Appendix 1 - Full rationale behind the formulated recommendations

This appendix belongs to the publication "Guidelines on the classification of foot ulcers in people with diabetes (IWGDF 2023)" by Monteiro-Soares and colleagues and uses the same references.

From the 28 systems we have retrieved with our systematic review and taking into account all the elements of the Summary of Judgments (see Methods section and Table 1), the group considered that only six had the potential to be recommended to be used in clinical practice. In this Appendix we will explain the reasons that supported this decision.

Conditionally or strongly recommended scores or classifications

DIAFORA

The Diabetic Foot Risk Assessment (DIAFORA) score includes four foot-related variables, used to predict ulcer onset, and four foot ulcer-related variables that are used to predict amputation in people with diabetes. This classification was derived in a study from Portugal, with a low risk of bias, providing accuracy measures for overall and major amputation (17). Using the same cohort, the authors concluded that this classification had similar accuracy in comparison to the 10 others used for amputation (42) and mortality (43). However, for the latter, no statistically significant association was found. In addition, this classification was considered to be probably acceptable and feasible. Due to the low risk of bias, the direct comparison of its accuracy with other widely used systems for more than one clinical outcome and its ease of use we have decided to recommend the use of this system even though more substantial evidence is needed.

IDSA/IWGDF

The Infectious Diseases Society of America (IDSA) and International Working Group of the Diabetic Foot classification, was created by expert consensus, to assess the severity of infected DFUs. Twelve studies (26, 42-52) have evaluated the ability of this classification to predict healing, amputation, risk of hospitalisation, risk of readmission, length of stay and mortality. The results were somewhat inconsistent, but sufficient for the group to consider this classification to be accurate and with moderate potential desirable effects. Also, one study has reported moderate agreement measures for this classification (53).

Specific equipment and blood sampling may be required for the assessment of patients at higher grades and this could reduce equity. However, this will correspond to a small proportion of people at the more severe end of the spectrum of the condition, i.e. those at higher risk of being hospitalised; a setting

where such equipment is usually readily available. So, overall we considered this classification to be probably acceptable and feasible.

SINBAD

The SINBAD score was a modification of the S(AD)SAD classification, including the same variables plus site, dichotomising each characteristic into absent or present features and creating an easy-to-use scoring system that can achieve a maximum of 6 points (15). Thus, it may be considered too simplistic for some specific settings and to guide clinical management.

This score (overall or each variable individually) was associated with several clinical outcomes (namely, healing, healing time, LEA, hospital admissions, being alive and ulcer-free, and costs) in cohorts from several countries (China, Germany, Pakistan, Portugal, Republic of Korea, Spain, Syria, Tanzania, United Kingdom) (15, 31, 38, 42, 48, 54-56).

Four studies, mainly with a low risk of bias, demonstrated an association between this score and amputation but with variable accuracy (AUC 0.52 to 0.88). In the only study having mortality as an outcome, no statistically significant association was observed for any of the assessed classifications (43).

This score allows a very simple description of a foot ulcer in people with diabetes, which is one of the main reasons why it was selected for the National Diabetes Foot Care Audit (NFDA) by the National Health Service (NHS) (38) in the United Kingdom. The national audit showed an 18% increase in the number of ulcers submitted into the registry when comparing the 2015-16 period to the 2016-17 period, which may indicate that, with an increase in the dissemination and in the level of experience due to the use in clinical practice, this system can be easily adopted by healthcare professionals. Furthermore, in 2021, the same authors reported that since data collection started in 2014, the NDFA has consistently found that faster referral to the specialist foot care service is associated with fewer severe ulcers and better 12 weeks outcomes.

In the 2019 audit associations of lower SINBAD score (of 0-2, rather than 3 or above) with being alive and ulcer free at 12 and 24 weeks, as well as with fewer hospital admissions and revascularizations were observed. Furthermore, two studies have confirmed that this score is reliable (57, 58).

The available evidence, besides supporting the reliability and the accuracy of this score to predict several clinical outcomes, has shown that it can be easily adopted by a large number of professionals in several countries and that the adoption of this classification in daily foot care can improve referral and clinical outcomes. Thus, we consider it to be feasible and probably acceptable to be applicable in all clinical settings and to have expected moderate positive effects.

Texas University Classification

Developed at the University of Texas Health Science Centre in San Antonio this system classifies diabetes-related foot ulcers using a bi-dimensional 4 × 4 matrix according to depth and presence of infection, ischaemia, or both, and is simple to apply. Infection is defined as frank purulence and/or two or more of the following local signs: warmth, erythema, lymphangitis, lymphadenopathy, oedema, pain, and loss of function. Ischaemia is diagnosed using a combination of clinical signs and symptoms (claudication, rest pain, absent pulses, atrophic integument [skin], absence of pedal hair, dependent rubor, or pallor on elevation) plus one or more non-invasive criteria [transcutaneous oxygen measurements of <40 mmHg, ankle brachial index (ABI) of <0.80 or toe systolic pressure of <45 mmHg]. Loss of protective sensation and ulcer size (area) is not included in this classification.

With one internal validation study (19) and 30 external validation studies from many different healthcare systems, this score showed an association with several clinical outcomes: healing, time to healing, lower extremity amputation (total, major and minor), hospital admissions, mortality, and costs. Most of the studies were considered at high risk of bias, however. Reliability was considered moderate in the four studies which assessed such (53, 57, 58).

Although widely used and validated in many healthcare systems it was considered that the use of more complex diagnostic tests for the assessment of ischaemia (ankle or toe pressure, or transcutaneous oxygen measurement) decreased equity and increased costs. However, the group considered that there was adequate evidence to prove the accuracy, reliability and applicability of this classification.

(Meggitt-)Wagner

This classification has six grades but is based largely on wound depth and tissue viability. Other variables (such as LOPS) are not considered, and infected and/or ischaemic foot ulcers cannot be adequately differentiated by this classification system.

This classification was the most commonly evaluated, with 74 articles having assessed the association between this classification and several of the foot ulcer-related outcomes or its predictive accuracy (24), presenting much more evidence, when compared to all of the other classifications identified.

Although most studies were considered to be at high risk of bias and some inconsistency was observed, the group considered that there was adequate evidence to state that this classification may have a valid role in predicting healing at a group level, amputation, in-hospital mortality, health-related quality of life and cost in people with diabetes. Also, moderate agreement measures were reported, which is comparable to others (53, 57, 59, 60).

This is a very simple classification to use, and the number of articles validating it indicates that it is feasible and widely accepted by healthcare professionals in the field.

WIfI

The purpose behind the creation of this system was to provide a more precise description of limb-related disease burden and thereby enable more accurate assessment of outcomes between patients with similar characteristics and thereby serve as a guide for the selection of therapies. The members of the Society of Vascular Surgery Lower Extremity Guidelines Committee used a Delphi technique to create strata of possible combinations of outcome predictors leading to four clinical limb stages corresponding, respectively, to a very low, low, moderate, or high risk of any individual requiring an lower extremity amputation within one year and very low, low, moderate, or high likelihood of benefiting from or requiring revascularization (assuming infection can be controlled first) (18).

Infection is graded using the IDSA/IWGDF criteria. The wound area is not considered quantitatively. Ischaemia characterization requires moderate expertise and equipment, which, depending on the clinical setting worldwide, may not always be available. On the other hand, one of the main purposes of this classification was to increase the detail of perfusion status and allow more objective decision-making regarding the need for revascularization.

This score is somewhat complex and the existence of an app may facilitate the use of this score, but requires access to a smart phone.

This classification has been shown to predict multiple pertinent diabetes-related foot ulcer outcomes, including the extent of healing, time to heal, amputation occurrence, amputation-free survival, need for revascularization, maintenance of ambulatory and independent living status, costs, and mortality (24). Its use has since been endorsed by many centres and societies worldwide.

Of the 14 articles found (4), five of them were conducted in the same setting and with a substantial overlap of samples. The group considered that this score is associated with healing and amputation and that it can help in clinical management, but that there is still insufficient evidence for hospital admissions, costs, and quality of life.

Not recommended classifications or scores

In this section, we provide the rationale for considering the classifications below as not suitable to be recommended for clinical care at the present time.

The Bates-Jensen wound assessment tool (BWAT) was developed to assess the progress of healing in pressure ulcers (61). It was validated only once in the diabetic foot context for healing at 4 weeks in a study considered to be at high risk of bias (62). It includes 13 items that are somewhat subjective and that require an accurate assessment of the dimensions of the ulcer, undermining and amounts (to calculate proportions) of specific types of tissue. Although probably feasible, this classification was considered to be unlikely to be accepted by clinicians specialised in the diabetic foot.

The Diabetic Foot Ulcer Assessment Scale (DFUAS) score was developed to assess the status of a foot ulcerin people with diabetes over time in order to evaluate the effectiveness of interventions (62), but

includes several of the variables used in systems developed to assess the progress of healing in pressure ulcers. The DMIST score is a simplification of the DFUAS score, diminishing the number of variables to be assessed from 11 to seven. This score was also validated once in people with diabetes and a foot ulcer, by a study at high risk of bias (63).

The Pressure Ulcer Scale for Healing (PUSH) score was also created for monitoring the healing of pressure ulcers (64). This score includes only three variables (area, exudate and tissue type) and was considered to be too simplistic. Two studies, with a high risk of bias and partial overlap of samples, provided high areas under the ROC curve to predict healing at 4 weeks (62, 63). No further evidence was found.

Similarly, the DESIGN was also developed for pressure ulcers, includes seven similar items, and was validated in the same study as the DMIST and PUSH scores (63).

The score developed by Chetpet and colleagues, in 2018 (44), includes 13 items with scores that can vary from 0 to 2-11 per item. This study, at high risk of bias, provided accuracy in the prediction of amputation at 12 months in people with diabetes and an infected foot ulcer. Besides the lack of evidence on accuracy and impact on clinical management, we have also considered it to be unacceptable for healthcare professionals and unfeasible.

The Curative Health Services (CHS) system is a descriptive wound classification system, created by Margolis et al., that has been included in several models derived by the team using the same database throughout the years (65-67). None of the versions of the models that included this system were externally validated, nor were reliability measures ever provided. When compared to 10 other classifications, the count model (the easiest version to be applied in the clinical setting) showed similar accuracy for amputation (42).

Similarly, in the Eurodiale cohort study, van Battum and colleagues (68) and Pickwell and colleagues (46) have proposed three different models: one for overall amputation, one for minor amputation and another for amputation excluding lesser toes. None of these models was externally validated. In addition, the group considered that having different models to predict one specific clinical outcome would be unfeasible.

The Clinical Signs and Symptoms checklist (CSSC) includes 12 clinical signs and symptoms of infection and was developed, in 2001, to assess infected chronic wounds (69). This checklist does not provide a score or ways of stratifying people. Only one article was found reporting on the reliability of CSSC, with kappa values highly variable for each item. In addition to the lack of evidence, we considered that the inconsistency of some of the findings would make this system unacceptable for the management of diabetic foot ulcers.

The DEPA score is considered to be fairly simple, but the lack of detailed definitions of some components may undermine its reliable application. One study assessed the use of this score to predict healing (70); three studies showed that the accuracy to predict amputation of this classification is comparable with others (42, 56, 71); although it does not predict mortality (43). There is no information

for other critical outcomes, nor about the reliability of this score. Due to the lack of sufficient evidence, oversimplicity and subjectivity in the appraisal of some of the variables included in the score, we could not currently recommend its use.

The Diabetic Ulcer Severity Score (DUSS) classification, in the same way as the DEPA score, could be considered too simplistic but does include more objective variables. Six studies (from China, India, Portugal and South Korea) reported an inconsistent association between this score and several of the foot ulcer-related clinical outcomes (42, 43, 56, 72-74). For these reasons we could not currently recommend its use.

The Diabetic Foot Infection (DFI) score includes 10 items that can be scored between 0 and 3 up to 10, focusing on signs and symptoms of inflammation and wound measurements. We found only one study (75), at low risk of bias, evaluating the association between this score and healing, amputation and death. However, when adjusting for potential confounders only the association between the score and healing remained statistically significant.

In 2016, a wound assessment algorithm, for the 10 most common types of wound, embedded in a mobile application to help guide management, was created by Jun and colleagues (76). We found two studies (71, 77), with a low risk of bias, validating the DIRECT score in South Korea. Both studies presented a comparable accuracy. In addition, some of the included items were not clearly defined and required a smartphone (app format), which may not be feasible or acceptable in all parts of the world. Its advantage however, may be that it can help guide management for each case and evaluate the progression of care more objectively. In 2022, Lee et al. proposed a refinement of this score by adding C-reactive protein and the presence of hypertension. This would require collecting blood samples from all patients with a foot ulcer and rapid access to the biomarker result.

Similarly, to calculate the LRINEC score it is necessary to determine a blood sample for six different biomarkers (C-reactive protein, white blood cells, haemoglobin, sodium, creatinine and glucose) and have rapid access to the biomarker(s) result(s). However, these analyses may not be possible in all contexts and would involve additional costs of care. Furthermore, this score does not include any variable that describes the actual foot ulcer, as it was originally developed to ascertain the risk of necrotising infection/ necrotising fasciitis. After being developed in 2004, by Wong et al (78), this classification was validated in a diabetic foot population only once in 2021 (79). In this single external validation study, with a high risk of bias, Sen and colleagues (79) reported an association between the score and amputation and in-hospital mortality, although with low accuracy. Overall, it was decided that the use of this classification could not be recommended.

The PEDIS classification was created for the selection of participants for clinical research. Having this specific purpose in mind, detailed and complex definitions of all components are provided which makes it difficult to use in all contexts. Because it was designed as an aid for prospective research, it does not provide a defined outcome against which to assess ulcer types.

This classification was later transformed into a score (55), but its accuracy was never directly compared with other classifications. It has also shown an inconsistent association with healing and amputation. Two studies, with a high risk of bias, showed an association of the score with mortality (80, 81) and two others, at low risk of bias reported that this classification presented moderate agreement (57, 58).

The S(AD)SAD classification was developed mainly to improve prospective clinical research on diabetic foot management, in other words, to allow clinical audit. It uses an acronym to facilitate the memorization of this classification, but each of the composing variables (area, depth, sepsis, arteriopathy and denervation) has four possible categories. The components are the same as those used in the PEDIS system, although each component is graded on a four-point scale according to severity. In 2004, Treece et al proposed the transformation of this descriptive classification into a score, according to the grade of each composing variable (82). To detect and define ischaemia, the authors proposed the use of palpation of foot pulses and the presence of gangrene, which can be criticized. In the original description, the assessment of LOPS was performed using the Neurotip, which may not translate well into current clinical practice as the most used tool is the 10-g monofilament and/or tuning fork. Moreover, Charcot foot is included in the most severe grade of LOPS, but the aetiology of this deformity is complex with likely multiple causes.

Six studies assessed the association between each composing variable with time to healing, healing or amputation(24). No reliability assessment was conducted. A clear variation occurred between the studies in terms of the association presented between the outcome and individual components or the full classification. Furthermore, the studies from Treece et al (82), Chipchase et al (83) and Ince et al (84) were conducted in the same setting with some overlap between the included samples. This classification includes several variables that describe the ulcer in some detail, although there are some controversies around the method of collection of some of the variables. Furthermore, it does not seem to allow for predicting clinical outcomes, guide ulcer management or adjust for variations in the efficacy of care due to differences in the population. For all these reasons, we could not recommend the use of this classification.

The Saint Elian Wound Score System (SEWSS) has also evolved from the PEDIS classification, by adding five more variables. Ischaemia is diagnosed using foot pulse palpation, ABI, or toe-brachial index (TBI). Neuropathy is identified using the 10-g monofilament or a 128-Hz tuning fork. This score was considered to be too complex and not feasible to use in daily clinical practice.

Besides the derivation/ internal validation study (85), this classification was assessed by four other studies. One (86), with a high risk of bias, reported an association between the score and chances of healing and median time to heal. Another study (87), with high risk of bias, found an association between the score and the risk of having a major amputation ($p<0.01$ and AUC of 0.89) but not with the chance of having a completed wound closure.While another study, with a low risk of bias, reported one of the lowest AUC values for this classification when compared to 10 others (42) along with the

Scottish Intercollegiate Guideline Network (SIGN)/SCI-DC classification for the prediction of amputation.

The length of stay in hospital was higher for those with a higher score in one study (88), with high risk of bias, but such association was lost when adjustment for other variables were made.

The SEWSS classification has been endorsed by the International Diabetes Federation (in 2017) and by Public Health Services of countries within Latin America.

The SIGN/SCI-DC classification was created to predict the risk of developing an ulcer in people with diabetes and so does not include any DFU-related variable. In 2007, Leese et al. also tested its association with healing and amputation (89). This classification provided inconsistent results and was considered by the group to have potentially small desirable effects and no impact on equity.

The Tardivo score combines the (Meggitt-)Wagner system with the perfusion component of PEDIS and also adds the location of the ulcer, providing a final score and stratifying patients into four grades of wound severity. In the internal validation study (90), a cut-off of 11 was proposed to best predict the risk of amputation, but in the external validation study (91), it was suggested to use a score of 16. No other clinical outcome was assessed, nor reliability measures reported.

The van Acker/Peter classification system, proposed in 2002, is a modified version of the University of Texas system, using also a two-dimensional matrix that grades depth on the vertical axis and foot characteristics on the horizontal axis. Loss of protective sensation is defined by abnormal monofilament and/or vibration perception, abnormal deep tendon reflexes, or abnormal electromyography. Ischaemia is defined by any of the following criteria being present: systolic ankle blood pressure less than 50 mmHg, ABPI less than 0.9, or transcutaneous oximetry (TcPO2) less than 50 mmHg. The extent of infection is classified as 1: extremely superficial ulcer without important signs of infection, 2: small ulcer with cellulitis without the involvement of tendons and bone, 3: more severe infected ulcer with involvement of tendons and/or bone with/without abscess, 4: periostitis, involvement of the bone without signs of destructive osteomyelitis; typical, bone contact without visible defects on radiography, and 5: overt radiographic destructive osteomyelitis. Charcot foot was included, even though it is considered a different clinical entity. Only one study analysed the association between this classification with healing (92), amputation (42) and death (43), and none assessed its reliability. Due to the paucity of evidence, the use of a two-dimensional matrix, and some issues in the collection of some of the composing variables the use of this classification is not recommended.

The Wound Healing Index (WHI) was derived and internally validated in an article by Fife and colleagues, using the United States wound registry data (93). Several models are proposed in the same article. The one with baseline characteristics is composed of 10 variables, but no instructions on how to use or calculate the score are provided. We considered that we did not have enough evidence relating to the use of this index and that in addition it would not be possible to apply it in all clinical contexts, due to the dependence on a high number of variables and need to include data from an electronic health registry.

The Xie et al. model was developed in 2021 (40), using machine learning techniques, to predict in-hospital lower extremity amputation using commonly available information in electronic health records. The authors were able to create an explainable model with very good sensitivity to predict that clinical outcome.

This model requires access to an electronic health support system and the generalised use of personal computers by all the healthcare professionals responsible for foot ulcer care in people with diabetes. Additionally, this model requires information about more than 35 variables. Due to the complexity of this model, the required equipment and infrastructure, and the lack of evidence about other clinical outcomes, other settings and reliability; the use of this classification is not recommended.

Appendix 2 – Summary of judgments and voting process

This appendix belongs to the publication “Guidelines on the classification of foot ulcers in people with diabetes (IWGDF 2023)” by Monteiro-Soares and colleagues. It contains the detailed information of all summaries of judgments for each of the systems, organized by alphabetical order, evaluated for the guidelines’ development as well as the the voting process behind the selection of conditionally or strongly recommended systems.

302 Supplementary Table 1: Summary of Judgments for BWAT (Bates Jensen wound assessment tool) system

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

303 Final voting results: 100% voted as against, 100% voted as conditional

304

Supplementary Table 2: Summary of Judgments for Chetpet system

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
	Very complex and hard to remember the variable points and only for infected DFUs				
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

Final voting results: 100% voted as against, 100% voted as conditional

Supplementary Table 3: Summary of Judgments for CHS (Curative Health Services) score

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

Final voting results: 100% voted as against, 83% voted as conditional

311 Supplementary Table 4: Summary of Judgments for CSSC (Clinical Signs and Symptoms checklist)

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	No answer
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

312 Final voting results: 100% voted as against, 100% voted as strong

313 Supplementary Table 5: Summary of Judgments for DEPA (Depth, extent of bacterial colonisation, phase of healing, and associated aetiology)
314 score

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate		Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

315 Final voting results: 100% voted as against, 100% voted as conditional

316 Supplementary Table 6: Summary of Judgments for DESIGN score

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

317 Final voting results: 100% voted as against, 100% voted as conditional

318

Supplementary Table 7: Summary of Judgments for DFI (Diabetic foot infection) score

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

Final voting results: 100% voted as against, 100% voted as conditional

322 Supplementary Table 8: Summary of Judgments for DFUAS (Diabetic foot ulcer assessment scale)

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

323 Final voting results: 100% voted as against, 100% voted as strong

324

Supplementary Table 9: Summary of Judgments for DIAFORA (Diabetic foot risk assessment) tool

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
	No external validation studies and no reliability studies				
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

Final voting results: 100% voted as in favour, 100% voted as conditional, Matilde Monteiro-Soares abstained as author

328 Supplementary Table 10: Summary of Judgments for DIRECT system

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

329 Final voting results: 100% voted as against, 100% voted as strong

330

331 Supplementary Table 11: Summary of Judgments for DMIST system

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

332 Final voting results: 100% voted as against, 100% voted as strong

333

334 Supplementary Table 12: Summary of Judgments for DUSS (Diabetic Ulcer Severity Score) score

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

335 Final voting results: 100% voted as against, 100% voted as conditional

336

337 Supplementary Table 13: Summary of Judgments for Eurodiale scores

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

338 Final voting results: 100% voted as against, 83% voted as conditional

339

Supplementary Table 14: Summary of Judgments for IDSA/IWGDF(Infectious Diseases Society of America/ International Working Group of the Diabetic Foot) system

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

Final voting results: 100% voted as in favour, 100% voted as conditional

343 Supplementary Table 15: Summary of Judgments for LRINEC (Laboratory Risk Indicator for Necrotising Fasciitis) score

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

344 Final voting results: 100% voted as against, 100% voted as conditional

345

346 Supplementary Table 16: Summary of Judgments for PEDIS system

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

347 Final voting results: 100% voted as against, 100% voted as conditional

348

Supplementary Table 17: Summary of Judgments for PUSH (Pressure Ulcer Scale for Healing) score

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

Final voting results: 100% voted as against, 67% voted as conditional

Supplementary Table 18: Summary of Judgments for S(AD)SAD system

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

Final voting results: 100% voted as against, 100% voted as conditional, Fran Game and William Jeffcoate abstained as co-authors

355 Supplementary Table 19: Summary of Judgments for SEWSS (Saint Elian Wound Score System) score

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

356 Final voting results: 60% voted as neutral and 40% as against, 100% voted as conditional

357

Supplementary Table 20: Summary of Judgments for SIGN (Scottish Intercollegiate Guideline Network)/SCI-DC system

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

Final voting results: 100% voted as against, 100% voted as strong

361 Supplementary Table 21: Summary of Judgments for SINBAD (Site, ischaemia, neuropathy, bacterial infection, area, and depth) score

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

362 Final voting results: 100% voted as in favour, 100% voted as strong, Fran Game and William Jeffcoate abstained as co-authors

363

Supplementary Table 22: Summary of Judgments for Tardivo score

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

Final voting results: 100% voted as against, 100% voted as strong

Supplementary Table 23: Summary of Judgments for UTCS (Texas University classification)

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

Final voting results: 100% voted as in favour, 100% voted as conditional

370 Supplementary Table 24: Summary of Judgments for van Acker/Peter classification

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

371 Final voting results: 100% voted as against, 100% voted as conditional

372

373 Supplementary Table 25: Summary of Judgments for (Meggitt-)Wagner classification

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

374 Final voting results: 100% voted as in favour, 100% voted as conditional

375

376 Supplementary Table 26: Summary of Judgments for WHI (Wound Healing Index) score

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

377 Final voting results: 100% voted as against, 100% voted as strong

378

379 Supplementary Table 27: Summary of Judgments for WIfI (Wound, Infection, foot Ischemia) score

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	
TYPE OF RECOMMENDATION	Strong Against Intervention	Conditional Against Intervention	Conditional for either intervention or comparison	Conditional for the intervention	Strong for the intervention

380 Final voting results: 100% voted as in favour, 100% voted as conditional, Joseph Mills abstained as author

381

382 Supplementary Table 28: Summary of Judgments for Xie model

PROBLEM priority	No	Probably no	Probably Yes	Yes	
TEST ACCURACY	Very inaccurate	Inaccurate	Accurate	Very Accurate	
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large	
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial	
CERTAINTY OF EVIDENCE OF TEST ACCURACY	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF TEST'S EFFECTS (direct benefits, adverse effects or burden)	Very low	Low	Moderate	High	
CERTAINTY OF EVIDENCE OF MANAGEMENT'S EFFECT (Guided by test result)	Very low	Low	Moderate	High	
CERTAINTY OF THE EVIDENCE of link between TEST RESULT/MANAGEMENT	Very low	Low	Moderate	High	
CERTAINTY OF EFFECTS	Very low	Low	Moderate	High	
How much people VALUE the main outcome	Important uncertainty or variability	Possibly important uncertainty or variability	Probable no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	Moderate	High	
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased
ACCEPTABILITY (to stakeholders)	No	Probably no	Probably Yes	Yes	
FEASIBILITY	No	Probably no	Probably Yes	Yes	

383 Final voting results: 100% voted as against, 100% voted as conditional

384

Appendix 3 – Voting process for each clinical scenario

This appendix belongs to the publication "Guidelines on the classification of foot ulcers in people with diabetes (IWGDF 2023)" by Monteiro-Soares and colleagues. In this Appendix we will describe the voting process behind the selection of one or two systems to be used in each clinical scenario and the strength of the recommendation (Strong versus Conditional)

To aid communication between healthcare professionals

First line: SINBAD
Final voting results: 100% voted as in favour, 100% voted as strong, Fran Game and William Jeffcoate abstained as co-authors

Second line: WIfI

Final voting results: 100% voted as in favour, 100% voted as conditional, Joseph Mills abstained as co-author

For clinical prediction of the outcome of an individual ulcer

Final voting results: 100% voted as against the use of any system

For characterising a person with the assessment of an infected ulcer

First line: IDSA/IWGDF
Final voting results: 100% voted as in favour, 83% voted as strong

Second line: WIfI
Final voting results: 100% voted as in favour, 100% voted as conditional, Joseph Mills abstained as co-author

For characterising a person with peripheral artery disease

No voting done, it was adopted the decision made by Peripheral Artery Working Group (see (11))

For the audit of outcome(s) of populations

First line: SINBAD
Final voting results: 100% voted as in favour, 100% voted as strong, Fran Game and William Jeffcoate abstained as co-authors

Second line:
Final voting results: 100% voted as against the use of any other system