**Prognostic factors and models for predicting work absence in adults with musculoskeletal conditions consulting a healthcare practitioner: A systematic review**

Gwenllian Wynne-Jones1, Elaine Wainwright2,3, Nicola Goodson4, Joanne L. Jordan1, Amardeep Legha1, Millie Parchment5, Ross Wilkie1, George Peat6

1. School of Medicine, Faculty of Medicine and Health Sciences, and Versus Arthritis/Medical Research Council Centre for Musculoskeletal Health and Work, Keele University, Keele ST5 5BG
2. Aberdeen Centre for Arthritis and Musculoskeletal Health (Epidemiology Group), School of Medicine, Medical Sciences and Nutrition, Versus Arthritis/Medical Research Council Centre for Musculoskeletal Health and Work, University of Aberdeen, UK
3. Centre for Pain Research, University of Bath, UK
4. Department of Rheumatology, Liverpool University Hospitals, Versus Arthritis/Medical Research Council Centre for Musculoskeletal Health and Work, Liverpool L9 7AL, UK
5. NIHR Applied Research Collaboration for Greater Manchester, University of Manchester, M13 9PL
6. Centre for Applied Health & Social Care Research (CARe), Sheffield Hallam University, UK

**Abstract**

*Purpose*

It is difficult to predict which employees, in particular those with musculoskeletal pain, will return to work quickly without additional vocational advice and support, which employees will require this support and what levels of support are most appropriate. Consequently, there is no way of ensuring the right individuals are directed towards the right services to support their occupational health needs. The aim of this review will be to identify prognostic factors for duration of work absence in those already absent and examine the utility of prognostic models for work absence.

*Methods*

Eight databases were search using a combination of subject headings and key words focusing on work absence, musculoskeletal pain and prognosis. Two authors independently assessed the eligibility of studies, extracted data from all eligible studies and assessed risk of bias using the QUIPS or PROBAST tools, an adapted GRADE used to assess the strength of the evidence.

To make sense of the data prognostic variables were grouped according to categories from the Disability Prevention Framework and the SWiM framework was utilised to synthesize findings.

*Results*

A total of 23 studies were included in the review, including 13 prognostic models and a total of 110 individual prognostic factors. Overall, the evidence for all prognostic factors was weak, although there was some evidence that older age and better recovery expectations were protective of future absence and that previous absence was likely to predict future absences. There was weak evidence for any of the prognostic models in determining future sickness absence.

*Conclusion*

Analysis was difficult due to the wide range of measures of both prognostic factors and outcome and the differing timescales for follow-up. Future research should ensure that consistent measures are employed and where possible these should be in-line with those suggested by Ravinskaya *et al* (2023).

**Key words**: Systematic Review; Absence,

**Introduction**

Sickness absence remains a significant cost to developed countries accounting for between 0.5% and 2% of lost GDP in European countries alone.[1] There is evidence that the majority of employees taking a period of absence will only take a short time away from the workplace [2]. However, across all working ages there is a small minority that go on to take longer-term absence, variously defined as greater than 4 weeks absence, greater than 6 weeks, and up to 3 months.[3], [4], [5] However long-term absence is defined, it is the small proportion of people going on to long-term absence who make up the majority of the costs associated with absence from the workplace.[6]

At present it is difficult to predict which employees, in particular those with musculoskeletal pain, will return to work quickly without additional vocational advice and support, which employees will require this support and what levels of support are most appropriate. Consequently, there is no way of ensuring the right individuals are directed towards the right services to support their occupational health needs. There is a growing evidence base around the usefulness of stratified care approaches to delivering healthcare, whereby prognostic information is used to allocate individuals to subgroups with matched recommended treatments or interventions[7], [8], stratified care has also been demonstrated to be a cost effective model.[9], [10] This approach has not been developed in occupational health yet, but the principles behind it could be used to ensure that scarce occupational health resources are targeted towards those individuals who need more support, whilst also providing reassurance to those for whom sickness absence is unlikely to be become longer-term. To allow a stratified care approach to be developed it is important to identify which factors predict work absence and to examine the utility of current prognostic models or tools[11].

It is anticipated that prognostic factors for work absence will be varied. Sickness absence is a complex concept, influenced not only by an individual’s health (or severity of health condition), but also by psychosocial variables, macro system variables (e.g. health services and workplace systems) and wider societal systems (e.g. sickness benefits policies) [12]. For many individuals, decisions about sickness absence will be made in the context of their own health, their own workplace and their own attitudes and beliefs. To support the management of the variety of prognostic variables anticipated in this systematic review, it may be possible to identify some common “core” concepts that can be used to predict the likelihood that individuals will go on to longer-term absence. These concepts can be organised around a framework such as the disability prevention framework [13], which structures the impacts on the health and work relationship into the “core” concepts of personal systems, healthcare systems, workplace systems and compensation system. Within each of these core concepts are sub-groups which allow an examination of the potential predictors of work absence on a more granular level.

Whilst there is a body of literature examining predictors of sickness absence [14], [15], [16], Lilley, there have been no systematic reviews that comprehensively consider which factors are predictive of work absence or the usefulness of prognostic models or measurement tools in identifying those who will have longer-term work absence. Furthermore, there is no evidence focussed on the prediction of absence duration in those that are already absent from work and presenting to primary care. This is a key timepoint in which to be able to provide evidence-based advice and guidance or to refer patients to appropriate services to support them with their health and work in particular those with long-terms conditions such as musculoskeletal pain (NICE 2019). Therefore, the primary aim of this review will be to identify prognostic factors for duration of work absence in those already absent and examine the utility of prognostic models for work absence.

**Methods**

This systematic review is reported using the PRISMA guidance [17] and the recommendations of Riley *et al* (2019)[18]for undertaking systematic review of prognostic factors. The review was prospectively registered with PROSPERO (CRD42020219452).

***Search strategy***

An experienced information specialist designed and conducted the searches using a combination of subject headings and key text words. The full search strategy is reported in the online supplement. The following eight databases were searched from their inception to 6th October 2020: MEDLINE; EMBASE; CINAHL; AMED; PsycINFO; HMIC; Business Source Complete; Cochrane Library (CENTRAL), a full updated search was run on 18th September 2023.

***Inclusion and exclusion criteria***

*Participants/population*

Studies including employed adults who were on sick leave and seeking or receiving healthcare for a musculoskeletal condition were included. If studies reported on participants who were unemployed, not on sick leave or working modified or alternative duties they were excluded. Studies where participants did not have musculoskeletal conditions or where these were as a result of acute trauma or injuries (such as fractures) were excluded as were studies where the participants had inflammatory arthritis or surgical intervention for their condition.

*Study setting*

Studies set in primary (first contact) care, community care and workplace settings where employees have sought healthcare have been included. Studies conducted in hospital populations, emergency care, tertiary care or rehabilitation centres were excluded.

*Study type*

Cohort studies (prospective and retrospective) with an integrated health and work focus were included. Additionally, prognosis studies based on randomised controlled trial data and / or case control studies were included alongside those papers that reported on tools or models used to predict work absence and summarise the predictive performance of the tool or model used. All other study designs were excluded.

***Prognostic factors***

The predictive performance of all identified prognostic factors or prognostic models were evaluated. We did not limit the factors that could be included allowing a full exploration of the breadth of prognostic factors examined in the literature.

***Outcomes***

The outcome of interest for this review was work absence. Prognostic factors for RTW will be reported in a separate publication. Work absence could be measured in any way (e.g. self-report, employer records, or insurance records), and at any follow-up time point. Definitions of absence were extracted from the studies to allow a comparison of outcome measures.

The strength of association of individual prognostic factors with the outcome were extracted from studies. Where the outcome was binary (absence from work yes versus no) the odds ratio, relative risk or time to event data were extracted, where the outcome was continuous (e.g. number of days absent from work) the mean differences were extracted.

***Screening and data extraction***

All screening and data extraction was undertaken by pairs of review authors independently. Any disagreements were resolved through consensus bringing in a third reviewer if necessary. The screening of titles and abstracts was undertaken using Rayaan software and the full text screening and data extraction undertaken using Covidence software.

A standardised data extraction form was developed and tested using MS Excel before being used to extract data from the included studies. Study level data were collected on: study design (primary care, community, population/national based, health records (primary care), health records (secondary care), health records (insurance), occupational health, outpatients, hospital/rehabilitation, other secondary care, other setting(defined)). Data were also collected on inclusion criteria, population description, definition of outcome, outcome data type (binary, continuous, time-to-event), follow-up time period, prognostic factor and description, variables used in adjustment of the analyses, adjusted and unadjusted estimates of the association between the prognostic factor and the outcome. Where studies reported on a model, measures of the model’s performance were also extracted.

***Quality assessment***

The Quality In Prognosis Studies (QUIPS) tool [19]was used to assess potential bias in prognostic factor studies and Prediction model Risk Of Bias ASsessment Tool (PROBAST)[20] for studies reporting prognostic models. Two authors independently assessed risk of bias for each study reported as unclear, high, or low risk of bias, for each domain of the tools (QUIPS and PROBAST). These were compared between each pair and any disagreements resolved through discussion or by consulting a third reviewer if necessary.

***Assessment of the strength of evidence***

GRADE was used to assess the strength of the evidence. This method takes into account a number of factors allowing a judgement to be made on the body of evidence overall rather than focusing on individual studies as with risk of bias.

For each of the groups of prognostic factors reported below, GRADE was used to assess the risk of bias with evidence downgraded where more than half of studies had moderate or high risk of bias. Additionally, evidence was downgraded where there was inconsistency in estimates of effect and / or heterogeneity between studies in the definition of the prognostic factor. Downgrading was also applied if there was any indirectness defined as follows: not all the participants were absent and separate results were not reported for those that were; only a subset of the population was represented (e.g. just males / females); the prognostic factor was not fully represented e.g. only a subset of those reporting absence were included. Finally, evidence was downgraded if there was any imprecision which included fewer than 2 studies in each prognostic factor grouping or if there was an insufficient sample size to detect a difference for the prognostic factor in most of the studies.

When considering the strength of the evidence around the prognostic models predicting absence from work an adapted GRADE was used. This was primarily to ensure that appropriate consideration of the performance of the included models was included, the guidance from Foroutan *et al* 2022 [21]was used. Evidence was downgraded where calibration was imprecise with wide variation in point estimates overall and wide confidence intervals.

Evidence was deemed to be high quality if none of the domains were downgraded, moderate quality if one of the domains was downgraded, low quality if two were downgraded and very low quality if three or four were downgraded.

***Data synthesis***

A narrative synthesis was planned to allow for variation in outcome measures, settings, and prognostic factors included in the studies within this review. Whilst the Popay narrative synthesis framework [22] had been planned to be used, the more recent Synthesis Without Meta-analysis (SWiM) framework was used to structure the data synthesis [23], this framework provides a guide with which to group, describe and report the results of this systematic review and was considered a more appropriate approach to synthesising the evidence in this review.

*Grouping of prognostic factors*

Due to the wide variation in prognostic factors measured within the studies in this review, they were grouped into broad domains. In total there were 110 individual factors identified which grouped via discussion within the team into 17 broad categories. These categories were further grouped for synthesis to broadly fit the categories of the Disability Prevention Framework[13] however, there were no variables that could be grouped into the compensation system concepts and just one variable reporting a healthcare prognostic factor (Figure 1).

**Figure 1: Overarching groups and categories of prognostic factors**

{Here}

*Description of standardised metric*

This paper aimed to identify prognostic factors for work absence and therefore a range of metrics were extracted and recorded. For binary outcomes odds ratios (OR), relative risk (RR) and risk reduction were recorded, for continuous outcomes mean differences were recorded, and for time to event outcomes hazard ratios (HR) were recorded. Where available the adjusted and unadjusted effect estimates were recorded. These metrics for reporting prognostic factors are recommended in the CHARMS-PF checklist. [24], [25]

*Methods of synthesis*

There was significant inconsistency across prognostic factors in terms of measurement and analysis and also inconsistency in outcome measure so a formal meta-analysis was not possible. However, data were sufficient to report the range and distribution of observed effects as well as identifying whether there was evidence of an effect in one or more studies examining the same prognostic factor and also to explore the direction of any effects seen.

**Results**

The searches returned 1608 references. Following de-duplication 1589 references remained. After completing screening of titles and abstracts 281 full texts were retrieved for assessment of eligibility with 22 studies included in the current systematic review and a further 44 studies identified for inclusion in a separate review reporting RTW as the outcome (Figure 2).

**Figure 2: Flow diagram of study selection**

{Here}

**Results: prognostic factors**

Across all 23 studies 111 individual prognostic factors were identified. There was considerable inconsistency in study design, prognostic factor measurement, outcome measurement, time point of follow-up and analysis methods across the included studies meaning a meta-analysis was not appropriate. Prognostic factors were grouped into 18 themes for ease of management and reporting (Figure 1) which can be considered within three domains of the disability prevention framework: personal system, workplace system and healthcare system.

***Description of included studies: prognostic factors***

Table 1 provides an overview of the included studies. In summary the studies were mainly conducted in North America (9 in Canada and 4 in the USA), this was followed by the Netherlands with 3 studies, 3 originating in Australia, with the rest from other European countries (5 studies).

The majority of studies were conducted using records from healthcare insurance databases (10 studies) with 3 studies conducted in primary care, 4 in occupational health and the remaining 6 studies from other settings. Prospective cohort studies were the most common design (15 studies) with retrospective cohorts (n=5) and health record reviews (n=2) being less frequently employed.

The outcome measure of work absence was defined differently in all studies, although the number of days absence from work was the most commonly used metric, this was calculated differently across studies, from:

- A simple count of days from company records as reported by Abenheim *et al* (1995)[26] and Bosman *et al* (2019)[27]

- Working or not working at 6 months as reported by Okurowski *et al* (2003)[28]

- To more complex calculations such as that reported by Nordin *et al* (1997)[29] where the days of absence were recorded from phone interviews with participants, or where compensated days were calculated as reported by Abenheim *et al* (1995[26] and Lederer *et al* (2014)[30]

The length of follow-up also varied ranging from 6 months or less in five studies[28], [29], [31], [32], [33], 12 months in five studies [34], [35], [36], [37], [38], two years in four studies [26], [39], [40], [41] and two with longer term follow-up at three [42] and five years [43].One study was not clear in the reporting of duration of absence, Shiels *et al* (2004)[44], however participants were followed-up for greater than one month.

Most of the studies reported prognostic factors only, however there were 13 prognostic models identified [27], [28], [33], [35], [37], [40], [41], [42], [45], [46], [47], [48], [49], some were models that had already been developed and were being tested in new populations and others were developed within a specific population.

***Risk of Bias: prognostic factors***

The summary judgements for each domain of the QUIPS tool are reported in Figure 3. Six of the included studies had at least one domain that was considered high risk [28], [36], [37], [38], [39], [42] and a further study was considered high risk overall due to the number of domains scoring moderate risk of bias [41]. The most common reason for a high risk of bias was study attrition, either through a large number of participants being lost to follow-up or studies not reporting attrition or the potential effect of this on the studies’ findings. The high risk of bias of these studies is reflected in the GRADE assessment. Just three studies were considered at low risk of bias [26], [35], [40], with the remaining five studies at moderate risk of bias with a lack of consideration or reporting of potential confounding being the most common domain to be reported as moderate risk.

**Figure 3: Risk of Bias (QUIPS) – Domain summary assessments**

{Here}

***Strength of the evidence: prognostic factors***

GRADE was used to assess the strength of the evidence and reported by grouping the prognostic factors into the 17 broad categories reported above (figure 1). For each of the categories the supporting research was assessed using an adapted GRADE criteria and an overall judgement was agreed (Table 3).

None of the categories were judged to be strong, with the strength of the evidence being low and with one theme “function” having a very low grading. The only factor that demonstrated an overall protective effect was age where an increasing age was associated with a lower risk of absence (reported in 9 studies).[28], [32], [35], [36], [38], [39], [40], [41], [44] All other themes were associated with a higher risk of absence however, comparisons within and between themes were impeded by the differing measures used across the studies.

***Summary of findings by personal systems***

Most of the categories included in personal systems (Figure 1) reported on inconsistent measures and outcomes and therefore provided a very mixed picture in terms of the contribution of that category to predicting work absence. There were some specific categories that warrant a fuller reporting as the direction of effect tended towards a more consistent direction, these are age, sex, recovery expectations and previous work absence.

*Age:* Age was reported in many different ways across each of the studies, however it did demonstrate an overall protective effect where increasing age was associated with a lower risk of absence. For example, Steenstra *et al* 2016 [48]reported age in 10 year increments from 15-25 years through to 55-65 years and found that for time on benefits there was a dose response effect when compared to the 25-35 year age group, with the 15-25 year age group reporting increased absence (hazard rate ratio 1.27; 95%CI 1.00, 1.60), and the older age groups reporting a lower risk of absence with increasing age.

*Sex:* Those studies reporting sex as a prognostic factor demonstrated no consistent direction of effect. [32], [35], [39], [40], [41], [44], [49] For example Abasolo *et al* 2008[39] found that women were less likely to experience temporary work disability when compared to men (HR 0.84 95%CI 0.78,0.90), however they was no difference in recurring work disability (HR 1.13 95%CI 0.97,1.32). Richter *et al* (2001)[35] reported that men were more likely to experience absence at follow-up when compared to women although not statistically significantly (HR 1.59 95%CI 0.78, 1.22). Steenstra *et al* (2015)[41] found that women were more likely to experience a recurrence of absence over two years follow-up (hazard rate ratio 1.36 (95%CI 1.09, 1.70).

*Recovery expectations:* Four of the studies included reported on recovery expectations, broadly the better a participants’ recovery expectations the better the outcome.[33], [34], [35], [48] For example, Turner *et al* (2006)[33] reported a dose response effect with recovery expectation of 0 (on a 0-10 scale) having an odds ratio of 9.18 (95%CI 5.00, 16.84) for 6-month work disability (defined as number of days on wage replacement) when compared to those with a very high recovery expectation. This odds ratio reduced to 1.95 (95%CI 1.18, 3.20) for those reporting a high recovery expectation of 8-9 (on a 0-10 point scale). However, it should be noted that there was no protective effect of recovery expectations in the study by Turner *et al* (2006). People who were unable to identify when they would return to work had a poorer outcome in the study by Richter *et al* (2011)[35] HR 0.23 (95%CI 0.15, 0.34) and those who reported they would return to work over a month later reported a HR of 0.24 (95%CI 0.15,0.38) indicating that a poorer recovery expectation was associated with a reduced “risk” of getting back to work.

*Previous absence:* Four studies reported on previous work absence and whether it can predict future absence.[32], [38], [39], [41] All studies measured previous work absence as a previous “claim”, the general direction of effect was of previous work absence being predictive of future work absence. For example Lederer *et al* (2014)[32] reported a HR of 0.91 (95%CI 0.87, 0.94) for previous claim history in the past 5 years (for return to work) and Van Dujin *et al* (2005)[38] found that prior sick leave (in the past 12 months) had a HR of 1.50 (95%CI 1.03, 2.17) at univariate analysis, but this variable was not included in the multivariable analyses.

There was some evidence that mental health may contribute to absence with Turner *et al* (2006)[33] reporting that mental health below the population mean (measured using the SF-36-v2) was associated with increasing absence however these were not significant results (<2 standard deviations (SD) below the mean OR 1.59 (95%CI 0.82, 2.08), 1-2 SD below the mean OR 1.84 (95%CI 0.99, 3.42) <1 SD below the mean OR 1.66 (95%CI 0.91, 3.03). Turner *et al* (2006) found no effect on absence related to catastrophising or blame. A high fear of movement was not reported to be associated with absence by Richter *et al* (2011)[35] OR 0.94 (95%CI 0.67, 1.33) however Turner *et al* (2006)[33] found that a high fear avoidance at 5-6 points measured with the Fear Avoidance Behaviour Questionnaire was associated with absence OR 4.64 (95%CI 1,57, 13.70) and at 3-5.9 points OR 2.96 (95%CI 0.98, 8.90).

General health and quality of life were reported by Richter *et al* (2011)[35] (general assessed with one question good versus poor) and Selander *et al* (2007)[36] (using the SF36) but there was no evidence of a relationship with work absence. Van Dujin *et al* (2005)[38] found that those participants who reported their musculoskeletal pain to be a chronic condition were more likely to experience absence OR 1.6 (95%CI 1.2, 2.32) however none of the other studies looked at this prognostic factor.

Pain was measured by three studies Richter *et al* (2011)[35], Lotters *et al* (2006)(low back pain and “other” MSK pain)[34], and van Dujin *et al* (2005)[38] using a 0-10 point likert scale, and all indicated that an increase in pain was significantly associated with work absence with effect sizes (OR) ranging between 1.1 to 1.3.

Steenstra *et al* (2015)[41] demonstrated a dose response effect with worsening functional ability, measured using a 0-4 scale, associated with time on absence benefits and risk of absence recurrence.

***Summary of findings by workplace systems***

Work schedule was examined by Absolo *et al* (2008)[39] who reported that being self-employed was protective of absence whilst having an indefinite work contract or being a “general” worker was associated with poor absence outcomes. Absolo *et al* (2008)[39] also found some indication that specific work demands related to movement e.g. frequent kneeling, flexion and rotation of the trunk was associated with absence however the effect sizes whilst generally significant were very small with OR between 1.05 to 1.39. Work culture was assessed by two studies both using different measures however both indicated that poor relationships at work and employer doubt about pain were indicators of absence. [33], [41] Richter *et al* (2011)[35] found that not being satisfied at work was again associated with absence. However, the availability of modified duties during sick leave[38] and continued salary during absence [41] were also indicative of absence.

**Results: prognostic models**

There was some overlap with studies reporting both individual prognostic factors and developing prognostic models. Overall, there were 13 prognostic models identified, some were models that had already been developed and were being tested in new populations and others were developed within a specific population.

***Description of included studies: prognostic models***

Table 2 reports the descriptive factors of the prognostic model studies. The majority of studies were undertaken using insurance health records (9 studies in total [28], [35], [37], [40], [41], [46], [47], [48], [49]), two studies in occupational health settings, [27], [45] one each in a general population [33] and primary care setting.[42] As with the prognostic factor papers the measures of absence were varied, there was no consistency in reporting and all studies used a different outcome measure.

*Reporting of models*

There was wide variation in the reporting of the models included in the review. Multivariable logistic regression was used by 3 studies [27], [28], [46] with logistic regression also reported by 3 studies [33], [42], [49] and one study reporting negative binomial regression.[40] A further 5 studies reported that Cox regression had been used for analysis.[35], [37], [41], [45], [48]

Validation was carried out in only half of the included studies with 7 studies reporting that internal validation had been undertaken, however, validation was not reported in 6 studies.[28], [33], [35], [40], [42], [46]

There was no consistency in the reporting of the models’ performance with most studies reporting the area under the curve or c-statistic [27], [28], [37], [41], [42], [46], [47], [48] and the other studies reporting the sensitivity and specificity, [27], [37], [42], [47] positive and negative predictive value.[27], [28], [47] Five studies did not report any measure of their models’ performance.[33], [35], [40], [45], [49]

None of the prognostic model papers reported the calibration of the models developed, so no observed:expected ratio or calibration slopes were presented and therefore no assessment on the calibration of the models included here could be made.

***Risk of Bias: prognostic models***

Figure 4 presents the overall judgement of risk of bias based on the domains of the PROBAST tool. Overall 62% of studies had a low risk of bias and also performed well in judgement of the domains assessing participants, predictors and outcomes. The main area for concern was the analysis where 38% of studies were at high risk of bias, this was often due to a lack of information reported in individual studies meaning assessment of how the analysis was performed was not able to be made.

**Figure 4: Risk of Bias (PROBAST) summary judgements**

{Here}

***Strength of the evidence: prognostic models***

Using the adapted GRADE to take account of the performance of the prognostic models included in this review, identified that the evidence for the use of the prognostic models was low (Table 4). This was primarily due to poor reporting of the models’ performance and one study who included a small percentage of participants who were not absent from work affecting indirectness.[42] Whilst the threshold for downgrading due to the risk of bias assessment was not met, it is worth noting that five of the 13 studies had a high or unclear risk of bias (not quite meeting the 50% required to downgrade).

**Discussion**

*Summary of main results*

A total of 23 studies were included in this review that all reported on prognostic factors for work absence in populations with musculoskeletal pain who were absent from work. Within these 23 studies 13 had developed prognostic models aimed at predicting absence from work. A total of 110 individual prognostic factors were identified and these were grouped into those related to personal systems and workplace systems aligned with the Disability Prevention Framework, within this overarching framework groups of prognostic factors were categorised for ease of comparison, these categories included all prognostic factors measuring the same concept (Figure 1). Overall, for both prognostic factors and prognostic models the strength of the evidence was low to very low. This grading of the evidence is due to the heterogeneous nature of the studies where prognostic factors, outcomes and timing of outcome measurement were different across studies, furthermore, reporting of model performance was also mixed with different statistics reported or performance measures not reported at all. The Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) Statement was published in 2015[50] and so was available for just two of the papers included in the review which may account for the issues in reporting that were seen when the prognostic models were synthesised. [27], [48]

*Study strengths*

We have followed the recommendations of each of the appropriate reporting checklists including; Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist and the CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies (CHARMS).[24] Furthermore, we have followed the guidance of Damen *et al* (2022)[51] who report a step-by-step guide to conducting systematic reviews of prognostic model studies including assessment of the performance of the models.

This review has comprehensively searched the literature on prognostic factors for work absence in those with musculoskeletal conditions. By considering how to group or categorise prognostic factors a priori using the Disability Prevention Framework,[13] we have ensured that sense could be made of the large number of prognostic factors identified. Being able to frame the impact of specific groups of prognostic factors within a framework mediated the impact of the heterogeneous nature of the studies, whilst we were unable to compare “like with like” we were able to assess the concepts overall and consider their contribution to predicting work absence.

The use of the GRADE system adapted to assess the strength of the evidence in prognostic factor and prognostic model studies has allowed summary judgements to be made and highlighted the inconsistencies in measurement and reporting of the studies included in the review. The adapted GRADE to include an assessment of the prognostic models’ performance has ensured that all available and pertinent data has been incorporated into the assessment of the strength of the evidence.[21]

*Study limitations*

There are some limitations to the current study, principally related the heterogeneity of the studies identified as part of this review. Whilst an individual patient data meta-analysis is often considered the gold standard (Cochrane https://methods.cochrane.org/ipdma/about-ipd-meta-analyses) in assessing the influence of a factor on an outcome it is not always possible when the quality of the studies is low. Due to the heterogeneous nature of the studies included and given that studies have controlled for different potential confounders, we were unable to consider any kind of meta-analysis, nor would this be wholly appropriate for this type of review. To address this and make sense of the varied measurements and outcomes, we aimed to categorise prognostic factors a priori and as far as possible assess the contribution of each category to predicting sickness absence. It was therefore important to ensure that synthesis of findings was as structured and transparent as possible. We had planned to use a narrative synthesis[22] but felt that the Synthesis without Meta-analysis (SWiM) framework was more suitable for this review as it provides a guide with which to group, describe and report the results of systematic reviews. The SWiM framework provides a more transparent method on how the studies’ findings were synthesised allowing a clear description of the findings to be reported and a more standardised approach to be followed when considering metrics and summaries of data.

*Comparison with other studies*

There are a number of reviews that are similar but focus on narrower populations. Kuijer *et al* (2006)[52] reviewed the literature exploring the prediction of sickness absence in patients with chronic low back pain, and found the same problems identified in our review with variable measurement of predictors, timing of follow-ups and differing definitions of outcome. Kuijer *et al* (2006)[52] concluded that no common set of core variables could be used to predict work absence in this specific population with chronic low back pain, the current review also noted that there was no common set of core predictor variables or even outcome measures or follow-up points, indicating that little has changed since the Kuijer *et al* (2006) review.[52] A recent Cochrane review by Hayden *et al* (2019)[53] focussed specifically on whether recovery expectations predict outcomes including work participation for which absence is a measure, in a population with non-specific low back pain. Hayden *et al* reported that there was moderate quality evidence that positive recovery expectations are strongly associated with better work participation. This finding is in part supported by the results of this review where broadly the better a participants’ recovery expectations the better the outcome. Other research has also identified previous absence as a predictor of future absences and whilst the evidence was weak in the current review the general direction of effect seen in this review supported this finding.[15], [54], [55]

A recent review from Ravinskaya *et al* (2022)[56] which assessed the reporting of work outcomes in randomised controlled trials also reported variability in work participation outcomes including work absence which was measured in the following ways: return to work rate, time to return to work, sick leave rate and sick leave duration. The authors concluded that a core outcome set for measurement of work participation is required and have gone on to develop that core outcome set recommending that studies including participants who are absent from work should report on the proportion of workers that return to work and time to return to work (Ravinskaya *et al* 2023).[57] This core outcome set would ensure that comparisons between studies can better be made and may allow more pooling of data to strengthen the body of evidence.

All the studies included in this review meet the criteria for exploratory prognostic studies and models in that they are describing associations and developing prediction models as described by Kent *et al* (2020).[58] Exploratory prognostic studies are usually carried out where little is known about a condition and they are an essential early step towards a confirmatory study.[58] However, given the number of studies included in this review and the number of prognostic factors measured it is difficult to argue that little is known about what predicts work absence in those with musculoskeletal pain. Whilst there will be important predictors not measured in these studies, our review indicates that there are commonalities in the concepts that may predict work absence but there is a wide variety in how the specific prognostic factors within these concepts are measured, the main concept that indicated any predictive ability was age, however age was measured in a variety of ways including “per year”,[39] in 5 year increments,[28] and in various categories[32], [35], [44] making meaningful comparisons between studies difficult. However, most prognostic studies within the field of musculoskeletal conditions are exploratory at present indicating that further research is needed to move this field forward.[59] In particular, by examining why there are differences in the extent to which models and factors predict absence.

**Conclusions**

This study has systematically reviewed the evidence for prognostic factors of future sickness absence in those with musculoskeletal conditions who are currently experiencing absence. Overall, the evidence for all prognostic factors was weak, although there was some evidence that older age and better recovery expectations were protective of future absence and that previous absence was likely to predict future absences. There was weak evidence for any of the prognostic models in determining future sickness absence. Analysis was difficult due to the wide range of measures of both prognostic factors and outcome and the differing timescales for follow-up. Future research should ensure that consistent measures are employed and where possible these should be in-line with those suggested by Ravinskaya *et al* (2023).[57]

**References**

[1] K. Zheltoukhova, “Musculoskeletal Disorders and Work: Results of a survey of  individuals living with Musculoskeletal Disorders in six European countries.,” London, 2013.

[2] G. Wynne-Jones *et al.*, “Absence from work and return to work in people with back pain: A systematic review and meta-analysis,” *Occup Environ Med*, vol. 71, no. 6, 2014, doi: 10.1136/oemed-2013-101571.

[3] N. Krause, L. K. Dasinger, L. J. Deegan, L. Rudolph, and R. J. Brand, “Psychosocial job factors and return-to-work after compensated low back injury: A disability phase-specific analysis,” *Am J Ind Med*, vol. 40, no. 4, pp. 374–392, Oct. 2001, doi: 10.1002/ajim.1112.

[4] P. M. Dekkers-Sánchez, W. Haije, J. K. Sluiter, and M. H. W. Frings-Dresen, “A qualitative study of perpetuating factors for longterm sick leave and promoting factors for return to work: Chronic work disabled patients in their own words,” *J Rehabil Med*, vol. 42, no. 6, pp. 544–552, Jun. 2010, doi: 10.2340/16501977-0544.

[5] “Overview | Workplace health: long-term sickness absence and capability to work | Guidance | NICE.” Accessed: Nov. 24, 2023. [Online]. Available: https://www.nice.org.uk/guidance/ng146

[6] C. Black, “Working for a  healthier tomorrow,” London, 2008.

[7] R. B. Keller *et al.*, “Programmatic Precision Oncology Decision Support for Patients With Gastrointestinal Cancer,” 2023, doi: 10.1200/PO.22.

[8] J. Delgadillo *et al.*, “Stratified Care vs Stepped Care for Depression: A Cluster Randomized Clinical Trial,” *JAMA Psychiatry*, vol. 79, no. 2, pp. 101–108, Feb. 2022, doi: 10.1001/jamapsychiatry.2021.3539.

[9] P. Williams, A. J. Dowson, A. M. Rapoport, and J. Sawyer, “The Cost Effectiveness of Stratified Care in the Management of Migraine,” 2001.

[10] J. C. Hill *et al.*, “Comparison of stratified primary care management for low back pain with current best practice (STarT Back): A randomised controlled trial,” *The Lancet*, vol. 378, no. 9802, pp. 1560–1571, Oct. 2011, doi: 10.1016/S0140-6736(11)60937-9.

[11] E. W. Steyerberg *et al.*, “Prognosis Research Strategy (PROGRESS) 3: Prognostic Model Research,” *PLoS Med*, vol. 10, no. 2, 2013, doi: 10.1371/journal.pmed.1001381.

[12] S. A. Ruhle and S. Süß, “Presenteeism and Absenteeism at Work—an Analysis of Archetypes of Sickness Attendance Cultures,” *J Bus Psychol*, vol. 35, no. 2, pp. 241–255, Apr. 2020, doi: 10.1007/s10869-019-09615-0.

[13] P. Loisel *et al.*, “Disability Prevention New Paradigm for the Management of Occupational Back Pain.”

[14] L. Hussey, K. Thorley, and R. Agius, “Reporting and prediction of work-related sickness absence by general practitioners,” *Occup Med (Chic Ill)*, vol. 66, no. 8, pp. 662–668, Nov. 2016, doi: 10.1093/occmed/kqw108.

[15] C. A. M. Roelen, P. C. Koopmans, J. A. H. Schreuder, J. R. Anema, and A. J. van der Beek, “The history of registered sickness absence predicts future sickness absence,” *Occup Med (Chic Ill)*, vol. 61, no. 2, pp. 96–101, Mar. 2011, doi: 10.1093/occmed/kqq181.

[16] R. Lilley, G. Davie, and S. Derrett, “Are the early predictors of long-term work absence following injury time dependent? Results from the Prospective Outcomes of Injury Study,” *BMJ Open*, vol. 7, no. 11, Nov. 2017, doi: 10.1136/bmjopen-2017-017390.

[17] M. J. Page *et al.*, “The PRISMA 2020 statement: An updated guideline for reporting systematic reviews,” *The BMJ*, vol. 372. BMJ Publishing Group, Mar. 29, 2021. doi: 10.1136/bmj.n71.

[18] R. D. Riley *et al.*, “A guide to systematic review and meta-analysis of prognostic factor studies,” *BMJ (Online)*, vol. 364. BMJ Publishing Group, Jan. 30, 2019. doi: 10.1136/bmj.k4597.

[19] J. A. Hayden, D. A. van der Windt, J. L. Cartwright, P. Côté, and C. Bombardier, “Assessing Bias in Studies of Prognostic Factors,” *Ann Intern Med*, vol. 158, no. 4, p. 280, Feb. 2013, doi: 10.7326/0003-4819-158-4-201302190-00009.

[20] R. F. Wolff *et al.*, “PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies,” *Ann Intern Med*, vol. 170, no. 1, pp. 51–58, Jan. 2019, doi: 10.7326/M18-1376.

[21] F. Foroutan *et al.*, “GRADE concept paper 2: Concepts for judging certainty on the calibration of prognostic models in a body of validation studies,” *J Clin Epidemiol*, vol. 143, pp. 202–211, Mar. 2022, doi: 10.1016/j.jclinepi.2021.11.024.

[22] J. Popay, L. Arai, and N. Britten, “Guidance on the conduct of narrative synthesis in systematic reviews: A product from the ESRC Methods Programme,” 2006, doi: 10.13140/2.1.1018.4643.

[23] M. Campbell *et al.*, “Synthesis without meta-analysis (SWiM) in systematic reviews: Reporting guideline,” *The BMJ*, vol. 368, Jan. 2020, doi: 10.1136/bmj.l6890.

[24] K. G. M. Moons *et al.*, “Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist,” *PLoS Med*, vol. 11, no. 10, 2014, doi: 10.1371/journal.pmed.1001744.

[25] T. P. A. Debray *et al.*, “A guide to systematic review and meta-analysis of prediction model performance,” *BMJ*, vol. 356, p. 6460, Jan. 2017, doi: 10.1136/BMJ.I6460.

[26] L. Abenhaim, M. Rossignol, D. Gobeille, Y. Bonvalot, P. Fines, and S. Scott, “The Prognostic Consequences in the Making of the Initial Medical Diagnosis of Work-Related Back Injuries,” *Spine (Phila Pa 1976)*, vol. 20, no. 7, pp. 791–795, Apr. 1995, doi: 10.1097/00007632-199504000-00010.

[27] L. C. Bosman, J. W. R. Twisk, A. S. Geraedts, and M. W. Heymans, “Development of Prediction Model for the Prognosis of Sick Leave Due to Low Back Pain,” *J Occup Environ Med*, vol. 61, no. 12, pp. 1065–1071, Dec. 2019, doi: 10.1097/JOM.0000000000001749.

[28] L. Okurowski, G. Pransky, B. Webster, W. S. Shaw, and S. Verma, “Prediction of Prolonged Work Disability in Occupational Low-Back Pain Based on Nurse Case Management Data,” *J Occup Environ Med*, vol. 45, no. 7, pp. 763–770, Jul. 2003, doi: 10.1097/01.jom.0000079086.95532.e9.

[29] M. Nordin *et al.*, “Early Predictors of Delayed Return to Work in Patients with Low Back Pain,” *J Musculoskelet Pain*, vol. 5, no. 2, pp. 5–27, Jan. 1997, doi: 10.1300/J094v05n02\_02.

[30] V. Lederer, M. Rivard, and S. D. Mechakra-Tahiri, “Gender Differences in Personal and Work-Related Determinants of Return-to-Work Following Long-Term Disability: A 5-Year Cohort Study,” *J Occup Rehabil*, vol. 22, no. 4, pp. 522–531, Dec. 2012, doi: 10.1007/s10926-012-9366-0.

[31] C. P. Gabel, M. Melloh, B. Burkett, J. Osborne, and M. Yelland, “The Örebro Musculoskeletal Screening Questionnaire: Validation of a modified primary care musculoskeletal screening tool in an acute work injured population,” *Man Ther*, vol. 17, no. 6, pp. 554–565, Dec. 2012, doi: 10.1016/j.math.2012.05.014.

[32] V. Lederer and M. Rivard, “Compensation benefits in a population-based cohort of men and women on long-term disability after musculoskeletal injuries: costs, course, predictors,” *Occup Environ Med*, vol. 71, no. 11, pp. 772–779, Nov. 2014, doi: 10.1136/oemed-2014-102304.

[33] J. A. Turner *et al.*, “Worker Recovery Expectations and Fear-Avoidance Predict Work Disability in a Population-Based Workers’ Compensation Back Pain Sample,” *Spine (Phila Pa 1976)*, vol. 31, no. 6, pp. 682–689, Mar. 2006, doi: 10.1097/01.brs.0000202762.88787.af.

[34] F. Lötters and A. Burdorf, “Prognostic Factors for Duration of Sickness Absence due to Musculoskeletal Disorders,” *Clin J Pain*, vol. 22, no. 2, pp. 212–221, Feb. 2006, doi: 10.1097/01.ajp.0000154047.30155.72.

[35] J. Richter, B. Blatter, J. Heinrich, E. de Vroome, and J. Anema, “Prognostic factors for disability claim duration due to musculoskeletal symptoms among self-employed persons,” *BMC Public Health*, vol. 11, no. 1, p. 945, Dec. 2011, doi: 10.1186/1471-2458-11-945.

[36] J. Selander, S.-U. Marnetoft, and M. Åsell, “Predictors for successful vocational rehabilitation for clients with back pain problems,” *Disabil Rehabil*, vol. 29, no. 3, pp. 215–220, Jan. 2007, doi: 10.1080/09638280600756208.

[37] M. Truchon, M.-È. Schmouth, D. Côté, L. Fillion, M. Rossignol, and M.-J. Durand, “Absenteeism Screening Questionnaire (ASQ): A New Tool for Predicting Long-term Absenteeism Among Workers with Low Back Pain,” *J Occup Rehabil*, vol. 22, no. 1, pp. 27–50, Mar. 2012, doi: 10.1007/s10926-011-9318-0.

[38] M. van Dujin, F. Lötters, and A. Burdorf, “Influence of modified work on return to work for employees on sick leave due to musculoskeletal complaints,” *J Rehabil Med*, vol. 37, no. 3, pp. 172–179, 2005.

[39] L. Abásolo *et al.*, “Prognostic factors in short-term disability due to musculoskeletal disorders,” *Arthritis Rheum*, vol. 59, no. 4, pp. 489–496, Apr. 2008, doi: 10.1002/art.23537.

[40] P. M. Smith, O. Black, T. Keegel, and A. Collie, “Are the Predictors of Work Absence Following a Work-Related Injury Similar for Musculoskeletal and Mental Health Claims?,” *J Occup Rehabil*, vol. 24, no. 1, pp. 79–88, Mar. 2014, doi: 10.1007/s10926-013-9455-8.

[41] I. A. Steenstra *et al.*, “Predicting Time on Prolonged Benefits for Injured Workers with Acute Back Pain,” *J Occup Rehabil*, vol. 25, no. 2, pp. 267–278, Jun. 2015, doi: 10.1007/s10926-014-9534-5.

[42] A. Westman, S. J. Linton, J. Öhrvik, P. Wahlén, and J. Leppert, “Do psychosocial factors predict disability and health at a 3‐year follow‐up for patients with non‐acute musculoskeletal pain? A validation of the Örebro Musculoskeletal Pain Screening Questionnaire,” *European Journal of Pain*, vol. 12, no. 5, pp. 641–649, Jul. 2008, doi: 10.1016/j.ejpain.2007.10.007.

[43] L. R. Sheehan, M. Di Donato, S. E. Gray, T. J. Lane, C. van Vreden, and A. Collie, “The Association Between Continuity of Care With a Primary Care Physician and Duration of Work Disability for Low Back Pain,” *J Occup Environ Med*, vol. 64, no. 10, pp. e606–e612, Oct. 2022, doi: 10.1097/JOM.0000000000002643.

[44] C. G. M. B. F. M. F. Shiels, “Patient factors associated with duration of certified sickness absence and transition to long-term incapacity,” *Br J Gen Pract*, vol. 54, no. 499, pp. 86–91, 2004.

[45] E. N. Branton, K. M. Arnold, S. R. Appelt, M. M. Hodges, M. C. Battié, and D. P. Gross, “A Short-Form Functional Capacity Evaluation Predicts Time to Recovery but Not Sustained Return-to-Work,” *J Occup Rehabil*, vol. 20, no. 3, pp. 387–393, Sep. 2010, doi: 10.1007/s10926-010-9233-9.

[46] M. Du Bois, M. Szpalski, and P. Donceel, “Patients at risk for long-term sick leave because of low back pain,” *The Spine Journal*, vol. 9, no. 5, pp. 350–359, May 2009, doi: 10.1016/j.spinee.2008.07.003.

[47] D. Fulton-Kehoe *et al.*, “Development of a Brief Questionnaire to Predict Long-Term Disability,” *J Occup Environ Med*, vol. 50, no. 9, pp. 1042–1052, Sep. 2008, doi: 10.1097/JOM.0b013e31817d361e.

[48] I. A. Steenstra, R.-L. Franche, A. D. Furlan, B. Amick, and S. Hogg-Johnson, “The Added Value of Collecting Information on Pain Experience When Predicting Time on Benefits for Injured Workers with Back Pain,” *J Occup Rehabil*, vol. 26, no. 2, pp. 117–124, Jun. 2016, doi: 10.1007/s10926-015-9592-3.

[49] M. Truchon, D. Côté, M.-È. Schmouth, J. Leblond, L. Fillion, and C. Dionne, “Validation of an Adaptation of the Stress Process Model for Predicting Low Back Pain Related Long-term Disability Outcomes,” *Spine (Phila Pa 1976)*, vol. 35, no. 13, pp. 1307–1315, Jun. 2010, doi: 10.1097/BRS.0b013e3181c03d06.

[50] G. S. Collins, J. B. Reitsma, D. G. Altman, and K. G. M. Moons, “Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement,” *BMC Med*, vol. 13, no. 1, Jan. 2015, doi: 10.1186/s12916-014-0241-z.

[51] J. A. A. Damen, K. G. M. Moons, M. van Smeden, and L. Hooft, “How to conduct a systematic review and meta-analysis of prognostic model studies,” *Clinical Microbiology and Infection*, vol. 29, no. 4. Elsevier B.V., pp. 434–440, Apr. 01, 2023. doi: 10.1016/j.cmi.2022.07.019.

[52] W. Kuijer, J. W. Groothoff, S. Brouwer, J. H. B. Geertzen, and P. U. Dijkstra, “Prediction of sickness absence in patients with chronic low back pain: A systematic review,” *Journal of Occupational Rehabilitation*, vol. 16, no. 3. pp. 439–467, Sep. 2006. doi: 10.1007/s10926-006-9021-8.

[53] J. A. Hayden, M. N. Wilson, R. D. Riley, R. Iles, T. Pincus, and R. Ogilvie, “Individual recovery expectations and prognosis of outcomes in non-specific low back pain: prognostic factor review,” *Cochrane Database of Systematic Reviews*, vol. 2019, no. 11. John Wiley and Sons Ltd, Nov. 25, 2019. doi: 10.1002/14651858.CD011284.pub2.

[54] R. LoMartire *et al.*, “Predictors of Sickness Absence in a Clinical Population With Chronic Pain,” *Journal of Pain*, vol. 22, no. 10, pp. 1180–1194, Oct. 2021, doi: 10.1016/j.jpain.2021.03.145.

[55] R. J. Reis, M. Utzet, P. F. La Rocca, F. B. Nedel, M. Martín, and A. Navarro, “Previous sick leaves as predictor of subsequent ones,” *Int Arch Occup Environ Health*, vol. 84, no. 5, pp. 491–499, Jun. 2011, doi: 10.1007/s00420-011-0620-0.

[56] M. Ravinskaya *et al.*, “Extensive variability of work participation outcomes measured in randomized controlled trials: a systematic review,” *J Clin Epidemiol*, vol. 142, pp. 60–99, Feb. 2022, doi: 10.1016/j.jclinepi.2021.10.013.

[57] M. Ravinskaya *et al.*, “Which outcomes should always be measured in intervention studies for improving work participation for people with a health problem? An international multistakeholder Delphi study to develop a core outcome set for Work participation (COS for Work),” *BMJ Open*, vol. 13, no. 2, Feb. 2023, doi: 10.1136/bmjopen-2022-069174.

[58] P. Kent, E. Boyle, C. Cancelliere, J. D. Cassidy, A. Kongsted, and A. Kongsted, “A conceptual framework for prognostic research,” *BMC Med Res Methodol*, vol. 20, no. 1, Jun. 2020, doi: 10.1186/s12874-020-01050-7.

[59] R. W. Wingbermühle, E. van Trijffel, P. M. Nelissen, B. Koes, and A. P. Verhagen, “Few promising multivariable prognostic models exist for recovery of people with non-specific neck pain in musculoskeletal primary care: a systematic review,” *J Physiother*, vol. 64, no. 1, pp. 16–23, Jan. 2018, doi: 10.1016/j.jphys.2017.11.013.

**Funding**

This work was funded as part of the Centre for Musculoskeletal Health and Work Medical Research Council/Versus Arthritis grant number: 22090

AL was funded through and Economic and Social Research (ESRC) Council Doctoral Training Program