	87/197	22/91(24	..	..	49/113	10/50(20	..	..

	All participants				Low–risk participants				Medium–risk participants				High–risk participants			
	Intervention group (n=568)	Control group (n=283)	Point estimate (95% CI)	p value	Intervention group (n=148)	Control group (n=73)	Point estimate (95% CI)	p value	Intervention group (n=263)	Control group (n=131)	Point estimate (95% CI)	p value	Intervention group (n=157)	Control group (n=79)	Point estimate (95% CI)	p value
satisfied	(40%)	3%)			(28%)	%)			(44%)	%)			(43%)	%)		
Quite satisfied'	139/416 (33%)	81/197(41%)	..	..	33/106 (31%)	21/56(38%)	..	..	63/197 (32%)	39/91(43%)	..	..	43/113 (38%)	21/50(42%)	..	..
Not satisfied	111/416 (27%)	70/197(36%)	..	..	43/106 (41%)	21/56(38%)	..	..	47/197 (24%)	30/91(33%)	..	..	21/113 (19%)	19/50(38%)	..	..

Data are number (%), unless otherwise indicated. PCS=Pain Catastrophizing Scale. TSK=Tampa Scale of Kinesiophobia. HADS=Hospital Anxiety and Depression Scale. SF12=Short Form 12. RMDQ=Roland and Morris Disability Questionnaire. * Mean difference—ie, mean for the intervention group minus mean for the control group (by use of linear regression adjusted for age, sex, baseline score, baseline RMDQ, and duration of back pain). † Odds ratio represents the proportional odds in terms of ordered response categories for the intervention group relative to the control group (by use of ordinal logistic regression adjusted for age, sex, baseline RMDQ, and duration of back pain). ‡ Analysis of time off work (days) is based on a total subsample of 298 of 567 responders who reported being currently employed at 12 months follow-up, and relates to leave due to low back pain during the period between baseline and follow-up at 12 months. § Incidence rate ratio for the extra work days lost in the control group relative to the intervention group (by use of Poisson regression adjusted for age, sex, baseline RMDQ, and duration of back pain). ¶ Analysis was based on 613 responders to the question about satisfaction with care at 4 months follow-up: 416 in the intervention group and 197 in the control group.

Jull et (2013) – Neck Disability Index

Patients were stratified into groups based on 3 factors: pain and disability (NDI score), symptoms of moderate post stress (IES score), and abnormal sensory responses. The NDI and IES are described in the embedded spreadsheet above. This information was collected for people within four weeks of sustaining a whiplash injury.

Other information collected included pain on a visual analogue scale (VAS), the General Health Questionnaire, Pictorial Fear of Activity Scale and range of motion (ROM) for flexion/extension and rotation.

Quality Appraisal

Quality appraisal was performed using the Cochrane assessment of bias. Jull et al (2013) was deemed to contain a high risk of bias (Table 19).

Table 19. Quality appraisal of Jull et al (2013)

	Sequence generation	Allocation concealment	Blinding of participants	Incomplete outcome data	Selective outcome reporting	Other sources of bias	Overall assessment
Jull et al (2013)	No	Yes	No	No	Yes	No	High risk of bias

What information was collected

Patients were stratified into groups based on 3 factors: pain and disability (NDI score), symptoms of moderate post stress (IES score), and abnormal sensory responses. The NDI and IES are described in the embedded spreadsheet above. This information was collected for people within four weeks of sustaining a whiplash injury.

Other information collected included pain on a visual analogue scale (VAS), the General Health Questionnaire, Pictorial Fear of Activity Scale and range of motion (ROM) for flexion/extension and rotation.

How was the information interpreted

Based on the information collected, participants were allocated to four subgroups: 1. NDI score <30%; 2. NDI score 30% and greater; 3. IES score greater than 26; 4. Sensory disturbance positive on 2 out of three measures (thermal pain threshold, pressure pain threshold, sympathetic vasoconstrictor response. How these groups defines the nature of the intervention applied in the intervention group is not clear. It appears that the information may have influenced the treating practitioners decisions related to treatment, but the stratification into groups did not seem to impact the referral patterns applied in the intervention group – i.e. pragmatic care was applied that could have been very similar to usual care. It is possible to determine that a high IES score (greater than 26) did result in a six week cognitive behavioural therapy (CBT) program.

What action was initiated

The intervention is described as ‘pragmatic care’ and participants could receive any combination of medical, physiotherapeutic and psychological care. The nature and extent of treatments prescribed was based on the presenting sensory, sensorimotor and psychological impairments, although how these factors influenced most treatments is unclear. Participants with an IES score greater than 26 were referred to a 6 week CBT therapy program. At a minimum, treatment included medical consultation and physiotherapy management for pain and physical symptoms associated with the injury.

Medical management was provided by a general practitioner in the trial centre. Physiotherapy was provided by experienced physiotherapists located in the community. Psychological management was provided by experienced psychologists located in the community. The intervention description suggests that treatment decisions were at the discretion of the treating practitioner from set options, however it is not clear how often these options were applied or adhered to.

Outcomes

Outcome data for the primary outcomes are provided in Table 20. Due to the high likelihood of bias in study results, only primary outcomes are presented in detail. No significant differences were found between the intervention and control groups for any outcomes at any of the follow up points.

It is not clear how many participants filled each of the diagnostic categories and therefore it is not possible to determine if the results were due to a lack of treatment effect of the intervention or reduced statistical power due to a small sample size. The analysis applied combined the results across the different diagnostic groups and may have reduced the ability to detect differences in the treatment groups applied. As a result, it is not possible to determine the overall effectiveness of the method of targeted care applied in this study.

Table 20. Primary outcomes (Neck Disability Index) for Jull et al (2013)

Reference	N	Outcomes	Baseline –		Post care –			Estimated treatment effect (95% CI)
Jull et al 2013	101 (baseline) [I:49;C:52]	NDI	Intervention	Control	11 weeks	6 months	12 months	
			36.0 (18.7) (n=49)	29.8 (16.9) (n=43)	I: 19.3 (17.9) (n=43) C: 16.1 (16.9) (n=36	I: 16.8 (15.2) (n=44) C: 13.5 (14.4) (n=43)	I: 16.9 (15.3) (n=45) C:13.5 (15.4) (n=47)	5.33 (-0.46-11.13) No difference between groups

No differences were found between groups on other outcomes of pain, Impact of Events Scale score, General Health Questionnaire, Pictorial Fear of Activity Scale, Pressure Pain Threshold or Range of Motion, ($p<0.05$)

DISCUSSION

The greatest unknown in relation to targeting remains the question of how to implement action based on the collection and interpretation of individual-level information. A wide range of evidence-based tools for the collection of information exist, but even in the case of more frequently applied tools such as the OMPQ, recommending a decision point to trigger action is not yet possible. Even if it were possible to recommend a particular score to commence intervention, the nature of the intervention to apply is difficult to determine when the tool is comprised of a mixture of constructs.

Despite the wide variety of evidenced-based tools available in the literature, this review found an absence of any clear patterns of use that could guide decision making related to the ideal information collection methods to be used. For example, characteristics of tools such as the length of the tool, the constructs measured by the tool and the level of research support for its measurement properties along with utilisation factors such as the timing of application and the setting did not provide clear indications of better or worse predictive performance. Therefore, it appears implementation factors, such as ease of application in a case management setting, should guide tool selection. The SBST provides a clear example of a tool that was implemented effectively with positive outcomes for people with low back pain. Even though analysis of the tool's measurement properties suggest the ability to differentiate pain and disability outcomes are modest at best, the stratified intervention improved pain, disability and work outcomes for medium and high risk participants.

Studies that allow evaluation of the impact of targeted care are rare. Of the two examples located only one allowed identification of the true impact of applying targeted care. Hill et al (2011) demonstrated improved outcomes when using the SBST to stratify care. This study only consisted of people with low back pain and was conducted in primary care outside of a compensation structure. It is unclear whether the tool can be applied to other conditions (even with condition specific modification) or in a compensation setting. One of the major challenges to producing evidence of this kind is the cost and time required to conduct such a trial, and including multiple conditions greatly increases the number of participants required to allow appropriate analysis. This is likely one reason behind why the volume of research related to targeted care is focused on the nature of the information collected rather than the element of action. Another reason could be that implementation occurs in a manner that is less conducive to academic research, or is applied in an environment where results are not published to protect commercial sensitives of insurance companies attempting to improve models of management.

A remaining challenge to providing models of targeted care is the effectiveness of the interventions applied. The large number of tools located and described in this review suggest that identification of important factors in recovery from injury is not only possible, but can be performed in a number of ways. What is less clear is whether interventions aimed at some of the factors typically identified are effective in reducing the factor in question. A number of tools included in this review are diagnostic of conditions (e.g. depression) and using the tool early in the course of the injury is found to be predictive of a high score on the tool at a later follow up. It may be that the interventions applied in the period between assessments is ineffective in addressing the problem so high scores on the diagnostic tools are maintained.

The provision of targeted care remains a challenge of implementation. When considering introducing models of targeted care, it appears a sensible approach would be to apply the Collect, Interpret, Act framework to the intention of the model of care. For example, if the aim is to address depression in claimants injured in motor vehicle accidents, there are a number of potential tools that could be applied. Matching the available tools to key implementation factors (e.g. will the tool be applied by a case manager or a treating practitioner) seems a better approach than relying on

factors related to measurement properties. Furthermore, identifying an action (such as a referral to a specialised program) that is available and readily applied for the majority of cases will guide interpretation and the steps required to initiate action.

IMPLICATIONS FOR THE TAC

This evidence review identified a lack of high level evidence to guide TAC decision making related to identifying and responding to client needs using existing tools. The implication of these findings for the TAC are discussed below in terms of both identifying client need and responding appropriately (targeting care).

Identifying need:

- A range of validated tools can be applied to identifying client need, some with high level predictive ability and research support. However, no clear factors emerged to clearly direct tool selection for use in practice.

  Recommendation: factors related to implementation should guide decisions on appropriateness of a validated tool for identifying client need.

- There are gaps in the evidence around tools to identify client need. Many tools have been developed but not rigorously tested in an applied setting. At this stage the evidence does not exist to support some factors as predictors of outcomes. Absence of evidence is not evidence of absence, so there may be further constructs and tools that may be important in the client need identification context, but the evidence does not yet exist to support their application.

  Recommendation: the application of validated instruments to identify relevant constructs remains a valid strategy for TAC to pursue. Should other constructs be deemed important in the identification of client need, the TAC should endeavour to contribute to the evidence base for the tool or construct through careful evaluation of the measures applied.

- The complexity of identifying client need was demonstrated by the number and variety of the tools identified in this review. Limited information was available related to the resources required to implement methods of identifying client needs in practice.

  Recommendation: it is reasonable to expect that any approach to client need identification will require training for employees directly interacting with claimants and this should be considered in the development of a client need approach.

Responding to client need:

- There is limited evidence to determine effectiveness of targeted care approaches. Only one example allowed evaluation of the impact of the approach applied, which was specific to low back pain applied outside a compensation setting. The evidence does not yet exist to clearly support a targeted care approach.

  Recommendation: an approach applying validated tools to identify and respond to client needs in order to provide early and appropriate intervention is an appropriate pathway to providing an evidence-based model of care. The TAC should endeavour to contribute to the evidence base and ensure that evaluation of the effectiveness of the model applied is possible.

- No examples exist that allow evaluation of a targeted care approach in a compensation setting. There is an opportunity to implement a world-leading approach and set the standard for research in this area.

  Recommendation: Trials should be set up that allow for evaluation of the impact of the effectiveness of the approach implemented. Evaluation should form an integral part of the planning process for pilot trials and larger scale implementation.

CONCLUSION

Evidence does not yet exist related to the implementation of targeted approaches to care in a compensation setting. There are a number of validated tools appropriate for identifying client needs, however the effectiveness of using these tools as a basis for intervention has not been rigorously evaluated.

In the absence of clear indicators for the selection of tools to identify client needs, implementation factors (such as the length of the tool, the skill of the person collecting the information and the avenues available to act on the information collected) should guide the nature of the information collected.

Incorporation of evaluation into the design of a new approach to providing care will not only reveal the effectiveness of the approach applied, but simultaneously build the evidence base behind targeted models of care.

APPENDIX 1: COLLECT, INTERPRET, ACT FRAMEWORK

Collection	Who was the target?	Being clear on who the candidates for targeting were is essential as it determines the overall approach to be used. For example, being clear on who care is being targeted to will assist in the decision between an approach that identifies high-level issues (e.g. need to prioritise management of a claim) vs identifying specific service needs (e.g. identifying whether referral to psychologist is required) Once the target has been identified, the aims of targeting care can include triage for claim allocation, identification of problem claims to manage costs, to target appropriate intervention or to stratify for treatment.
	What information was collected?	Typically targeted care includes some measure of psychosocial aspects (e.g. depression, recovery expectations or fear avoidance beliefs) that indicate a higher risk of poor outcomes. The inherent complexity of workers' compensation suggests that targeting care based on information covering a single domain is not likely to be as effective as addressing multiple areas (e.g. work-related, socio-demographic and health-related factors) that may impact outcomes
	When was information collected?	Time influences many of the aspects typically targeted in providing tailored care (e.g. development of secondary psychosocial problems). Some tools used in approaches of tailoring care are developed to be applied a certain time after injury (e.g. Orebro was developed to be applied 6 weeks after injury).
	How was it collected?	Questionnaires are generally easiest to administer, but offer little flexibility in how the content is described or administered. Conversational sources offer maximum flexibility but vary in consistency. Administrative sources are restricted to the variables gathered for the purpose and often do not capture the type of information required for targeting care.
	Who was it collected by?	The skill level of the person collecting the information may determine the nature of the information collected. A person who has had the opportunity to develop rapport with the target may be in a better position to collect more sensitive or personal information.
	What resources were used in collection?	Gathering new information on all claims requires more resources compared to those who are yet to return to work after three or four weeks. Interview-based tools will provide the greatest depth of information, but take longer to administer.
Interpretation	What was the output?	If the output is an overall score, cut offs must be determined. At some point the result of gathering the information must be converted into a format that can be applied to determining appropriate action.

	Who interpreted the output?	This may be the person who performed the scoring of a questionnaire. Treating practitioners may have the highest level of skill in terms of interpreting outcomes of an assessment, but present difficulties in ensuring the consistent application in practice. Case managers may represent consistent points of contact in order to tailor care, but high turnover and varying ability to interpret findings may impact results.
	What was needed to link output to action?	Commencing action may require the input of another health professional. A summary of treatment to date may need to be provided as part of referral.
Action	Who initiated action?	A health practitioner may be able to initiate action immediately. Alternatively it may be another party, such as a case manager, who commences the required action.
	When did action commence?	Interventions for musculoskeletal conditions are suggested to be more effective in the sub-acute phase of the condition. The processes leading up to action should be of an appropriate length so that action is appropriately timed.
	What resources were needed for action?	Some interventions are more expensive than others and consideration of the resources available may determine whether the action proposed will suit other purposes.

APPENDIX 2: INITIAL SEARCH STRATEGY

Searches	Results: Ovid MEDLINE(R) 1946 to Present with Daily Update	Type
1	Musculoskeletal conditions.mp.	1121
2	Musculoskeletal diseases.mp.	11047
3	Musculoskeletal disorders.mp.	4175
4	Musculoskeletal diseases/ or musculoskeletal diseases/ or fasciitis, plantar/ or foot deformities, acquired/ or heel spur/ or posterior tibial tendon dysfunction/ or hand deformities, acquired/ or exp temporomandibular joint disorders/ or bursitis/ or joint deformities, acquired/ or joint instability/ or joint loose bodies/ or patellofemoral pain syndrome/ or shoulder impingement syndrome/ or synovitis/ or compartment syndromes/ or anterior compartment syndrome/ or ischemic contracture/ or contracture/ or dupuytren contracture/ or muscle cramp/ or myofascial pain syndromes/ or exp tendinopathy/ or tennis elbow/	86315
5	musculoskeletal pain/ or exp back pain/ or chronic pain/ or neck pain/ or pain, intractable/	50319
6	((exp arm injuries/ or exp back injuries/ or contusions/ or exp dislocations/ or exp fractures, bone/ or fractures, cartilage/ or exp hand injuries/ or exp hip injuries/ or exp leg injuries/ or exp neck injuries/ or occupational injuries/ or soft tissue injuries/ or exp spinal injuries/ or exp sprains/) and "strains"/) or exp tendon injuries/	21145
7	((Pain* or tear or tears or injur* or sprain* or strain* or dislocation*) adj (musc* or joint or back or spine or spinal or neck or cervical or pelvic or hip or rotator cuff or knee or ankle or elbow or shoulder)).mp.	8885
8	(carpal tunnel or frozen shoulder or shoulder impingement or chronic pain or myofascial pain or patellofemoral pain or regional pain disorder* or whiplash).mp.	43966
9	Depression/ or Depressive Disorder/ or Mental Disorders/	287910
10	Stress Disorders, Post-Traumatic/ or Stress Disorders, Traumatic, Acute/	25478
11	Anxiety/ or Anxiety Disorders/	88051
12	or/1-11	544767
13	(osteoporosis or (diabet* and ulcer*) or fibromyalgia or ankylosing spondilytis or RA or arthritis or osteomyelitis).ti.	117600
14	exp *Osteoporosis/	35201
15	exp *Diabetic Foot/	5807
16	exp *Fibromyalgia/	6248
17	exp *Arthritis/	187432
18	exp *Osteomyelitis/	16183
19	or/13-18	252624
20	12 not 19	529879
21	(Motor vehicle accident or motor vehicle collision or MVC or motor vehicle crash or work accident or injury).mp.	530329
22	accidents/ or accidents, occupational/ or accidents, traffic/	69176
23	(wounds and injuries).mp.	113571
24	"wounds and injuries"/ or abdominal injuries/ or amputation, traumatic/ or arm injuries/ or back injuries/ or fractures, bone/ or fractures, cartilage/ or hand injuries/ or hip injuries/ or leg injuries/ or neck injuries/ or occupational injuries/ or shock, traumatic/ or soft tissue injuries/ or spinal cord injuries/ or "sprains and strains"/ or tendon injuries/ or trauma, nervous system/	220041
25	"Compensation and Redress"/	2685
26	or/21-24	742015
27	or/21-25	744111
28	screening.mp.	418876
29	(primary prevention or early prevention or secondary prevention).mp.	51565
30	Decision Support Systems, Clinical/	5901
31	decision making, computer assisted/ or decision support.mp.	27605

Searches	Results: Ovid MEDLINE(R) 1946 to Present with Daily Update	Type
32	decision making/ and (model or models or classification or subgroup* or sub-group* or algorithm*).mp.	14629
33	decision support.mp.	25908
34	clinical prediction rule*.mp.	753
35	decision system*.tw.	166
36	treatment based classification.tw.	32
37	knowledge-base*.tw.	9177
38	treatment rule*.tw.	79
39	treatment selection.tw.	1894
40	targeted treatment*.tw.	2850
41	(treatment algorithm* or management algorithm*).tw.	4989
42	(orebro adj4 (musculoskeletal or questionnaire* or pain*)).tw.	40
43	STarT Back.tw.	39
44	Acute Low Back Pain Screening.tw.	9
45	((support* or guide or aid* or rule* or tool*) adj4 decision).tw.	22306
46	clinical decision-making/ or diagnosis, computer-assisted/ or "diagnostic techniques and procedures"/ or mass screening/	113332
47	triage/ or preventive health services/ or rehabilitation/	38405
48	prognostic model.mp.	1398
49	risk screening.mp.	1206
50	Pain Measurement/ or Low Back Pain/ or stratified care.mp. or Physical Therapy Modalities/ or Primary Health Care/	176180
51	(stratification or stratified medicine).mp.	36844
52	subgroup.tw.	86908
53	(family practice or primary care).mp.	134460
54	Yellow Flags.tw.	45
55	Brief Pain Inventory.tw.	1083
56	Kessler 10.tw.	131
57	DASS 21.tw.	152
58	Pain Catastrophizing Scale.tw.	316
59	Pain Self-efficacy Questionnaire.tw.	52
60	Roland-Morris Disability Questionnaire.tw.	518
61	Recovery expectation*.tw.	57
62	or/28-61	954487
63	reliability.mp.	107743
64	validity.mp.	114802
65	validation.mp.	172802
66	sensitivity.mp.	867929
67	(specificity and sensitivity).mp.	401065
68	(positive predictive values or negative predictive values).mp.	12030
69	"Reproducibility of Results"/ or "Surveys and Questionnaires"/	658489
70	accuracy.mp.	242900
71	or/63-70	1709051
72	(return to work or sick leave or disability leave or recovery or quality of life or pain or psychological injury or distress or anxiety or depression or PTSD or driving phobia).mp.	1449920
73	Intervertebral Disc Displacement/ or Return to Work/ or Occupational Diseases/ or Employment/ or Low Back Pain/	151043
74	Neck Pain/ or Back Pain/ or Pain Measurement/ or Shoulder Pain/ or Pain Perception/ or Pain Management/ or Pain/ or Facial Pain/ or Pain, Referred/ or Musculoskeletal Pain/ or Nociceptive Pain/ or Pain Clinics/ or Pain, Intractable/ or Low Back Pain/ or Chronic Pain/ or Pain, Postoperative/ or Myofascial Pain Syndromes/ or Complex Regional Pain Syndromes/	250004
75	Stress, Psychological/pc, rh [Prevention & Control, Rehabilitation]	7630

Searches	Results: Ovid MEDLINE(R) 1946 to Present with Daily Update	Type
76	Sick Leave/	4693
77	disability leave.mp. or Sick Leave/	4720
78	"Quality of Life"/	143613
79	claim indicators.mp.	0
80	"Insurance Claim Review"/ or Insurance Claim Reporting/	9411
81	(duration or legal action).mp.	428017
82	cost.mp.	349149
83	"Compensation and Redress"/	2685
84	or/72-82	2219825
85	or/72-83	2221988
86	20 and 26 and 62 and 71	1075
87	20 and 27 and 62 and 85	4063
88	20 and 27 and 62 and (71 or 85)	4197
89	20 and 26 and 62 and (71 or 85)	4190
90	limit 86 to (english language and humans and yr="1995 -Current")	952
91	limit 87 to (english language and humans and yr="1995 -Current")	3266
92	20 and 27 and 62 and (71 or 84)	4195
93	limit 92 to (english language and humans and yr="1995 -Current")	**3374**
94	93 not (spinal cord injury or traumatic brain injury).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]	3031

APPENDIX 3: REFERENCE DETAILS FOR INCLUDED TOOLS

Tool	Relevant citation
BDI	(Silove et al., 2003)
CAPSIV	(Coffey, Gudmundsdottir, Beck, Palyo, & Miller, 2006)
CCI	(Skogstad et al., 2014)
CES-D	(Nota, Bot, Ring, & Kloen, 2015)
CPG	(Roy, MacDermid, Tang, & Beaton, 2013)
DASS-21	(Dahm, Ponsford, Wong, & Schonberger, 2011; Wiseman, Curtis, Lam, & Foster, 2015)
FABQ	(Holden, Davidson, & Tam, 2010)
FCE	(Branton et al., 2010; Gross & Battié, 2006; Trippolini et al., 2014)
FIM	(Kiely, Brasel, Weidner, Guse, & Weigelt, 2006)
FRI	(Casey, Feyer, & Cameron, 2011)
HADS	(Dahm et al., 2011)
IEQ	(Scott, Trost, Milioto, & Sullivan, 2013; Sullivan et al., 2008b)
IES- revised	(Coffey et al., 2006)
IPQ	(Gehrt et al., 2015)
McGill Pain Q	(Zogaria & Bohannon, 2003)
NDI	(Jull, Kenardy, Hendrikz, Cohen, & Sterling, 2013; Sterling et al., 2012; Sterling, Jull, Vicenzino, Kenardy, & Darnell, 2005; Williamson, Williams, Gates, & Lamb, 2015)
NRS	(Gopinath et al., 2015)
OMPQ	(Dunstan, Covic, Tyson, & Lennie, 2005; Gabel, Melloh, Burkett, Osborne, & Yelland, 2012; Law et al., 2013; Lee et al., 2013; Linton & Boersma, 2003; Margison & French, 2007; Oncu, Iliser, & Kuran, 2016; Westman, Linton, Ohrvik, Wahlen, & Leppert, 2008)
PAS	(O'Donnell et al., 2008)
PBPI	(Bostick, Carroll, Brown, Harley, & Gross, 2013)
PCS	(Casey, Feyer, & Cameron, 2015)
PSS-SR	(Coffey et al., 2006)
PTSD-RI	(Zatzick et al., 2006)
RMDQ	(Baldwin, Butler, Johnson, & Cote, 2007)
SBST	(Hill et al., 2011)
SF 12	(Baldwin et al., 2007)
SF36	(Kiely et al., 2006)
SMFA	(Nota et al., 2015)
STEPP	(van Meijel et al., 2015)
STEPP-Aus	(Nixon, Ellis, Nehmy, & Ball, 2010b)
TSQ	(Mouthaan, Sijbrandij, Reitsma, Gersons, & Olff, 2014)
WRREQ	(Gross & Battie, 2010)

APPENDIX 4: MEASUREMENT PROPERTIES SEARCHES

Details of searches to identify studies to establish measurement properties of included tools

Tool	Development & Validation studies; Title search	Systematic reviews
Work-related Recovery Expectations Questionnaire	**Title:** Measures of patients' expectations about recovery: a systematic review. [Review] (2015)	**1**
Visual analogue Scale (VAS)	"Visual analogue Scale" and valid* - 27 studies	0
Trauma screening questionnaire (TSQ)	"Trauma screening questionnaire" and valid* - 2 studies	0
Orebro	8 studies; Validation studies in several language	**1**
Tampa Scale for Kinesiophobia (TSK)	9 studies;	**1**
Survey of Pain Attitudes (SOPA)	5 studies	0
SF-36	138 studies	**7**
Screening tool for predictors of PTSD	1 - study	0
Ronald Morris Disability Questionnaire (RMDQ)	Validation studies in several languages. For some reason Medline is not showing	0
Recovery Related Anxiety Questionnaire (RRAQ)	Preliminary validation -1	0
PTSD Symptom Scale – Self Report (PSS)	1 study cited by 1981; The validation of a self-report measure of posttraumatic stress disorder: The Posttraumatic Diagnostic Scale. (Foa et a., 1997)	0
Posttraumatic Stress Diagnostic Scale (PDS)	Same authors as above; PSS is the earlier version	0
Post traumatic Adjustment Scale (PAS)	1 study cited by 81 papers	0
pain DETECT	10 studies; validation studies in several languages. More than 300,000 pts were assessed;	0
Pain Catastrophi?ing Scale (PCS)	20 studies	0
Pain Beliefs and Perception Inventory (PBPI)	Validation studies in several languages. For some reason Medline is not showing	0
Numeric Rating Scale (NRS)	5 studies	0
New York PTSD Risk Score (NYPRS)	1 study	0
Neck Disability Index (NDI)	35 studies	**2**
McGill Pain Questionnaire (MPQ)	24 studies	**1**
Life Orientation Test (LOT)	4 studies	0

Tool	Development & Validation studies; Title search	Systematic reviews
STarT Back	6 studies	0
Injustice Experience Questionnaire (IEQ)	3 studies	0
Impact of Event Scale (IES)	13 studies	0
Illness Perception Questionnaire (IPQ)	11 studies	2
Hospital Anxiety and Depression Scale (HADS)	49 studies	2
Functional Rating Index (FRI)	8 studies	0
Functional Independence Measure (FIM)	18 studies	1
Fear Avoidance Beliefs Questionnaire (FABQ)	10 studies	0
Davidson Trauma Scale (DTS)	4 studies	0
Depression Anxiety Stress Scales (DASS)	6 studies	0
Chronic Pain Grade (CPG)	2 studies	0
Center for Epidemiologic Studies Depression Scale (CES-D)	23 studies	0
Casualty Chain Inventory (CCI)	1 study	0
Beck Depression Inventory (BDI)	67	1
4-item SPAN	1	0

APPENDIX 5: DECISION RULES FOR PREDICTIVE STRENGTH

Decision rules applied to predictive strength rating for each tool

Across included studies different statistical approaches were applied to determine the predictive ability of the tools examined (Area Under the Curve – AUC, sensitivity and specificity, Odds Ratios - OR). Below is summary of the different categories used to determine whether the predictive strength was rated as excellent, good, moderate, correlation only or non-significant.

Excellent

- AUC 0.9-1.0 (excellent classification)
- Sensitivity or specificity > 0.95 (to enable use of tool to rule in or out)
- OR > 2.0 (indicates tool indicates greater than double likelihood of poor outcome)

Good

- AUC 0.8-0.9 (good classification)
- Sensitivity or specificity 0.90 - 0.95
- OR 1.50 – 2.0

Moderate

- AUC 0.7-0.8 (fair classification)
- Sensitivity or specificity 0.80-0.90
- OR 1.0 – 1.5

Correlation only

- Evidence of correlation with outcome or other measures only

Not significant

- Tool found not to predict outcome at any level

Decision rules applied to the level of evidence to support the tool

Where possible, reviews of included tools were located in order to summarise the level of support for the tool and the quality of measurement properties. Three levels were determined based on the characteristics examined in review of the tools. These characteristics included Cronbach's alpha (measure of internal consistency), the intra-class coefficient (measure of the reliability of the tool), weighted kappa (measure of reliability of the tool) and correlation coefficients (measure of how the tool compares to existing tools and outcomes).

High level of support

To be rated as high, reports of the measurement characteristics of the tool had to consistently fall under the highest (excellent) rated categories for each of the statistics examined (rules for each are outlined below).

Medium level of support

To be rated as medium, reports of the measurement characteristics had to consistently fall into the next highest (good) rated categories for each of the statistics examined OR contained a mix between the highest and moderate level categories. Where there were suggestions of strong characteristics from limited evidence (either number of studies or type of supporting evidence), a level of medium support was applied.

Low level of support

To be rated as low, reports of the measurement characteristics had to consistently fall into the lower rated categories for the statistics examined. Where there were large inconsistencies in the ratings provided this level of support was applied.

Not available –N/A

Where no reviews of the measurement properties of the tool were available, N/A was applied to indicate that further research is required to be able to make recommendations on the general properties of the tool.

Decision rules applied in determination of level of support

Cronbach's alpha

Cronbach's alpha	Internal consistency
$\alpha > 0.9$	Excellent
$0.9 > \alpha > 0.8$	Good
$0.8 > \alpha > 0.7$	Acceptable
$0.7 > \alpha > 0.6$	Questionable
$0.6 > \alpha > 0.5$	Poor
$0.5 > \alpha$	Unacceptable

Intra-class correlation Coefficient (ICC)

Cronbach's alpha	Reliability
> 0.9	Excellent
0.9 – 0.75	Good
0.75 – 0.5	Moderate
< 0.5	Poor

Weighted Kappa

Kappa	Interpretation
1.0 – 0.81	Almost perfect agreement
0.8 – 0.61	Substantial agreement
0.6 – 0.41	Moderate agreement
0.4 – 0.21	Fair agreement
0.2 - 0	Slight agreement
<0	Poor agreement

Correlation coefficients

Size of correlation	Interpretation
1.00 – 0.90	Very high (positive or negative)
0.90 – 0.70	High (positive or negative)
0.70 – 0.50	Moderate (positive or negative)
0.50 – 0.30	Low (positive or negative)
0.30 – 0	Negligible (positive or negative)

APPENDIX 6: SUMMARY OF STUDIES INCLUDED IN PART 2

Reference	Country	Setting	Sample description	Interventions	Quality appraisal summary
Hill et al (2011)	UK	Primary health care Proportion compensation claimants = Not Reported	Patients with acute/chronic low back pain (Range - less than one month to greater than 3 years) Key exclusion criteria: patients with potentially serious, disorders (e.g., cauda equine compression, inflammatory arthritis, and malignancy), serious illness or comorbidity (including those undergoing treatment for a prevalent axis 1 or 3 mental health disorder according to the DSM-IV criteria), who had spinal injury in the past 6 months, who were pregnant, who were receiving back treatments (except primary care) and who were unable or unwilling to attend.	SERVICE DETAILS: The baseline clinical assessments and treatment sessions, decisions about referral were made by use of the STarT Back Screening tool classification. *All risk groups*: The 30 min assessment and initial treatment were delivered according to an agreed protocol, with advice focusing on promotion of appropriate levels of activity, including return to work, and a pamphlet about local exercise venues and self-help groups. Participants were shown a 15 min video entitled Get Back Active given the Back Book. - Low risk pts were only given this clinic session; - Medium and high risk pts were referred for further physiotherapy-led treatment sessions. Follow up sessions for medium risk pts was standardized physiotherapy. i.e., to address symptoms and function. - High risk pts were referred to psychologically informed physiotherapists to address psychosocial obstacles to recovery. TRAINING: Intervention – Clinic physiotherapists received 1 day of training about the Start Back Screening Tool, and to standardize advice and data gathering from the case report forms. Follow up therapists: those providing service to the medium risk were given 3 days of additional training and those delivering the high risk group had 6 days of additional training (9 days). Control - Clinic physiotherapists received a half day training to familiarize them with study procedures. Follow-up therapists: had general training in physical therapies and some training in	Overall rating: Low Risk of Bias We used Cochrane risk of bias tool (modified) for quality assessment of randomized controlled trials. http://handbook.cochrane.org/chapter_8/8_assessing_risk_of_bias_in_included_studies.htm

Reference	Country	Setting	Sample description	Interventions	Quality appraisal summary
				more complex psychologically informed treatments, and none underwent special training or instructions related this study.	
Jull et al (2013)	AUS	General medical practice & Physio & Psych practice located in the community Proportion compensation claimants = Reported compensation claim lodged (n=95); I – 27.7%; C – 29.2%	Patients with acute whiplash <4 weeks post injury as a result of MVC. Referred by GP or other HCPs Key exclusion criteria: If volunteers' neck pain was not related to a motor vehicle crash; was classifiable as WAD I, III, IV; if there was a previous history of whiplash or other neck pain for which treatment had been sought; if they lacked fluency in spoken and written English for independent questionnaire completion, if they were unwilling to receive either pragmatic treatment or usual care	SERVICE DETAILS – Pragmatic intervention - individualised multi-professional management (medical with pharmacotherapeutic options, physiotherapeutic and psychological) Physiotherapy - The number of treatment occasions was not fixed and was at the discretion of the treating physiotherapist based on the participant's responsiveness, but the maximum intervention period was 10 wks. Psychological mgmt. – 6 wk to 10 wks CBT TRAINING – general practitioner provided medical management in consultation with the trial's medical pain specialist and included advice, assurance and medication	Overall rating: High Risk of Bias For sequence generation, blinding of participants and personnel, incomplete outcome data, and other sources of bias. We used Cochrane risk of bias tool (modified) for quality assessment of randomized controlled trials. http://handbook.cochrane.org/chapter_8/8_assessing_risk_of_bias_in_included_studies.htm

REFERENCES

Baldwin, M. L., Butler, R. J., Johnson, W. G., & Cote, P. (2007). Self-reported severity measures as predictors of return-to-work outcomes in occupational back pain. *Journal of Occupational Rehabilitation, 17*(4), 683-700.

Bishop, A., Thomas, E., & Foster, N. E. (2007). Health care practitioners' attitudes and beliefs about low back pain: a systematic search and critical review of available measurement tools. *Pain, 132*(1-2), 91-101. doi: 10.1016/j.pain.2007.01.028

Bostick, G. P., Carroll, L. J., Brown, C. A., Harley, D., & Gross, D. P. (2013). Predictive capacity of pain beliefs and catastrophizing in Whiplash Associated Disorder. *Injury, 44*(11), 1465-1471. doi: http://dx.doi.org/10.1016/j.injury.2012.10.007

Bouffard, J., Bertrand-Charette, M., & Roy, J. S. (2016). Psychometric properties of the Musculoskeletal Function Assessment and the Short Musculoskeletal Function Assessment: a systematic review. *Clinical Rehabilitation, 30*(4), 393-409. doi: 10.1177/0269215515579286

Branton, E. N., Arnold, K. M., Appelt, S. R., Hodges, M. M., Battie, M. C., & Gross, D. P. (2010). A short-form functional capacity evaluation predicts time to recovery but not sustained return-to-work. *Journal of Occupational Rehabilitation, 20*(3), 387-393. doi: http://dx.doi.org/10.1007/s10926-010-9233-9

Broadbent, E., Wilkes, C., Koschwanez, H., Weinman, J., Norton, S., & Petrie, K. J. (2015). A systematic review and meta-analysis of the Brief Illness Perception Questionnaire. *Psychol Health, 30*(11), 1361-1385. doi: 10.1080/08870446.2015.1070851

Carleton, R. N., Thibodeau, M. A., Teale, M. J., Welch, P. G., Abrams, M. P., Robinson, T., & Asmundson, G. J. (2013). The center for epidemiologic studies depression scale: a review with a theoretical and empirical examination of item content and factor structure. *PLoS ONE [Electronic Resource], 8*(3), e58067. doi: 10.1371/journal.pone.0058067

Casey, P. P., Feyer, A. M., & Cameron, I. D. (2011). Identifying predictors of early non-recovery in a compensation setting: The Whiplash Outcome Study. *Injury, 42*(1), 25-32. doi: http://dx.doi.org/10.1016/j.injury.2010.07.234

Casey, P. P., Feyer, A. M., & Cameron, I. D. (2015). Course of recovery for whiplash associated disorders in a compensation setting. *Injury, 46*(11), 2118-2129. doi: http://dx.doi.org/10.1016/j.injury.2015.08.038

Coffey, S. F., Gudmundsdottir, B., Beck, J. G., Palyo, S. A., & Miller, L. (2006). Screening for PTSD in motor vehicle accident survivors using the PSS-SR and IES. *Journal of Traumatic Stress, 19*(1), 119-128.

Dahm, J., Ponsford, J., Wong, D., & Schonberger, M. (2011). Utility of the Depression Anxiety Stress Scales in assessing depression and anxiety following traumatic brain injury. *Brain Impairment, 12*, 72.

Dunstan, D. A., Covic, T., Tyson, G. A., & Lennie, I. G. (2005). Does the Örebro Musculoskeletal Pain Questionnaire predict outcomes following a work-related compensable injury? *International Journal of Rehabilitation Research, 28*(4), 369-370. doi: 10.1097/00004356-200512000-00012

Feise, R. J., & Menke, J. M. (2010). Functional Rating Index: literature review. *Med Sci Monit, 16*(2), RA25-36.

Gabel, C. P., Melloh, M., Burkett, B., Osborne, J., & Yelland, M. (2012). The Orebro Musculoskeletal Screening Questionnaire: validation of a modified primary care musculoskeletal screening tool in an acute work injured population. *Manual Therapy, 17*(6), 554-565. doi: http://dx.doi.org/10.1016/j.math.2012.05.014

Gehrt, T. B., Wisbech Carstensen, T. B., Ornbol, E., Fink, P. K., Kasch, H., & Frostholm, L. (2015). The role of illness perceptions in predicting outcome after acute whiplash trauma: a multicenter

12-month follow-up study. *Clinical Journal of Pain, 31*(1), 14-20. doi: http://dx.doi.org/10.1097/AJP.0000000000000085

Glenny, C., & Stolee, P. (2009). Comparing the functional independence measure and the interRAI/MDS for use in the functional assessment of older adults: a review of the literature. *BMC Geriatr, 9*, 52. doi: 10.1186/1471-2318-9-52

Gopinath, B., Jagnoor, J., Nicholas, M., Blyth, F., Harris, I. A., Casey, P., & Cameron, I. D. (2015). Presence and predictors of persistent pain among persons who sustained an injury in a road traffic crash. *European Journal of Pain, 19*(8), 1111-1118. doi: http://dx.doi.org/10.1002/ejp.634

Gross, D. P., & Battie, M. C. (2010). Recovery expectations predict recovery in workers with back pain but not other musculoskeletal conditions. *Journal of Spinal Disorders & Techniques, 23*(7), 451-456. doi: http://dx.doi.org/10.1097/BSD.0b013e3181d1e633

Gross, D. P., & Battié, M. C. (2006). Does functional capacity evaluation predict recovery in workers' compensation claimants with upper extremity disorders? *Occupational & Environmental Medicine, 63*(6), 404-410.

Hawker, G. A., Mian, S., Kendzerska, T., & French, M. (2011). Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). *Arthritis Care Res (Hoboken), 63 Suppl 11*, S240-252. doi: 10.1002/acr.20543

Hill, J. C., Whitehurst, D. G., Lewis, M., Bryan, S., Dunn, K. M., Foster, N. E., . . . Hay, E. M. (2011). Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. *Lancet, 378*(9802), 1560-1571. doi: http://dx.doi.org/10.1016/S0140-6736(11)60937-9

Hingorani, A. D., Windt, D. A., Riley, R. D., Abrams, K., Moons, K. G., Steyerberg, E. W., . . . Group, P. (2013). Prognosis research strategy (PROGRESS) 4: stratified medicine research. *BMJ, 346*, e5793. doi: 10.1136/bmj.e5793

Holden, J., Davidson, M., & Tam, J. (2010). Can the Fear-Avoidance Beliefs Questionnaire predict work status in people with work-related musculoskeltal disorders? *Journal of Back & Musculoskeletal Rehabilitation, 23*(4), 201-208.

Jull, G., Kenardy, J., Hendrikz, J., Cohen, M., & Sterling, M. (2013). Management of acute whiplash: a randomized controlled trial of multidisciplinary stratified treatments. *Pain, 154*(9), 1798-1806. doi: http://dx.doi.org/10.1016/j.pain.2013.05.041

Karran, E. L., McAuley, J. H., Traeger, A. C., Hillier, S. L., Grabherr, L., Russek, L. N., & Moseley, G. L. (2017). Can screening instruments accurately determine poor outcome risk in adults with recent onset low back pain? A systematic review and meta-analysis. *BMC Med, 15*(1), 13. doi: 10.1186/s12916-016-0774-4

Kiely, J. M., Brasel, K. J., Weidner, K. L., Guse, C. E., & Weigelt, J. A. (2006). Predicting quality of life six months after traumatic injury. *Journal of Trauma-Injury Infection & Critical Care, 61*(4), 791-798.

Law, R. K., Lee, E. W., Law, S. W., Chan, B. K., Chen, P. P., & Szeto, G. P. (2013). The predictive validity of OMPQ on the rehabilitation outcomes for patients with acute and subacute non-specific LBP in a Chinese population. *Journal of Occupational Rehabilitation, 23*(3), 361-370. doi: http://dx.doi.org/10.1007/s10926-012-9404-y

Lee, W. Y. A., Lee, W. C. E., Law, S. W., Lau, W. K. A., Leung, S. M., Sieh, K. M., . . . Law, K. Y. R. (2013). Managing Psychosocial Contributors in Low Back Pain Patients-A Randomised Controlled Trial. *Journal of Orthopaedics, Trauma and Rehabilitation, 17*(1), 46-51. doi: 10.1016/j.jotr.2012.12.002

Linton, S. J., & Boersma, K. (2003). Early identification of patients at risk of developing a persistent back problem: the predictive validity of the Orebro Musculoskeletal Pain Questionnaire. *Clinical Journal of Pain, 19*(2), 80-86.

Margison, D. A., & French, D. J. (2007). Predicting Treatment Failure in the Subacute Injury Phase Using the Orebro Musculoskeletal Pain Questionnaire: An Observational Prospective Study in a Workers' Compensation System. *Journal of Occupational and Environmental Medicine, 49*(1), 59-67. doi: http://dx.doi.org/10.1097/JOM.0b013e31802db51e

Mouthaan, J., Sijbrandij, M., Reitsma, J. B., Gersons, B. P., & Olff, M. (2014). Comparing screening instruments to predict posttraumatic stress disorder. *PLoS ONE [Electronic Resource], 9*(5), e97183. doi: http://dx.doi.org/10.1371/journal.pone.0097183

Nixon, R. D., Ellis, A. A., Nehmy, T. J., & Ball, S. A. (2010a). Screening and predicting posttraumatic stress and depression in children following single-incident trauma. *Journal of Clinical Child & Adolescent Psychology, 39*(4), 588-596. doi: 10.1080/15374416.2010.486322

Nixon, R. D., Ellis, A. A., Nehmy, T. J., & Ball, S. A. (2010b). Screening and predicting posttraumatic stress and depression in children following single-incident trauma. *Journal of Clinical Child & Adolescent Psychology, 39*(4), 588-596. doi: http://dx.doi.org/10.1080/15374416.2010.486322

Norton, S., Cosco, T., Doyle, F., Done, J., & Sacker, A. (2013). The Hospital Anxiety and Depression Scale: a meta confirmatory factor analysis. *J Psychosom Res, 74*(1), 74-81. doi: 10.1016/j.jpsychores.2012.10.010

Nota, S. P., Bot, A. G., Ring, D., & Kloen, P. (2015). Disability and depression after orthopaedic trauma. *Injury, 46*(2), 207-212. doi: http://dx.doi.org/10.1016/j.injury.2014.06.012

O'Donnell, M. L., Creamer, M. C., Parslow, R., Elliott, P., Holmes, A. C., Ellen, S., . . . Bryant, R. A. (2008). A predictive screening index for posttraumatic stress disorder and depression following traumatic injury. *Journal of Consulting & Clinical Psychology, 76*(6), 923-932. doi: http://dx.doi.org/10.1037/a0012918

Oncu, J., Iliser, R., & Kuran, B. (2016). Cross-cultural adaptation of the Orebro Musculoskeletal Pain Questionnaire among Turkish workers with low back pain. *Journal of Back & Musculoskeletal Rehabilitation, 29*(1), 135-143. doi: http://dx.doi.org/10.3233/BMR-150609

Roy, J. S., MacDermid, J. C., Tang, K., & Beaton, D. E. (2013). Construct and predictive validity of the chronic pain grade in workers with chronic work-related upper-extremity disorders. *Clinical Journal of Pain, 29*(10), 891-897. doi: http://dx.doi.org/10.1097/AJP.0b013e318278d455

Sattelmayer, M., Lorenz, T., Roder, C., & Hilfiker, R. (2012). Predictive value of the Acute Low Back Pain Screening Questionnaire and the Orebro Musculoskeletal Pain Screening Questionnaire for persisting problems. *European Spine Journal, 21 Suppl 6*, S773-784. doi: 10.1007/s00586-011-1910-7

Scott, W., Trost, Z., Milioto, M., & Sullivan, M. J. (2013). Further validation of a measure of injury-related injustice perceptions to identify risk for occupational disability: a prospective study of individuals with whiplash injury. *Journal of Occupational Rehabilitation, 23*(4), 557-565. doi: http://dx.doi.org/10.1007/s10926-013-9417-1

Shea, T. L., Tennant, A., & Pallant, J. F. (2009). Rasch model analysis of the Depression, Anxiety and Stress Scales (DASS). *BMC Psychiatry, 9*, 21. doi: 10.1186/1471-244X-9-21

Silove, D., Blaszczynski, A., Manicavasager, V., Tyndall, K., Petridis, A., & Hillman, K. (2003). Capacity of Screening Questionnaires to Predict Psychiatric Morbidity 18 Months After Motor Vehicle Accidents. *Journal of Nervous and Mental Disease, 191*(9), 604-610. doi: http://dx.doi.org/10.1097/01.nmd.0000087187.13408.e1

Skogstad, L., Toien, K., Hem, E., Ranhoff, A. H., Sandvik, L., & Ekeberg, O. (2014). Psychological distress after physical injury: a one-year follow-up study of conscious hospitalised patients. *Injury, 45*(1), 289-298. doi: http://dx.doi.org/10.1016/j.injury.2012.10.001

Smeets, R., Koke, A., Lin, C. W., Ferreira, M., & Demoulin, C. (2011). Measures of function in low back pain/disorders: Low Back Pain Rating Scale (LBPRS), Oswestry Disability Index (ODI),

Progressive Isoinertial Lifting Evaluation (PILE), Quebec Back Pain Disability Scale (QBPDS), and Roland-Morris Disability Questionnaire (RDQ). *Arthritis Care Res (Hoboken), 63 Suppl 11*, S158-173. doi: 10.1002/acr.20542

Steinberg, A. M., Brymer, M. J., Kim, S., Briggs, E. C., Ippen, C. G., Ostrowski, S. A., . . . Pynoos, R. S. (2013). Psychometric properties of the UCLA PTSD reaction index: part I. *Journal of Traumatic Stress, 26*(1), 1-9. doi: 10.1002/jts.21780

Sterling, M., Hendrikz, J., Kenardy, J., Kristjansson, E., Dumas, J. P., Niere, K., . . . Jull, G. (2012). Assessment and validation of prognostic models for poor functional recovery 12 months after whiplash injury: a multicentre inception cohort study. *Pain, 153*(8), 1727-1734. doi: http://dx.doi.org/10.1016/j.pain.2012.05.004

Sterling, M., Jull, G., Vicenzino, B., Kenardy, J., & Darnell, R. (2005). Physical and psychological factors predict outcome following whiplash injury. *Pain, 114*(1-2), 141-148.

Sullivan, M. J., Adams, H., Horan, S., Maher, D., Boland, D., & Gross, R. (2008a). The role of perceived injustice in the experience of chronic pain and disability: scale development and validation. *Journal of Occupational Rehabilitation, 18*(3), 249-261. doi: 10.1007/s10926-008-9140-5

Sullivan, M. J., Adams, H., Horan, S., Maher, D., Boland, D., & Gross, R. (2008b). The role of perceived injustice in the experience of chronic pain and disability: scale development and validation. *Journal of Occupational Rehabilitation, 18*(3), 249-261. doi: http://dx.doi.org/10.1007/s10926-008-9140-5

Sundin, E. C., & Horowitz, M. J. (2002). Impact of Event Scale: psychometric properties. *Br J Psychiatry, 180*, 205-209.

Trippolini, M. A., Dijkstra, P. U., Cote, P., Scholz-Odermatt, S. M., Geertzen, J. H., & Reneman, M. F. (2014). Can functional capacity tests predict future work capacity in patients with whiplash-associated disorders? *Archives of Physical Medicine & Rehabilitation, 95*(12), 2357-2366. doi: http://dx.doi.org/10.1016/j.apmr.2014.07.406

van Meijel, E. P., Gigengack, M. R., Verlinden, E., Opmeer, B. C., Heij, H. A., Goslings, J. C., . . . Lindauer, R. J. (2015). Predicting posttraumatic stress disorder in children and parents following accidental child injury: evaluation of the Screening Tool for Early Predictors of Posttraumatic Stress Disorder (STEPP). *BMC Psychiatry, 15*, 113. doi: http://dx.doi.org/10.1186/s12888-015-0492-z

Walters, J. T., Bisson, J. I., & Shepherd, J. P. (2007). Predicting post-traumatic stress disorder: validation of the Trauma Screening Questionnaire in victims of assault. *Psychological Medicine, 37*(1), 143-150. doi: 10.1017/S0033291706008658

Walton, D. M., Wideman, T. H., & Sullivan, M. J. (2013). A Rasch analysis of the pain catastrophizing scale supports its use as an interval-level measure. *Clinical Journal of Pain, 29*(6), 499-506. doi: 10.1097/AJP.0b013e318269569c

Wang, Y. P., & Gorenstein, C. (2013). Psychometric properties of the Beck Depression Inventory-II: a comprehensive review. *Rev Bras Psiquiatr, 35*(4), 416-431. doi: 10.1590/1516-4446-2012-1048

Westman, A., Linton, S. J., Ohrvik, J., Wahlen, P., & Leppert, J. (2008). Do psychosocial factors predict disability and health at a 3-year follow-up for patients with non-acute musculoskeletal pain? A validation of the Orebro Musculoskeletal Pain Screening Questionnaire. *European Journal of Pain, 12*(5), 641-649.

Williamson, E., Williams, M. A., Gates, S., & Lamb, S. E. (2015). Risk factors for chronic disability in a cohort of patients with acute whiplash associated disorders seeking physiotherapy treatment for persisting symptoms. *Physiotherapy, 101*(1), 34-43. doi: http://dx.doi.org/10.1016/j.physio.2014.04.004

Wiseman, T. A., Curtis, K., Lam, M., & Foster, K. (2015). Incidence of depression, anxiety and stress following traumatic injury: a longitudinal study. *Scandinavian Journal of Trauma, Resuscitation & Emergency Medicine, 23*, 29. doi: http://dx.doi.org/10.1186/s13049-015-0109-z

Yao, M., Sun, Y. L., Cao, Z. Y., Dun, R. L., Yang, L., Zhang, B. M., . . . Cui, X. J. (2015). A systematic review of cross-cultural adaptation of the neck disability index. *Spine (Phila Pa 1976), 40*(7), 480-490. doi: 10.1097/BRS.0000000000000788

Zatzick, D. F., Grossman, D. C., Russo, J., Pynoos, R., Berliner, L., Jurkovich, G., . . . Rivara, F. P. (2006). Predicting posttraumatic stress symptoms longitudinally in a representative sample of hospitalized injured adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry, 45*(10), 1188-1195.

Zogaria, J. R., & Bohannon, R. W. (2003). McGill Pain Questionnaire scores do not predict outcomes of outpatient therapy for patients with low back pain. *Journal of Manual and Manipulative Therapy, 11*(2), 110-116.