Journal of Physiotherapy xxx (2017) xxx–xxx

Journal of PHYSIOTHERAPY

journal homepage: www.elsevier.com/locate/jphys

Appraisal

Research Note

Understanding the usefulness of prognostic models in clinical decision-making

Introduction

Over the past decade, prognostic models have become increasingly available for musculoskeletal conditions. Researchers have developed models to predict outcomes from back, neck, shoulder, elbow, knee and ankle pain.[1] Low back pain alone has inspired the development of more than 30 prognostic models.[2] Although widely available, few prognostic models have been successfully translated to clinical practice.[3]

Prognostic models generate individual risk profiles, which can inform decisions about the type and intensity of early management. The quality of a prognostic model is usually assessed by its ability to predict outcome. However, accuracy does not guarantee that a model will improve decision-making: some models with moderate to high predictive accuracy would offer little value over simply treating everybody, regardless of their risk profile.[4] Furthermore, statistical measures of predictive accuracy, such as discrimination, calibration, and model fit, are often difficult for clinicians to interpret and give no indication as to whether the model will improve decision-making.

Researchers have begun to assess the ability of prognostic models to inform the clinical decision-making process by using a decision curve analysis (DCA).[5] The DCA is a statistical approach that estimates the net benefit of basing clinical decisions on a patient's prognostic score, and compares this to the value of other decision-making strategies.

This Research Note describes the principles and limitations of DCA, and the clinical consequences of using prognostic models for clinical decision-making in physiotherapy practice. A glossary of terms is provided in Appendix 1 (see eAddenda for Appendix 1).

Validating a prognostic model

Prognostic models work by allocating individual patients a probability for developing a health outcome. When health outcomes are binary (one of two outcomes are possible eg, development of chronic pain, mortality), prognostic models produce estimates of probability. Also known as 'absolute risk', these probabilities can be expressed as a number between 0 and 1 or a percentage chance (0 to 100) of a health outcome occurring. When health outcomes are continuous (outcome has multiple levels eg, level of motor function or quality of life after stroke, costs of hospitalisation, days to recovery), rather than probabilities, prognostic models provide other estimates of predicted outcome. Binary outcomes tend to be a more popular choice for prognostic models than continuous outcomes.[6]

The first step in the validation process is to determine if the predictions that a model makes are accurate. Traditionally, researchers assess accuracy by testing the predictions in a sample of patients that is different to the sample where the model was developed: a validation cohort. In this cohort, researchers can determine to what extent risk estimates from a prognostic model are higher for those who experience poor outcomes versus those who experience good outcomes (discrimination) and how well these predicted risks match observed risks (calibration).

Unfortunately, discrimination and calibration are not easy to interpret clinically. For example, the statistics offer no guidance about how well a model should discriminate between good and poor outcomes and how correct the calibration should be before a clinician should decide to use the prognostic model in practice.

The gold standard for assessing the clinical consequences of using a prognostic model is by an impact trial, where patients are randomly allocated to either prognostic screening with matched recommendations (stratified care) or to usual practice. The results of an impact trial provide an unbiased estimate of whether stratified care improves outcomes compared to usual practice. Although they are essential aspects of prognostic research, impact trials are costly and time consuming.

In 2006, Vickers and Elkin[5] proposed the DCA not as an alternative to an impact study, but as a step towards deciding whether a model is likely to be useful for decision-making or not.

Net benefit

Prognostic models are likely to be useful if they can be shown to lead to more benefit and less harm than a one-size-fits-all decision-making strategy. Benefits and harms, in this context, refer to the consequences of clinical decisions. At a fundamental level, benefits occur when clinicians recommend the appropriate intensity of treatment to a given patient; harms occur when clinicians recommend too much treatment (overtreatment) or too little treatment (undertreatment).

DCA quantifies the trade-off between benefit and harm by placing them on the same scale: the net benefit. The net benefit accounts for how an individual clinician, in their decision-making, might balance the benefit of early treatment with the harms of overtreatment or undertreatment. The net benefit statistic combines aspects of discrimination, calibration, and model fit. The interpretation is relatively easy: net benefit is the proportion of high-risk patients who would be recommended early intervention appropriately, without increasing the number of early interventions recommended to low-risk patients.

Basing clinical decisions on prognosis

Every clinician has to make a decision on the course of action they will recommend to a patient. They might decide to treat everyone with a particular health condition the same way. For example, they might recommend self-management to every patient with low back pain, and vastus medialis exercises to every patient with anterior knee pain.

An alternative to the one-size-fits-all strategy is to recommend treatments, particularly intensive treatments, according to a patient's prognosis. A clinician might decide to recommend intensive treatments to only those with a high risk of poor outcome. For example, they might only recommend cognitive behavioural therapy to patients at high risk of chronic pain, or intensive rehabilitation programs to elderly patients who are at high risk of falling.

http://dx.doi.org/10.1016/j.jphys.2017.01.003
1836-9553/© 2017 Published by Elsevier B.V. on behalf of Australian Physiotherapy Association. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

UNCORRECTED PROOF

Please cite this article in press as: Traeger AC, et al. Understanding the usefulness of prognostic models in clinical decision-making. *J Physiother.* (2017), http://dx.doi.org/10.1016/j.jphys.2017.01.003

Each of these decision strategies carries a trade-off between benefit and harm. The consequences of 'treat all' strategies are that all high-risk patients receive early intervention (benefit) but a large number of low-risk patients are overtreated (harm).[7] The consequences of 'treat none' strategies are that no high-risk patients receive early intervention (harm),[7] but low-risk patients are spared unnecessary treatment (benefit).

The key question for the clinician is: which approach (ie, treat all, treat none or use a prognostic model) leads to the highest net benefit?

Probability threshold

There is an added level of complexity when clinicians want to assess the value of a prognostic model: the cut-off score. Using a prognostic model with a very low cut-off score will have similar consequences to a treat all strategy. A high cut-off score will have similar consequences to a treat none strategy. Therefore, the cut-off score is of crucial relevance to the usefulness of a prognostic model.

To determine a relevant cut-off, clinicians need to decide exactly what level of risk warrants further intervention. For example, some clinicians might recommend intensive rehabilitation for patients with > 20% risk of a poor outcome. Other more risk-averse clinicians might use 10% as their threshold. In the DCA this cut-off score is known as the probability threshold (Figure 1).

Although slightly abstract, the probability threshold reflects an implicit step in the decision-making process for all physiotherapists. Even when they use gut feelings to make a decision, clinicians apply an assumption of risk to every patient.

The probability threshold represents the benefit to harm trade-off that was discussed earlier. As such, the threshold will depend on the following:

1. How invasive is the treatment in question?
2. What is the likely outcome if *all* patients were to be recommended treatment versus not recommended treatment?
3. What is more important: not missing high-risk patients (ie, not undertreating) or not treating low-risk patients unnecessarily (ie, not overtreating)?

Even if a clinician doesn't know their threshold, they can estimate it by asking themselves the following question: how many unnecessary courses of intensive intervention would I be happy to recommend in order to provide early intervention to one patient who experienced a poor outcome (Table 1)?

The probability threshold is the starting point for interpreting a DCA. Once the probability threshold has been decided upon, one can assess which decision-making strategy leads to the highest net benefit.

Example

In acute low back pain, guidelines suggest that every patient receives minimal early management. However, around one in every three patients with acute low back pain develops chronic low back pain.[8] That means that one in every three would be undertreated with a minimal early management strategy. On the other hand, intensive early management for all patients is impractical and risks exposing large numbers of low-risk patients to unnecessary intervention (overtreatment). Targeted intensive management for those with a poor prognosis is a promising alternative.

A clinician could choose from the following decision-making strategies:

1. Recommend early, intensive treatment to all patients with acute low back pain (treat all).
2. Do not recommend early, intensive treatment to any patients with acute low back pain (treat none).
3. Recommend early, intensive treatment based on a prognostic model (prognostic screening).

Table 1
What is my probability threshold?

Number of unnecessary episodes of early management that I would be willing to recommend in order to prevent one patient experiencing a poor outcome	Probability threshold
2	0.5
3	0.33
5	0.20
10	0.10
20	0.05
50	0.02

Question

When deciding whether to recommend early, intensive management to a patient with acute low back pain, does using a prognostic model to screen patients for their risk of chronic low back pain lead to a net benefit, compared to a treat all or treat none decision-making strategy?

This question was examined in a recent study.[9] It investigated whether a prognostic model for acute low back pain could lead to more appropriate care than a treat all or treat none strategy. A DCA was performed and the results are shown in Figure 2. The DCA showed that for clinicians with a probability threshold between 15 and 35% (those who were happy to recommend unnecessary intervention to between three and seven patients to prevent one case of chronic pain), using the prognostic model would provide the highest net benefit. Risk-averse clinicians (eg, those who were happy to recommend upwards of seven patients unnecessary treatment) would be better off using a treat all strategy than the prognostic model. Clinicians who were not risk-averse (eg, those who were unhappy to recommend more than three unnecessary interventions) would be better off using a treat none strategy.

Application to physiotherapy research

Until recently, the DCA has been applied primarily to cancer and cardiovascular research. For example, researchers in the United States used a DCA to test whether a prognostic model based on imaging findings, biopsy results, or both, could reduce the number

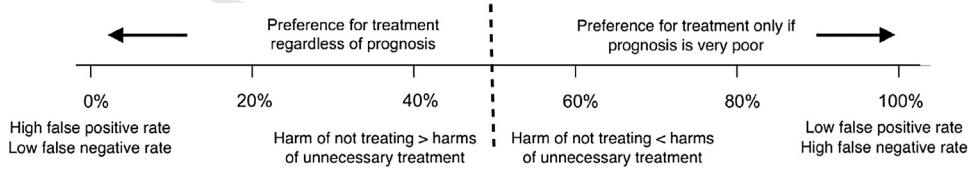

Figure 1. The probability threshold.
Number is the % risk of poor outcome, where a clinician would recommend further treatment. For example, a threshold of 20% means that a clinician would only recommend intensive treatment to patients with risk scores > 20%. A clinician with a low probability threshold (< 50%) weighs the consequences of undertreatment more heavily than the consequences of unnecessary treatment. Like low cut-off scores, low probability thresholds have a high false positive rate and low false negative rate (ie, they are sensitive but not specific). A clinician with a high probability threshold (> 50%) weighs the consequences of unnecessary treatment more heavily than the consequences of undertreatment. Like high cut-off scores, high probability thresholds have a low false positive rate and high false negative rate (ie, they are specific but not sensitive).

CORRECTED PROOF

Please cite this article in press as: Traeger AC, et al. Understanding the usefulness of prognostic models in clinical decision-making. *J Physiother.* (2017), http://dx.doi.org/10.1016/j.jphys.2017.01.003

Download English Version:

https://daneshyari.com/en/article/5564339

Download Persian Version:

https://daneshyari.com/article/5564339

Daneshyari.com