Classification systems for lower extremity amputation prediction in subjects with active diabetic foot ulcer: a systematic review and meta-analysis

CRD summary
The review concluded that there were numerous classification systems for diabetes foot ulcer outcome prediction, but few studies evaluated their reliability or external validity. The authors stated that further studies were needed to assess reliability and accuracy of the available systems and their composing variables. These conclusions appear reasonable given the limitations of the evidence.

Authors' objectives
To evaluate the accuracy of systems for predicting lower extremity amputation in patients with diabetes related foot ulcer.

Searching
MEDLINE, EBSCO, SCOPUS and ISI were searched for studies published up to May 2013. Some language restrictions were applied. Reference lists for included studies and relevant reviews were also consulted. Search terms were reported.

Study selection
Randomised controlled trials (RCTs), cohort and case–control studies that included people with diabetes and active diabetes foot ulcer were eligible for inclusion. Eligible prognostic factors were diabetes foot ulcer description or prognostic stratification classification systems. Studies had to evaluate the accuracy and/or reliability of these factors to predict lower extremity amputation occurrence.

Fifteen different classification systems were identified. Of those, eight used a description of diabetes foot ulcer (descriptive assessment where no specific prognostic was proposed) and seven used prognostic stratification classification systems (formal assessment where patients were grouped according to their risk of outcome occurrence). A range of prognostic factors were used (between 1 and 9 for each classification system), and the most common were: peripheral arterial disease; infection at the ulcer site; and/or ulcer depth. Variables included in the descriptive classification systems were more similar than those composing the prognostic stratification systems.

Two reviewers independently selected the studies. Disagreements were resolved by consensus.

Assessment of study quality
Two standard checklists were used to assess studies' quality of reporting: Reporting of Observational Studies in Epidemiology Studies (STROBE) and Standards for the Reporting of Diagnostic Accuracy studies (STARD). Studies were assigned a quality score for each checklist, with higher scores indicating better reporting quality. Reporting quality was assessed by a single reviewer.

Data extraction
Data on outcome prevalence (minor, major or total lower extremity amputation) were extracted to calculate reliability and prognostic accuracy measures, including sensitivity, specificity, positive and negative likelihood ratios, positive and negative predictive values. Prognostic accuracy cut-offs were extracted from the studies or calculated to evaluate the optimal balance between sensitivity and specificity. The authors did not report whether data were extracted in duplicate.

Methods of synthesis
Where possible, meta-analysis was conducted to calculate pooled unadjusted accuracy measures (sensitivity, specificity, LRs and area under the receiver operating characteristics (ROC) curve (AUC)) using a random effects model. Heterogeneity was assessed using $I^2$ and Cochran's $Q$. Patients who died during follow-up were removed from the analyses. Where a meta-analysis was not possible, results on reliability, validation and prognostic test accuracy were reported in tables and narratively.
Results of the review
Twenty five studies were included. Of those, 13 reported prognostic accuracy data. All except one study were prospective and retrospective cohort studies with moderate to good reporting. Reporting quality varied from 11 to 18 out of 22 (STROBE) and from 12 to 15 out of 25 (STARD).

Prevalence of lower extremity amputation ranged from 6% to 78%. Sensitivity of the classification systems ranged from 37.6% with the Curative Health Services wound grade scale (CHS), to 100% using the Depth of the Ulcer, Extent of bacterial colonization, Phase of ulcer and Association aetiology classification system (DEPA) score (1 study). Specificity ranged from 31% with the Scottish Intercollegiate Guidelines Network (SIGN) classification system, to 88% using the Size (Area, Depth), Sepsis, Arteriopathy, Denervation system (S(AD)SAD) score. Three studies reported major lower extremity amputation rates separately. Of those, Lipsky scores had higher sensitivity and specificity, followed by Meggit-Wagner and SIGN.

Meta-analysis was only possible for the accuracy of individual prognostic variables. Pooled analyses showed that diabetic peripheral neuropathy had the highest sensitivity (88%) and the lowest specificity (30%). On the other hand, gangrene had the lowest sensitivity (11%) and the highest specificity (95%). Positive likelihood ratios ranged from 1.22 (for diabetic peripheral neuropathy) to 5.50 (gangrene), negative likelihood ratios from 0.38 (diabetic peripheral neuropathy) to 0.91 (gangrene). Heterogeneity was generally high (Ι²>90%). Further results from the meta-analyses were reported.

Classification system reliability was evaluated in one study, and only four classification systems were validated by three or more studies. Further data on reliability and external validity were reported.

Authors' conclusions
There were numerous classification systems for diabetes foot ulcer outcome prediction, but few studies evaluated their reliability or external validity. Further studies were needed to assess reliability and accuracy of the available systems and their composing variables.

CRD commentary
The review question and selection criteria appeared clear. Several bibliographic sources were consulted, but unpublished studies were not sought. Studies were selected in duplicate, but it was unclear whether similar attempts were made to reduce error and bias for data extraction and quality assessment. Quality of study reporting was assessed and generally considered moderate. The actual quality of the conduct of the studies was unclear. A wide range of classification systems were used, and outcome prevalence and accuracy results were highly heterogeneous. Therefore for most results, no formal statistical synthesis was performed. Where meta-analyses were performed, sensitivity and specificity were analysed separately. This is not generally recommended and may lead to misinterpretation of the results.

The limited and heterogeneous nature of the data and the limited synthesis make the reliability of the results difficult to determine. However, given the limitations of the evidence, the authors' conclusions appear reasonable.

Implications of the review for practice and research
Practice: The authors stated that no classification system was ready for wide application and no independent predictive variables showed enough accuracy or consistency to propose a new classification or optimisation of the existing systems.

Research: The authors stated that further studies were needed to assess reliability and accuracy of the classification available systems and their composing variables. They stated that a systematic identifying all the studies evaluating predicting variables’ association with lower extremity amputation was needed to better understand the markers of outcome and optimise existing classifications.

Funding
One author was supported by Fundacao para a Ciencia e Tecnologia, (Portugal).

Bibliographic details

**PubMedID**
24523130

**DOI**
10.1002/dmrr.2535

**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Amputation /statistics & numerical data; Classification /methods; Diabetic Foot /classification /diagnosis; Humans; Lower Extremity /surgery; Models, Statistical; Prognosis; Reproducibility of Results; Risk Factors

**AccessionNumber**
12014012614

**Date bibliographic record published**
24/02/2014

**Date abstract record published**
14/03/2014

**Record Status**
This is a systematic review that meets the criteria for inclusion on DARE.