Risk stratification systems for diabetic foot ulcers: a systematic review

M. Monteiro-Soares · E. J. Boyko · J. Ribeiro ·
I. Ribeiro · M. Dinis-Ribeiro

Received: 20 September 2010 / Accepted: 6 December 2010 / Published online: 20 January 2011
© Springer-Verlag 2011

Abstract

Aims/hypothesis Several risk stratification systems have been proposed for predicting development of diabetic foot ulcer. However, little has been published that assesses their similarities and disparities, diagnostic accuracy and evidence level. Consequently, we conducted a systematic review of the existing stratification systems.

Methods We searched the MEDLINE database for studies (published until April 2010) describing the creation and validation of risk stratification systems for prediction of diabetic foot ulcer development.

Results We included 13 studies describing or evaluating the following different risk degree stratification systems: University of Texas; International Working Group on Diabetic Foot; Scottish Intercollegiate Guideline Network (SIGN); American

Diabetes Association; and Boyko and colleagues. We confirmed that five variables were included in almost all the systems: diabetic neuropathy, peripheral vascular disease, foot deformity, and previous foot ulcer and amputation. The number of variables included ranged from four to eight and the number of risk groups from two to six. Only four studies reported or allowed the calculation of diagnostic accuracy measures. The SIGN system showed some higher diagnostic accuracy values, particularly positive likelihood ratio, while predictive ability was confirmed through external validation only in the system of Boyko et al.

Conclusions/interpretation Foot ulcer risk stratification systems are a much needed tool for screening patients with diabetes. The core variables of various systems are very similar, but the number of included variables in each model and risk groups varied greatly. Overall, the quality of evidence for these systems is low, as little validation of their predictive ability has been done.

Keywords Clinical prediction rules · Diabetes · Diabetic foot · Diagnostic accuracy · Foot ulcer · Podiatry · Stratification systems · Systematic review

M. Monteiro-Soares (✉)

Serviço de Endocrinologia–Pé Diabético, Centro Hospitalar de Vila Nova de Gaia/Espinho EPE,
Unidade I, Rua Conceição Fernandes,
4434-502 Vila Nova de Gaia, Portugal
e-mail: mat.monteirosoares@gmail.com

E. J. Boyko

Department of Veterans Affairs Puget Sound Health Care System,
Seattle Epidemiologic Research and Information Center,
Seattle, WA, USA

E. J. Boyko

Department of Medicine, University of Washington,
Seattle, WA, USA

M. Monteiro-Soares · J. Ribeiro · M. Dinis-Ribeiro
Department of Biostatistics and Medical Informatics,
Oporto Faculty of Medicine, CINTESIS,
Oporto, Portugal

I. Ribeiro

Matosinhos Local Health Unit–Atlântida Extension,
Oporto, Portugal

Abbreviations

ABI	Ankle–brachial index
AUC	Area under the receiver operating characteristic curve
IWGDF	International Working Group on Diabetic Foot
PVD	Peripheral vascular disease
ROC	Receiver operating characteristic curve
STARD	Standards for the Reporting of Diagnostic Accuracy Studies
SIGN	Scottish Intercollegiate Guideline Network
STROBE	Strengthening of the Reporting of Observational Studies in Epidemiology

SWM	Semmes–Weinstein monofilament
UTFRS	University of Texas Foot Risk Stratification
VPT	Vibration perception threshold

Introduction

Diabetes mellitus is one of the most frequent metabolic disorders [1, 2], achieving an epidemic magnitude [2], of nearly 3% prevalence worldwide [3], with an expected increment to more than 4% in 2030 [3, 4]. This significant rise combined with insufficient healthcare resources will make it increasingly necessary to further improve prevention and treatment of diabetic foot complications [5].

In diabetic populations, amputations are 15 [3, 6] to 40 [1] times more frequent than in persons without diabetes. Foot ulcer is the major predisposing factor for non-traumatic foot amputations [4], preceding about 85% of them [4, 7]. Furthermore, after a lower limb amputation the risk of additional amputations is 50% in 5 years; the mortality rate is about 70% [8].

It has been reported that an effective evidence-based prevention programme (with early detection and control of independent risk factors for foot ulceration), patient and carer education, foot ulcer treatment by a multidisciplinary team and periodic surveillance can diminish the amputation rate by 49% to 85% [9]. Hence, various studies have concluded that amputation is always more expensive than its prevention [10]. Therefore, it is crucial to define a standardised and efficient approach to prevention of foot ulceration and consequently amputation [2]. The first step should be the correct identification of degree of risk for foot ulceration in all patients [11–13].

At present, numerous stratification systems using different methods have been proposed for this purpose [9, 14], but there are few validation studies [15], leading to the problem of how to select the best system for widespread implementation. Our aim was to conduct a systematic review of the existing risk stratification systems for the development of diabetic foot ulcer in order to compare them with regard to selection of variables, development of prediction model, diagnostic accuracy measures, validation and generalisability. Additionally we aimed to better understand the potential for this decision tool to impact clinical care.

Methods

To conduct this systematic review, we carried out a sensitive search in MEDLINE database (PubMed) for studies that were published up to 15 April 2010 and analysed diabetic foot ulcer risk stratification systems. The query used is shown in Fig. 1.

This search retrieved 2,275 studies. These were considered further if they met the following selection criteria: (1) publication date up to and including 15 April 2010; (2) published in the following languages: English, French, Italian, Spanish or Portuguese; (3) type of study: reviews, randomised controlled trials or cohort, case-control and cross-sectional studies; (4) studies that described the creation of or evaluated diabetic foot ulcer risk degree stratification systems; and (5) results that described the creation or modification (by the same group) and/or evaluated the effectiveness of one or several diabetic foot ulcer risk degree stratification systems.

Initially, articles were masked as to the identities of the authors, institutions and journals, and then selected by assessing their pertinence on the basis of titles and abstracts (when available) by two investigators (M. Monteiro-Soares, J. Ribeiro), who worked independently and were blinded to each other's assessments. In this phase the most common cause for exclusion was an article's theme.

Fig. 1 Systematic review: flow diagram of article selection process. Studies were retrieved using the following query: (“Diabetic Foot/blood”[Mesh] OR “Diabetic Foot/classification”[Mesh] OR “Diabetic Foot/complications”[Mesh] OR “Diabetic Foot/diagnosis”[Mesh] OR “Diabetic Foot/epidemiology”[Mesh] OR “Diabetic Foot/etiology”[Mesh] OR “Diabetic Foot/mortality”[Mesh] OR “Diabetic Foot/pathology”[Mesh] OR “Diabetic Foot/physiopathology”[Mesh] OR “Diabetic Foot/prevention and control”[Mesh] OR “Diabetic Foot/radiography”[Mesh] OR “Diabetic Foot/radionuclide imaging”[Mesh] OR “Diabetic Foot/surgery”[Mesh] OR “Diabetic Foot/ultrasonography”[Mesh] OR “Diabetic Foot/urine”[Mesh] OR (diabetes AND ulcer AND lesion)) AND ((predict*[tiab] OR predictive value of tests [mh] OR scor*[tiab] OR observ*[tiab] OR observer variation[mh]) OR (incidence[MeSH:noexp] OR mortality[MeSH Terms] OR follow up studies[MeSH:noexp] OR prognos*[Text Word] OR predict*[Text Word] OR course*[Text Word]) OR (sensitiv*[Title/Abstract] OR sensitivity and specificity[MeSH Terms] OR diagnos*[Title/Abstract] OR diagnosis[MeSH:noexp] OR diagnostic * [MeSH:noexp] OR diagnosis,differential[MeSH:noexp] OR diagnosis[Subheading:noexp]) OR (cohort OR case-control OR prospective OR “risk factor” OR screening))

In a second phase, the previously chosen articles ($n=37$) were examined in their entirety (with the respective reference list) and selected for inclusion for this review by the same two investigators who had performed the initial review, again acting independently and blinded. As in every stage, divergence was resolved by the decision of a third investigator (I. Ribeiro). At the end of this stage, seven articles were included in this systematic review.

Finally, after analysing the reference list of all the selected articles and relevant reviews that had been excluded, new articles were found. These were subjected to the first and second phases, and included or excluded from the study. This procedure was repeated until no new article was found through the reference list analysis, resulting in the inclusion of six more articles. In conclusion, 13 studies were included in this review (Fig. 1).

The review of title and/or abstract led to disagreements between the two reviewers in 36 cases, making for 98% inter-observer agreement and a kappa value of 0.61. The same occurred in the selection of papers reviewed in their entirety, where the two reviewers disagreed on the inclusion of four studies for 95% inter-observer agreement and a kappa value of 0.9.

Once article selection was completed, the following data were collected from each article using a checklist created for this review: (1) article identification: title, author(s), publication date, journal; (2) outcome definition; (3) methods: study design, setting, period(s) of data collection, inclusion and exclusion criteria, sources and methods of participant selection, sample size, clinical factors analysed, diagnostic tests analysed, potential bias; (4) results: study participant characteristics, outcome prevalence, method of statistical analysis, risk categorisation diagnostic accuracy measures; and (5) quality assessment. The articles' quality was assessed (by M. Monteiro-Soares) through the number of items fulfilled in the corresponding checklist, selected according to type of study, i.e. the Strengthening of the Reporting of Observational Studies in Epidemiology [STROBE] checklist for observational studies and the Standards for the Reporting of Diagnostic Accuracy Studies [STARD] checklist for diagnostic accuracy studies [16, 17]. Both checklists have multiple components per item, which caused difficulties in scoring. We therefore stipulated that total completion of an item should score 1 point, partial completion 1/2 point and null completion 0 points.

Results

Foot ulcer risk stratification systems identified

We retrieved five stratification systems, discussed in 13 papers (Table 1): (1) University of Texas Foot Risk

Stratification (UTFRS, $n=1$) [18]; (2) International Working Group on Diabetic Foot (IWGDF, $n=4$) [9, 14, 19, 20]; (3) Scottish Intercollegiate Guideline Network (SIGN) Risk Assessment ($n=2$) [6, 21]; (4) American Diabetes Association (ADA, $n=4$) [11, 12, 22, 23]; and (5) Boyko et al. model ($n=2$) [24, 25].

Examining Table 2, which lists the variables included in each stratification system, we observed that the majority had identical core variables, namely: diabetic neuropathy, peripheral vascular disease (PVD), foot deformity, previous ulcer and previous lower extremity amputation. On the other hand, data collection procedures differed greatly between studies for diagnosis of diabetic neuropathy and PVD.

The number of variables included varied from four [18] to eight [21] and the number of risk groups varied from two in the original ADA system [22, 23] to six in the IWGDF system modified by Lavery et al. [20] (Table 3).

The Leese et al. study had the biggest sample size [21], while the Boyko et al. study [24] had the longest follow-up (Table 1).

It was only possible to analyse or calculate diagnostic accuracy measures (sensitivity, specificity, predictive values) in three studies [18, 21, 25] (Table 4). In the studies where diagnostic accuracy measures or crude data were not displayed, they were calculated by or requested from the authors [18, 20, 24], respectively. Unfortunately, it was not possible to obtain these data. In the Boyko and colleagues study [24], only the area under the receiver operating characteristic curve (AUC) and cut-off values were available, which did not allow direct comparison with the effectiveness of other stratification systems.

Using the STROBE checklist, the Peters et al., Boyko et al. and Monteiro-Soares et al. studies had the best scores (all with 18 points out of 22) [19, 24, 25]. For studies where diagnostic accuracy measures were reported the STARD checklist was also applied. The Leese et al. and Boyko et al. studies had 15 items, while Monteiro-Soares et al. had 20 (out of 25) [21, 24, 25].

Only with the SIGN system was reliability assessed through the kappa value for inter-observer agreement calculation [21]. No validity testing of the ADA system has ever been performed, while the SIGN and UTFRS systems have been validated once [18, 21]. The IWGDF system suffered modifications twice and its validation was performed accordingly [19, 20]. The Boyko et al. system was the only one externally validated [24, 25]. Conversely, the IWGDF was the only group that applied worldwide dissemination techniques (manuals and CD distribution, website creation and others) for the system described in Apelqvist et al. articles [9, 14].

Table 1 Stratification systems: characterisation and classification of studies

Study [ref.]	Stratification	Creation method	Step	Sample size (n)	Mean follow-up (months)	Ulcer prevalence (%)	STROBE score
Lavery et al. [18]	UTRFS	Logistic regression model	Derivation	225	Cross-sectional	34	13
Apelqvist et al. [9]	IWGDF	International consensus	Description	NA	NA	NA	NA
Peters et al. [19]	IWGDF	International consensus	Modification proposal, evaluation	213	30	25	18
Lavery et al. [20]	IWGDF	International consensus	Modification proposal, evaluation	1,666	27	15	12
Apelqvist et al. [14]	IWGDF	International consensus	Description	NA	NA	NA	NA
SIGN [6]	SIGN	Literature review	Description	NA	NA	NA	NA
Leese et al. [21]	SIGN	Literature review	Modification proposal, evaluation	3,526	20	5	15 ^a
Mayfield et al. [22]	ADA	Literature review	Description	NA	NA	NA	NA
Mayfield et al. [23]	ADA	Literature review	Description	NA	NA	NA	NA
Boulton et al. [11]	ADA	Literature review	Modification proposal	NA	NA	NA	NA
Boulton et al. [12]	ADA	Literature review	Modification proposal	NA	NA	NA	NA
Boyko et al. [24]	Boyko	Logistic regression model	Derivation	1,285	40	17	18 ^a
Monteiro-Soares et al. [25]	Boyko	Logistic regression model	External validation, optimisation proposal	360	25	26	18 ^b

For the construction of this table, studies were ordered by stratification system (through their creation date) and within each one chronologically by publication date

^a With STARD checklist 15 points;

^b With STARD checklist 20 points

NA, not applicable

Foot ulcer risk stratification systems: data synthesis

The UTRFS system This system was described for the first and only time in 1998, by Lavery and colleagues, in a cross-sectional case–control study that enrolled 213 participants with diabetes: 76 cases with an existing or recently healed (<4 weeks) foot ulcer and 149 (controls) without active or previous foot ulcer. This study was performed in the setting with the highest foot ulcer prevalence (34%) [18].

First, the association between foot ulceration and several variables was evaluated through univariate analysis. Next, they analysed the cumulative risk associated with the significant variables more frequently available in daily practice: diabetic neuropathy, foot deformity and ulcer or lower extremity amputation history. This resulted in the stratification system presented in Table 3 (very similar to that proposed in 2000 by the IWGDF). For each added variable the cumulative risk increased. For category 1 the OR for foot ulceration was 1.7 (95% CI 0.7–4.3), for

category 2 it was 12.1 (95% CI 5.2–28.3) and for category 3 it was 36.4 (95% CI 16.1–82.3) in comparison with category 0 (reference category) [18]. Despite the use of logistic regression, no score for a risk group calculation was created.

A vibration perception threshold (VPT) >25 V, using a biothesiometer, indicated diabetic neuropathy. No foot ulcer definition was provided.

It was not possible to calculate any diagnostic accuracy measures due to a lack of cross-tabulation with the number of cases and controls in each risk stratification group; it was also not possible to retrieve these data from the article's first author.

The IWGDF system This stratification system was created through consensus involving 45 expert clinicians and researchers from 23 countries [9, 19]. Although there is an 8 year interval between them, both papers by Apelqvist and colleagues [9, 14] are very similar. They recommend use of the 10 g Semmes–Weinstein monofilament (SWM),

Table 2 Variables included in the diverse stratification systems

Stratification	Variables										
	DN	PVD	Foot deformity	Previous ulcer	Previous amputation	Visual impairment	Physical impairment	Callus	HbA _{1c}	Tinea pedis	Onychomycosis
UTFRS	O		O	O	O						
IWGDF	O/R	O/R	O/R	O/R	O/R						
SIGN	O/R	O/R	O/R	O/R	O/R	O/R	O/R	O/R			
ADA	O/R	O/R	O/R	O/R	O/R						
Boyko et al.	O/R			O/R	O/R	O/R			O/R	O/R	O/R

DN, diabetic neuropathy; O, present in the original stratification; R, present in the revised stratification

tuning fork and/or cotton wisp for detection of diabetic neuropathy [9, 14]. However, to our knowledge, no study has analysed the cotton wisp's diagnostic ability. Moreover, use of three tests simultaneously or alone presents different accuracy values.

This stratification system has never been validated in its original form for the prediction of foot ulcer development. In 2001, its effectiveness was evaluated, but by this time small modifications had been effected [19]. In a prospective cohort study with 213 participants followed for a mean period of 30 months, this stratification system (Table 1) was evaluated for prediction of diabetic foot ulceration, i.e. skin lesions distal to the ankle [19].

With the stratification system, there was a statistically significant increase in frequency of ulceration and amputation ($p<0.001$, χ^2 test) in the higher risk groups. Individuals in group 3 (higher risk) were 34.1 (95% CI 11.0–105.8) times more prone to foot ulcer occurrence during the follow-up period [19]. Although these results indicate good effectiveness, no diagnostic accuracy measures were reported, although it would have been possible to calculate them (Table 4).

Diabetic neuropathy was defined as one or more insensitive sites to the 10 g SWM or a VPT >25 V. An ankle-brachial index (ABI) inferior to 0.8 or any non-palpable pedal pulsation was defined as PVD [19]. The biothesiometer is not commonly available due to its cost. However, the authors stressed that a 128 Hz tuning fork can be used as an alternative, alleging good correlation, based on a single study [19].

Peters and colleagues proposed a subdivision in group 3, separating patients with history of foot ulceration from those with history of lower-extremity amputation [19]. In 2008, Lavery and colleagues included this modification in the stratification system and also proposed a subdivision for group 2 (Table 3). This prospective cohort study included 1666 consecutive participants followed for an average of 27 months [20]. An increase in the group risk was associated with more foot ulcerations ($p<0.001$, χ^2 test for association and trend) and more complications were

observed in group 2B than in 2A ($p<0.001$). This did not occur when comparing group 1 with 2A or group 3A with 3B [20]. No diagnostic accuracy measures were presented in the paper and due to a lack of data, it was impossible to determine them.

Foot ulcer was defined as a full-thickness wound involving the foot or ankle. Diabetic neuropathy was assessed using the 10 g SWM and the biothesiometer [20]. Although the authors did not describe how the diagnosis was established in this paper, they referred to another article where details of the diagnosis are given [26]. One non-palpable foot pulse combined with an ABI inferior to 0.8 indicated PVD [20].

The SIGN system This stratification system was created at the same time as that from the IWGDF through an evidence-based systematic review performed by a multidisciplinary group (Table 3) [6].

It has never been validated in its originally conceived form. In 2006, Leese and colleagues validated it with slight modifications in a prospective cohort study in a community setting (foot ulcer prevalence of 5%). In sum, individuals with no risk factors were considered at low risk of foot ulcer occurrence; those with one risk factor were at moderate risk; and individuals with two or more risk factors or with foot ulcer history were at high risk [21].

In this study, diabetic neuropathy was detected through the 10 g SWM. Inability to feel the monofilament on more than one of ten pre-defined sites was ranked as altered sensation. This study was the only one in this review to assess inter-observer agreement of a stratification system in 50 participants by two healthcare professionals, resulting in a kappa value of 0.95. The main quality of this stratification system is to identify individuals at very low risk of developing a foot ulcer. Thus patients in the low-risk group had a 99.6% probability of not developing a foot ulcer during follow-up [21].

The ADA system This system was created through a literature review. Initially, some variables were recognised

Table 3 Foot ulcer risk stratification system risk group description

UTRFS			IWGDF			SIGN			ADA		
Lavery et al. [18]			Apelqvist et al. [9, 14]			Peters et al. [19]			Lavery et al. [20]		
RG	Definition		RG	Definition		RG	Definition		RG	Definition	
0	Without DN	0	Without DN	0	Without DN	0	Without DN	0	Without DN or PVD	Low	No risk of FU
1	DN	1	DN	1	DN but without FD or PVD	1	DN but without FD or PVD	1	DN but without FD or PVD	Low	No risk of FU
2	DN and FD	2	DN and/or FD or PVD; and/or history of FU or LEA	2	DN and with FD or PVD	2A	DN and FD, no PVD	2	DN and/or PVD or significant VI or PI	Medium	DN or PVD or FD or VI or PI
3	DN and FD history of LEA	3	History of FU or LEA	3	History of FU or LEA	3A	History of foot ulcer	3	History of FU or LEA or PVD and DN or one of above, with callus or deformity	High	DN, altered biomechanics, evidence of increased pressure, FD, PVD, history of FU or LEA and/or severe nail pathology
						3B	LEA				

In this table, the Boyko et al. system was not included due to the fact that patient's risk group attribution depends on a score calculation with multiple possible variable combinations and not in a static number of risk factors being present

DN, diabetic neuropathy; FD, foot deformity; FU, foot ulcer; LEA, lower extremity amputation; PI, physical impairment; RG, risk group; SIGN, SIGN risk assessment; VI, visual impairment

Table 4 Diagnostic accuracy measures for each foot ulcer risk stratification system

Stratification system	Measures						
	Risk group	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR- (95% CI)	PPV (95% CI)	Accuracy (95% CI)
UTRFS [18]		NP	NP	NP	NP	NP	NP
IWGDF: Apelqvist et al. [9]		NA	NA	NA	NA	NA	NA
IWGDF: Peters et al. [19]	3	74 (62–86)	86 (81–92)	5.35 (3.52–8.14)	0.30 (0.19–0.47)	64 (58–70)	83 (78–88)
IWGDF: Peters et al. [19]	3+2	87 (78–96)	58 (51–66)	2.10 (1.70–2.59)	0.22 (0.11–0.45)	NA	66 (59–72)
IWGDF: Lavery et al. [20]		NP	NP	NP	NP	NP	NP
IWGDF: Apelqvist et al. [14]		NA	NA	NA	NA	NA	NA
SIGN [6]		NA	NA	NA	NA	NA	NA
SIGN: Leese et al. [21]	High risk	84 (79–90)	90 (89–91)	8.41 (7.45–9.49)	0.17 (0.12–0.25)	31 (29–33)	90 (89–91)
SIGN: Leese et al. [21]	High+moderate risk	95 (92–98)	67 (65–68)	2.87 (2.70–3.04)	0.07 (0.04–0.14)	NA	68 (67–70)
ADA [11, 12, 22, 23]		NA	NA	NA	NA	NA	NA
Boyko et al. [24]	ROC	NP	NP	NP	NP	NP	81
Monteiro-Soares et al. [25]	Highest risk	61 (51–70)	87 (83–91)	4.7 (3.33–6.76)	0.45 (0.35–0.58)	62 (57–67)	80 (76–84)
Monteiro-Soares et al. [25]	Highest+next-to-highest risk	84 (75–90)	70 (65–75)	2.83 (2.34–3.47)	0.23 (0.14–0.36)	NA	74 (69–79)
Monteiro-Soares et al. [25]	Highest+next-to-highest+next-to-lowest risk	95 (88–98)	50 (44–56)	1.88 (1.65–2.13)	0.10 (0.05–0.25)	NA	61 (56–66)
Monteiro-Soares et al. [25]	ROC	NA	NA	NA	NA	NA	83 (78–88)

LR-, negative likelihood ratio; LR+, positive likelihood ratio; NA, not applicable; NP, not possible to calculate with the available data; PPV, positive predictive value

as related to foot ulcer development (namely diabetic neuropathy, PVD, foot deformity, and foot ulcer or amputation history) and anyone presenting with any of these conditions was considered to be at high risk [22, 23]. In 2008, a modification was proposed. Using the same variables, Boulton and colleagues, proposed a stratification system that graded by estimated cumulative risk [11, 12] (Table 3).

Diabetic neuropathy screening was recommended using the 10 g SWM and one of the following other tests: 128 Hz tuning fork, pinprick sensation, ankle reflex or VPT. An abnormal result in one or more tests suggested loss of protective sensation. Absence of the posterior tibial and/or dorsalis pedis pulses indicated PVD [11, 12].

This stratification system has been described in four articles. However, the two articles by Mayfield et al. [22, 23] are identical, as are the two by Boulton and colleagues [11, 12]. None of the ADA stratification systems were validated for prediction of ulcer development [19].

The Boyko et al. system This stratification system was developed in a study that prospectively followed 1285 veterans (98% men) over more than 3 years, with re-evaluations at 12 to 18 months, with a view to evaluating the ‘individual and combined effects of commonly available clinical information in the prediction of diabetic foot ulcer occurrence’ [24]. Several available and pertinent variables were assessed at baseline. Using a Cox proportional hazards regression model, the association between baseline variables and foot ulcer occurrence was evaluated through univariate and multivariate analysis, resulting in the following risk score equation, where a one (1) was inserted when the characteristic in parentheses was present: $\text{score} = \text{HbA}_{1c} \times 0.0975 + 0.7101 \times (\text{diabetic neuropathy}) + 0.3888 \times (\text{poor vision}) - 0.3206 \times (\text{tinea pedis}) + 0.4579 \times (\text{onychomycosis}) + 0.7784 \times (\text{history of foot ulcer}) + 0.943 \times (\text{history of lower limb amputation})$ [24].

According to the resultant score, participants were stratified into the following risk groups: (1) lowest risk (score < 1.48); (2) next-to-lowest risk (score 1.48 to ≤ 1.99);

(3) next-to-highest risk (score 2.00 to ≤ 2.61); and (4) highest risk (score ≥ 2.62).

This was the only study that included HbA_{1c} as a predictive variable and assessed the stratification system's ability to predict foot ulcer occurrence through a receiver operating curve (ROC) at 1 and 5 years from the start of follow-up, resulting in AUCs of 0.81 and 0.76 respectively [24].

Analysing the ROC curve at 1 year, it can be seen that a specificity of 86% corresponded to a sensitivity of 60%, while 80% sensitivity corresponded to 60% specificity [24]. However, it was not possible to calculate any other diagnostic accuracy measures for the different groups or the AUC confidence intervals, due to lack of data.

Foot ulcer was defined as a full-thickness skin defect that needed more than 14 days to heal. Diabetic neuropathy was diagnosed by applying a 10 g SWM to nine sites in each foot. Insensitivity in one or more sites indicated altered sensation. In this stratification system, PVD was not included [24].

The Boyko et al. stratification system, as originally proposed, was externally validated in a 2010 retrospective cohort study including 360 participants [25]. They were followed for 25 months (mean) and 26% developed an ulcer (using the same definition as in the study of Boyko et al. [24]). Inability to feel the 10 g SWM at one or more of eight points (four in each foot) was considered to indicate diabetic neuropathy [25].

In univariate analysis, six of the seven variables included in the Boyko model were also significantly associated with foot ulcer development in the external validation study [25]. Tinea pedis, as in the Boyko et al. study [24], showed a statistically significant association only in multivariate analysis [25]. Diagnostic accuracy measures were described, and the resulting AUC and respective confidence intervals (AUC 0.83; 95% CI 0.78–0.88) [25] included the value reported by Boyko and colleagues [24] (Table 4). Additionally, an increase in the group risk was associated with a higher risk of foot ulcer development ($p < 0.001$ χ^2 test for association and trend). This study also demonstrated that the model proposed by Boyko et al., which had been originally developed in a predominantly male population, was equally accurate in both sexes; it also reported that including a variable referring to footwear could improve this model's accuracy, although not to a statistically significant degree (AUC 0.88, 95% CI 0.84–0.91) [25].

Discussion

Stratification systems are an essential tool for classifying patients according to a cumulative risk of foot ulcer development and consequently allowing the limited exist-

ing medical care resources to be distributed to those at most need [4, 18, 21]. Doing so may diminish the unreasonably high level of foot-related morbidity [11, 12]. However, no system has been unanimously adopted [4] and their implementation in clinical practice is scarce [21]. Consequently we felt the necessity to perform a systematic review in order to understand whether and how these systems could facilitate clinicians' and researchers' choice when it comes to future implementation and development.

Overall, we retrieved five stratification systems, but it was only possible to determine the effectiveness of three of them through diagnostic accuracy measures [19, 21, 25].

The UTRFS stratification system derivation study [18] is a cross-sectional case-control study and therefore has a very low evidence level and some possible bias. We believe that having as an outcome the presence of or recently healed ulcer (without definition) could introduce selection as well as information bias, due to the absence of blinding to presence of the condition. In addition, there are concerns about adequacy of sample size (taking in consideration the reported appropriated number of subjects for each predictive variable's detection) [27]. This study assessed, in univariate analysis, the association between 27 different variables and ulceration in a sample of 213 participants [18].

We reviewed four articles related to the IWGDF stratification system, two describing it [9, 14] and two evaluating its effectiveness [19, 20]. However, each study presents modifications (without reported statistical justification), suggesting that this stratification scheme is still under development. In addition, only one study allowed the calculation of diagnostic accuracy measures and ulcer definition, and diagnosis of diabetic neuropathy and of PVD was somewhat different in each study. As with the UTRFS system, the use of ABI and the biothesiometer is somewhat difficult in daily practice due to its cost and/or need of trained professionals. In the study by Peters et al. [19], patients with a diabetic foot ulcer that directly led to amputation were excluded in order to reduce selection bias, since these patients had a priori a higher risk of amputation.

The SIGN stratification system was validated in a prospective cohort of 3,526 participants in a community setting [21]. It is based on eight easy to use and inexpensive measurements and has great value in detecting patients who will not develop a foot ulcer. The ADA stratification system was never validated for foot ulcer development, only for amputations [19]. However, the variables included are the same as those in the IWGDF system.

The system derived by Boyko et al. [24], along with the UTRFS system [18], were created through multivariate regression modelling instead of literature review and/or consensus. The Boyko group, along with the SIGN group, sought to include only variables that are easy to collect and commonly available in daily clinical care.

The system of Boyko et al. was also the only one assigning a specific score to the presence of each variable associated with foot ulcer development, which allows an impact evaluation of each variable (vs a group of variables) on overall risk. Additionally, no other system has been externally validated using the same variables as the original study, or reported their results in terms of AUC, which is considered the best way to determine a model's discriminatory ability [28]. On the other hand, score calculation is somewhat complicated without the use of data processing (e.g. a personal digital assistant or personal computer), which may make implementation more difficult in daily practice.

The Boyko et al. study [24] was also the only one to include the time factor in their analysis, assessing the stratification system at 1 and 5 years. A limitation of the Boyko et al. study is that participants were mainly men. However, in the subsequent study by Monteiro-Soares et al. [25], the system had no statistical differences by sex in subgroup analysis of foot ulcer risk prediction. Limitations of this later study are its retrospective design and patient recruitment from a high-risk setting. Nevertheless, the Boyko et al. model was equally valid in both distinct contexts.

Comparison of the IWGDF stratification system [19] with that proposed by the SIGN group [21] shows that the latter presents a significantly higher positive likelihood ratio for prediction of foot ulcer development in the high- and moderate-risk groups, and significantly higher accuracy in the high-risk group. In the other diagnostic accuracy measures, no statistically significant difference occurred (Table 4). Comparing the Boyko et al. system [24], validated by Monteiro-Soares et al. [25], with that proposed by the IWGDF [19] revealed no statistical differences. However, comparison of it with the SIGN system [21] revealed that several measures are significantly inferior. Nevertheless, it should be noted that there was a difference in the number of groups and also that these results were retrieved from three different populations with varied foot ulcer prevalence (from 5% [21] to 34% [20]), context and participants characteristics.

These differences in foot ulcer prevalence and/or incidence across studies should be kept in mind when assessing the comparative value of different foot complication prediction systems. Prediction rules developed in persons at high risk such as those under the care of a foot specialist might have less value in lower risk patients receiving care in primary care or diabetes clinic settings.

Our systematic review has a number of strengths and weaknesses. One of the latter is that quality assessment, data analysis and extraction were performed by one reviewer only (M. Monteiro-Soares). Additionally, this reviewer was not blinded to authors or institutions for this phase of the review. Strengths include review of articles for fulfilment of inclusion criteria by two reviewers blinded to

identity of authors and institutions, with a third serving as tie breaker.

Although a serious problem for diabetes patients and their healthcare providers, the best method for assessment of risk stratification is not immediately apparent and comparatively little research has been performed on this topic compared with other serious micro- and macro-vascular complications of diabetes. The question of which system one should choose to apply to one's specific setting cannot, we believe, be answered clearly at present. This deficiency could be remedied with further testing of existing risk classification systems, with a view to assessing predictive ability overall and in well-defined patient subgroups. In addition, further expansion of such systems would be justified using other easy-to-measure characteristics that have been overlooked in existing research on this subject. Such research will require multi-centre collaboration based on a common protocol, much as is the case with research now being conducted on prediction of cardiovascular disease outcomes [29].

Acknowledgements VA Puget Sound provided support for E. J. Boyko's involvement in this research.

Duality of interest The authors declare that there is no duality of interest associated with this manuscript.

References

1. Fard AS, Esmaelzadeh M, Larijani B (2007) Assessment and treatment of diabetic foot ulcer. *Int J Clin Pract* 61:1931–1938
2. Khanolkar MP, Bain SC, Stephens JW (2008) The diabetic foot. *Q J Med* 101:685–695
3. Wild S, Roglic G, Green A, Sicree R, King H (2004) Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diab Care* 27:1047–1053
4. Frykberg RG, Zgnois T, Armstrong DG et al (2006) Diabetic foot disorders: a clinical practice guideline (2006 revision). *J Foot Ankle Surg* 45(Suppl 5):S1–S66
5. Leese G, Schofield C, McMurray B et al (2007) Scottish foot ulcer risk score predicts foot ulcer healing in a regional specialist foot clinic. *Diab Care* 30:2064–2069
6. [No authors listed] (2001) Scottish Intercollegiate Guideline Network (SIGN) Guideline 55: The management of Diabetes. Edinburgh. SIGN, Royal College of Physicians Edinburgh. Available from www.sign.ac.uk/guidelines/index.html, accessed 10 October 2008
7. Boulton AJ (2004) The diabetic foot: from art to science. The 18th Camillo Golgi Lecture. *Diabetologia* 47:1343–1353
8. Larsson J, Agardh CD, Apelqvist J, Stenström A (1998) Long-term prognosis after healed amputation in patients with diabetes. *Clin Orthop Relat Res* 350:149–158
9. Apelqvist J, Bakker K, van Houtum WH, for the International Working Group on Diabetic Foot et al (2000) International consensus and practical guidelines on the management and the prevention of the diabetic foot. *Diabetes Metab Res Rev* 16(Suppl 1):S84–S92

10. Boulton AJ, Meneses P, Ennis W (1999) Diabetic foot ulcers: a framework for prevention and care. *Wound Repair Regen* 7:7–16
11. Boulton AJ, Armstrong DG, Albert SF et al (2008) Comprehensive foot examination and risk assessment—a report of the Task Force of the Foot Care Interest Group of the American Diabetes Association with endorsement by the American Diabetes Association of Clinical Endocrinologists. *Diab Care* 31:1679–1685
12. Boulton AJ, Armstrong DG, Albert SF et al (2008) Comprehensive foot examination and risk assessment. *Endocr Pract* 14:576–583
13. Helfand AE, Hirt PR (1994) Caring for the diabetic: assessing the risk in the diabetic foot. *N J Med* 91:256–258
14. Apelqvist J, Bakker K, van Houtum WH, Schaper NC, International Working Group on the Diabetic Foot (IWGDF) Editorial Board (2008) Practical guidelines on the management and prevention of the diabetic foot—based upon the International Consensus on the Diabetic Foot (2007). *Diabetes Metab Res Rev* 24(Suppl 1):S18–S187
15. Bower VM, Hobbs M (2009) Validation of the basic foot screening checklist. A population screening tool for identifying foot ulcer risk in people with diabetes mellitus. *J Am Podiatr Med Assoc* 99:339–347
16. Vandenbrouke JP, von Elm E, Altman DG, STROBE initiative et al (2007) Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Ann Intern Med* 147:W163–W194
17. Bossuyt PM, Reistma JB, Bruns DE et al (2003) The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Ann Intern Med* 138:W1–W12
18. Lavery LA, Armstrong DG, Vela SA, Quebedeaux TL, Fleischli JG (1998) Practical criteria for screening patients at high risk for diabetic foot ulceration. *Arch Intern Med* 158:157–162
19. Peters EJ, Lavery LA, International Working Group on the Diabetic Foot (2001) Effectiveness of the diabetic foot risk classification system of the International Working Group on the Diabetic Foot. *Diab Care* 24:1442–1447
20. Lavery LA, Peters EJ, William JR, Murdoch DP, Hudson A, and International Working Group on the Diabetic Foot (2008) Reevaluating the way we classify the diabetic foot: restructuring the diabetic foot risk classification system of the International Working Group on the Diabetic Foot. *Diab Care* 31:154–156
21. Leese GP, Reed F, Green V et al (2006) Stratification of foot ulcer risk in patients with diabetes: a population based study. *Int J Clin Pract* 60:541–545
22. Mayfield JA, Reiber GE, Sanders LJ, Janisse D, Pagach LM (2003) Preventive foot care in people with diabetes—American Diabetes Association. *Diab Care* 26(Suppl 1):S78–S79
23. Mayfield JA, Reiber GE, Sanders LJ, Janisse D, Pagach LM, American Diabetes Association (2004) Preventive foot care in diabetes: American Diabetes Association. *Diab Care* 27(Suppl 1):S63–S64
24. Boyko EJ, Ahroni JH, Cohen V et al (2006) Prediction of diabetic foot ulcer occurrence using commonly available clinical information: the Seattle Diabetic Foot Study. *Diab Care* 29:1202–1207
25. Monteiro-Soares M, Dinis-Ribeiro M (2010) External validation and optimisation of a model for predicting foot ulcers in patients with diabetes. *Diabetologia* 53:1525–1533
26. Armstrong DG, Hussain SK, Middleton J et al (1998) Vibration perception threshold: are multiple sites of testing superior to single site testing on diabetic foot examination? *Ostomy Wound Manage* 44:70–74
27. Stiell IG, Wells GA (1999) Methodological standards for development of clinical decision rules in emergency medicine. *Ann Emerg Med* 33:437–447
28. Reynolds T (2001) Disease prediction models aim to guide medical decision making. *Ann Intern Med* 135:637–640
29. Bhatt DL, Eagle KA, Ohman M et al (2010) Comparative determinants of 4-year cardiovascular event rates in stable outpatients at risk of or with atherothrombosis. *JAMA* 304:E1–E8