

EPUAP classification system for pressure ulcers: European reliability study

Dimitri Beeckman, Lisette Schoonhoven, Jacqui Fletcher, Kátia Furtado, Lena Gunningberg, Hilde Heyman, Christina Lindholm, Louis Paquay, José Verdú & Tom Defloor

Accepted for publication 14 August 2007

Correspondence to D. Beeckman:
e-mail: dimitri.beeckman@ugent.be

Dimitri Beeckman MA RN
PhD Candidate
Nursing Science, Department of Public Health, Faculty of Medicine and Health Sciences, Ghent University and Research Staff, Nursing Department, University College Arteveldehogeschool, Belgium

Lisette Schoonhoven PhD RN
Assistant Professor
Nursing Science, Centre for Quality of Care Research (WOK), Radboud University Nijmegen Medical Centre, The Netherlands

Jacqui Fletcher BSc RN
Principal Lecturer
School of Nursing and Midwifery, University of Hertfordshire, Hatfield, UK

Kátia Furtado MA RN
Community Nursing Specialist
Centro de Saúde de Arronches, Portalegre, Portugal

Lena Gunningberg PhD RN
Leader of Nursing Research and Development
Surgery Division, Uppsala University Hospital and Assistant Professor, Department of Surgical Science, Uppsala University, Sweden

Hilde Heyman RN
Nursing Director
Nursing Home Sint Bartholomeus, Antwerp, Belgium

BEECKMAN D., SCHOONHOVEN L., FLETCHER J., FURTADO K., GUNNINGBERG L., HEYMAN H., LINDHOLM C., PAQUAY L., VERDÚ J. & DEFLOOR T. (2007) EPUAP classification system for pressure ulcers: European reliability study. *Journal of Advanced Nursing* 60(6), 682–691
doi: 10.1111/j.1365-2648.2007.04474.x

Abstract

Title. EPUAP classification system for pressure ulcers: European reliability study

Aim. This paper is a report of a study of the inter-observer reliability of the European Pressure Ulcer Advisory Panel pressure ulcer classification system and of the differential diagnosis between moisture lesions and pressure ulcers.

Background. Pressure ulcer classification is a valuable tool to provide a common description of ulcer severity for the purposes of clinical practice, audit and research. Despite everyday use of the European Pressure Ulcer Advisory Panel system, its reliability has been evaluated in only a limited number of studies.

Methods. A survey was carried out between September 2005 and February 2006 with a convenience sample of 1452 nurses from five European countries. Respondents classified 20 validated photographs as normal skin, blanchable erythema, pressure ulcers (four grades), moisture lesion or combined lesion. The nurses were familiar with the use of the European Pressure Ulcer Advisory Panel classification scale.

Results. Pressure ulcers were often classified erroneously ($\kappa = 0.33$) and only a minority of nurses reached a substantial level of agreement. Grade 3 lesions were regularly classified as grade 2. Non-blanchable erythema was frequently assessed incorrectly as blanchable erythema. Furthermore, the differential diagnosis between moisture lesions and pressure ulcers appeared to be complicated.

Conclusion. Inter-observer reliability of the European Pressure Ulcer Advisory Panel classification system was low. Evaluation thus needs to focus on both the clarity and complexity of the system. Definitions and unambiguous descriptions of pressure ulcer grades and the distinction between moisture lesions will probably enhance clarity. To simplify the current classification system, a reduction in the number of grades is suggested.

Keywords: classification scale, European Pressure Ulcer Advisory Panel, instrument validation, nursing, pressure ulcer, reliability

continued on page 683

Christina Lindholm PhD RN
Professor
Department of Clinical Nursing,
Kristianstad University, Kristianstad, Sweden

Louis Paquay MSc RN
Nursing Coordinator
Wit-Gele Kruis van Vlaanderen, Brussels,
Research Staff, Academic Centre for General
Practice, Kuleuven, Belgium

José Verdú PhD RN
Senior Lecturer
Nursing Science, Department of Community
Nursing, Preventive Medicine, Public Health
and History of Science, School of Nursing,
University of Alicante, Alicante, Spain

Tom Defloor PhD RN
Professor
Nursing Science, Department of Public
Health, Faculty of Medicine and Health
Sciences, Ghent University, Gent, Belgium

Introduction

A pressure ulcer is a localized area of tissue destruction occurring when soft tissue is compressed over bony prominences for prolonged periods of time. Tissue destruction occurs when the compressed tissue is deprived of oxygen (Wound, Ostomy, and Continence Nurses Society (WOCN) 2003). A pressure ulcer is caused primarily by unrelieved pressure, shearing, friction or a combination of these forces (European Pressure Ulcer Advisory Panel (EPUAP) (1999). The severity of pressure ulcers varies from erythema of intact skin to tissue destruction involving skin, subcutaneous fat, muscle and bone.

Since the first classification system for pressure ulcers, developed by Shea (1975), numerous systems have been developed to describe the different degrees of tissue damage, with varying numbers for grades ranging from a 0- to 5-grade classification to a 1- to 7-grade classification (Witkowski & Parish 1981, Dealey & Lindholm 2006). In the United States of America, the National Pressure Ulcer Advisory Panel (NPUAP) developed in 1989 a classification using four grades. This classification was adopted by the EPUAP in 1999 with some minor textual changes (Defloor & Schoonhoven 2004) (Table 1). The purpose of a classification system is to standardize record-keeping and provide a common description of ulcer severity for the purposes of clinical practice, audit and research (Nixon *et al.* 2005).

The major weakness of pressure ulcer classification systems is the lack of evidence to support their use, the most important factor being reliability (Sharp 2004, Dealey & Lindholm 2006). Reliability represents the variation of a classification system produced in repeated measurements. The less variation a classification system produces, the higher its reliability. Both inter-observer and intra-observer reliability can be measured. Inter-observer reliability reflects the degree to which two or more independent assessors assign an equal value during observation or measurement (Polit & Beck 2003). Intra-observer reliability measures the degree of reliability of a test score of a single assessor over time (Guggenmoos-Holzmann 1993).

Background

Despite everyday use by nurses, there are only a limited number of recent studies evaluating the inter-observer reliability of the EPUAP classification system (Bours *et al.* 1999, Russell & Reynolds 2001, Pedley 2004, Defloor & Schoonhoven 2004; Defloor *et al.* 2006). Intra-observer reliability is seldom studied (Defloor *et al.* 2006). Wide variability can be found in both simple percentage agreement and chance-corrected rater agreement (κ = Cohen's kappa).

In the study by Bours *et al.* (1999), pairs of nurses were asked to observe and classify the skin at the pressure points with 23 hospital patients and 45 nursing home patients using the EPUAP system. Inter-observer reliability was high

Table 1 Pressure ulcer classification (European Pressure Ulcer Advisory Panel (EPUAP). 1999)

Grade 1	Non-blanchable erythema of intact skin. Discolouration of the skin, warmth, oedema, induration or hardness may also be used as indicators, particularly in individuals with darker skin*.
Grade 2	Partial thickness skin loss involving epidermis, dermis or both. The ulcer is superficial and presents clinically as an abrasion or blister.
Grade 3	Full thickness skin loss involving damage to or necrosis of subcutaneous tissue that may extend down to, but not through, underlying fascia.
Grade 4	Extensive destruction, tissue necrosis or damage to muscle, bone, or supporting structures with or without full thickness skin loss.

*Whether the erythema can be blanched or not (by means of a finger or a transparent disk) is the most important distinction between a normal physiological reaction of the tissue to pressure and shearing forces, and grade 1 pressure ulcer.

($\kappa = 0.81-0.97$). However, the observations were not made independently. When observers worked independently of each other, the reliability was found to be much lower ($\kappa = 0.49$).

In a study by Russell and Reynolds (2001), the two-digit Stirling classification (percentage agreement = 30.2%) was found less reliable than the simpler EPUAP system (percentage agreement = 61.9%) when 12 pressure ulcer photographs were assessed by 200 nurses.

Defloor and Schoonhoven (2004) found a high degree of reliability of the EPUAP classification system when 56 photographs of pressure ulcers and moisture lesions were presented to 44 pressure ulcer experts ($\kappa = 0.80$). The authors concluded that it is likely that there would be less agreement amongst those with little experience.

Pedley (2004) measured inter-observer agreement of the Stirling scale (one- and two-digit version) and the EPUAP scale when 35 observations were made by two Registered Nurses with 30 patients in an elder care unit of an acute hospital. The levels of agreement obtained were poor (two-digit Stirling scale: $\kappa = 0.46$; one-digit Stirling scale: $\kappa = 0.37$; EPUAP scale: $\kappa = 0.31$).

Defloor *et al.* (2006) reported a study examining the inter- and intra-observer reliability of the EPUAP system and the ability to differentiate correctly between moisture lesions and pressure ulcers. Moisture lesions were defined as a result of prolonged exposure of the skin to excessive fluid because of urinary or faecal incontinence, profuse sweating or wound exudate (Maklebust & Siegreen 1995). Photographs were presented to 473 nurses. Both inter- ($\kappa = 0.37$) and intra-rater reliability ($\kappa = 0.38$) were low. If only pressure ulcer photographs were considered, the average kappa varied between 0.41 and 0.51. Ascertaining the differential diagnoses for pressure ulcers and other types of lesions appeared to be difficult.

In summary, the results of previous research are similar. Inter-observer reliability is low, but some variability can be found. In this study, inter-rater reliability of the EPUAP classification system was tested in clinical areas, with an additional focus on the differential diagnosis between pressure ulcers and moisture lesions.

The study

Aim

The aim of the study was to assess the inter-observer reliability of the EPUAP pressure ulcer classification system and of the differential diagnosis between moisture lesions and pressure ulcers.

Design

An instrument validation study was conducted, using a survey to study potential difficulties and indistinctness when classifying pressure ulcers and to differentiate between pressure ulcers and moisture lesions.

Participants

A survey was carried out between September 2005 and February 2006 with a convenience sample of 1452 nurses in Belgium, The Netherlands, Portugal, Sweden and the United Kingdom (UK).

Methods

A random selection of 40 photographs was divided in two sets (sets A and B). Both sets contained one photograph of normal skin, one of blanchable erythema, three for each pressure ulcer grade, three of moisture lesions and three of combined lesions. In a combined lesion, a pressure ulcer coincides with a moisture lesion. The photographs were graded and discussed by 12 trustees of the EPUAP, whose opinion is considered the gold standard. All experts had extended experience in the care of pressure ulcers and pressure ulcer classification.

The two sets of photographs were randomly presented to the study participants, who were asked to classify them as normal skin, blanchable erythema, non-blanchable erythema (grade 1 pressure ulcer), blister (grade 2 pressure ulcer), superficial pressure ulcer (grade 3), deep pressure ulcer (grade 4), moisture lesion or combined lesion. No further information was given.

Ethical considerations

The study was approved by the appropriate ethics review committees. The participating nurses received full information before the start of the study. In the questionnaire used for the study, the purpose, procedure, risks and benefits were fully explained, and anonymity and confidentiality were assured. Return of a completed questionnaire was taken as consent to participate.

Data analysis

For each nurse, the percentage of agreement and the Cohen's kappa statistic (κ) were calculated based on the comparison between their assessment and the gold standard. The median percentage of agreement and the median

Table 2 Interpretation of Cohen's kappa according to Landis and Koch (1977)

< 0.00	Poor
0.00–0.20	Slight
> 0.20–0.40	Fair
> 0.40–0.60	Moderate
> 0.60–0.80	Substantial
> 0.80–1.00	Almost perfect

Cohen's kappa were used as summaries of inter-observer reliability.

In contrast to percentage agreement, which measures the total number of occasions on which the raters agreed – including random guesses and chance agreements – the kappa statistic measures the degree of agreement over and above that which may be expected by chance alone. A κ of 0.0 represents agreement equivalent with chance alone, whereas a κ of 1.0 represents perfect agreement. The criteria for the κ statistic by Landis and Koch (1977) were used to interpret the results (Table 2).

The median Cohen's kappa, the interquartile range (IQR), and the median percentage of agreement were calculated using spss® 12.0 (spss® Inc., Chicago, IL, USA). Analyses included the Mann–Whitney *U*-test and Kruskal–Wallis test because of the non-Gaussian distribution of the results. For categorical data, the chi-squared test was used. To describe

the relationship between two variables, Spearman's rho (r_s) was calculated. Microsoft Office EXCEL® 2003 (Microsoft Corporation®, Redmond, WA, USA) was used for graphical presentation of the results. An α level of 0.05 was used for all statistical tests.

Results

A total of 1452 nurses from Belgium (45.9%; $n = 666$), The Netherlands (28.3%; $n = 411$), The UK (15.2%; $n = 221$), Sweden (7.4%; $n = 107$) and Portugal (3.2%; $n = 47$) were involved in this study. About 70% of the nurses were between 20 and 45 years old. A quarter was over the age of 45 years. Approximately, 70% had more than 10 years of experience and 30.1% had been active in nursing practice for more than 20 years (Table 3). All participants stated that they were familiar with use of the EPUAP classification scale.

The median Cohen's kappa for the entire group of nurses was 0.33 when they were asked to assess the total set of photographs (Table 4). To examine the level of inter-observer reliability, respondents were divided into six groups, based on the criteria for the κ statistic by Landis and Koch. About 22% achieved slight assessor agreement ($0 \leq \kappa \leq 0.20$); approximately one-third (37.3%) achieved fair agreement ($0.20 < \kappa \leq 0.40$), another third (33.3%) achieved moderate agreement ($0.40 < \kappa \leq 0.60$) and only 5.0% reached a

Table 3 Basic characteristics of the participating nurses

	Total	Belgium	The Netherlands	Portugal	Sweden	United Kingdom
<i>n</i>	1452	666	411	47	107	221
Gender (%)						
Female	1245 (85.7)	554 (83.2)	353 (85.9)	37 (78.7)	95 (88.8)	206 (93.2)
Age (sd)	38.7 (10.1)	37.1 (9.7)	40.1 (10.8)	34.7 (8.2)	41.1 (11.6)	40.5 (8.8)
Experience (%)						
< 5 years	244 (17.1)	131 (20.0)	61 (15.3)	8 (17.0)	19 (18.3)	25 (11.5)
5–10 years	228 (16.0)	112 (17.1)	58 (14.6)	12 (25.5)	18 (17.3)	28 (12.8)
10–20 years	523 (36.8)	234 (35.7)	155 (38.8)	20 (42.6)	30 (28.8)	84 (38.5)
> 20 years	428 (30.1)	178 (27.2)	125 (31.3)	7 (14.9)	37 (35.6)	81 (37.2)
Education (%)						
Undergraduate	711 (49.0)	275 (41.4)	264 (64.2)	41 (87.2)	29 (27.1)	102 (46.1)
Bachelor's degree	667 (46.1)	365 (55.1)	120 (29.2)	6 (12.8)	72 (67.3)	104 (47.1)
Master's degree	71 (4.9)	23 (3.5)	27 (6.6)	0 (0)	6 (5.6)	15 (6.8)
Expertise (%)						
Expert	57 (4.0)	12 (1.9)	35 (8.8)	0 (0)	1 (1.0)	9 (4.1)
Extensive	372 (26.4)	120 (18.6)	116 (29.0)	3 (6.4)	17 (16.2)	116 (54.0)
Basic	792 (56.1)	427 (66.3)	201 (50.2)	36 (76.6)	55 (52.3)	73 (34.0)
Limited	190 (13.5)	85 (13.2)	48 (12.0)	8 (17.0)	32 (30.5)	17 (7.9)
Work location (%)						
Hospital	727 (55.0)	344 (53.2)	201 (57.3)	6 (13.0)	70 (68.0)	106 (60.9)
Nursing home	245 (18.5)	134 (20.7)	85 (24.2)	1 (2.2)	11 (10.6)	14 (8.1)
Home care	286 (21.7)	149 (23.0)	48 (13.7)	38 (82.6)	8 (7.8)	43 (24.7)
Education	63 (4.8)	20 (3.1)	17 (4.8)	1 (2.2)	14 (13.6)	11 (6.3)

Table 4 Comparison of inter-rater reliability by country, experience, level of education, expertise and work location

	Classification of the total set of photographs		Distinction between pressure ulcers and moisture lesions*		Classification of the pressure ulcer photographs	
	κ (IQR)	Kruskal–Wallis' χ^2 <i>P</i> value	κ (IQR)	Kruskal–Wallis' χ^2 <i>P</i> value	κ (IQR)	Kruskal–Wallis' χ^2 <i>P</i> value
Total group	0.33 (0.21–0.47)		0.36 (0.20–0.51)		0.29 (0.14–0.47)	
Country						
Belgium	0.36 (0.24–0.48)	$\chi^2 = 111.92$	0.38 (0.20–0.53)	$\chi^2 = 63.86$	0.28 (0.14–0.47)	$\chi^2 = 83.93$
The Netherlands	0.38 (0.25–0.47)	<i>P</i> < 0.001	0.37 (0.23–0.51)	<i>P</i> < 0.001	0.37 (0.23–0.48)	<i>P</i> < 0.001
Portugal	0.37 (0.30–0.53)		0.46 (0.36–0.57)		0.27 (0.12–0.49)	
Sweden	0.23 (0.12–0.30)		0.26 (0.12–0.37)		0.19 (0.09–0.29)	
United Kingdom	0.24 (0.13–0.37)		0.28 (0.15–0.46)		0.20 (0.05–0.37)	
Experience						
< 5 years	0.30 (0.18–0.45)	$\chi^2 = 6.48$	0.31 (0.18–0.46)	$\chi^2 = 9.03$	0.29 (0.09–0.47)	$\chi^2 = 1.91$
5–10 years	0.32 (0.19–0.44)	<i>P</i> = 0.09	0.35 (0.18–0.49)	<i>P</i> = 0.03	0.27 (0.11–0.46)	<i>P</i> = 0.59
10–20 years	0.35 (0.24–0.47)		0.37 (0.21–0.51)		0.29 (0.15–0.48)	
> 20 years	0.33 (0.21–0.47)		0.37 (0.19–0.52)		0.29 (0.15–0.44)	
Education						
Undergraduate	0.32 (0.19–0.45)	$\chi^2 = 11.87$	0.34 (0.18–0.49)	$\chi^2 = 9.36$	0.29 (0.13–0.47)	$\chi^2 = 3.32$
Bachelor	0.35 (0.21–0.47)	<i>P</i> = 0.04	0.38 (0.22–0.52)	<i>P</i> = 0.009	0.28 (0.14–0.47)	<i>P</i> = 0.19
Master	0.39 (0.26–0.53)		0.42 (0.20–0.56)		0.34 (0.19–0.48)	
Expertise						
Expert	0.47 (0.36–0.53)	$\chi^2 = 63.33$	0.51 (0.36–0.59)	$\chi^2 = 65.01$	0.47 (0.32–0.56)	$\chi^2 = 36.19$
Extensive	0.36 (0.24–0.48)	<i>P</i> < 0.001	0.41 (0.25–0.54)	<i>P</i> < 0.001	0.31 (0.16–0.47)	<i>P</i> < 0.001
Basic	0.33 (0.19–0.45)		0.35 (0.19–0.49)		0.28 (0.12–0.47)	
Limited	0.26 (0.14–0.37)		0.27 (0.14–0.42)		0.25 (0.09–0.38)	
Work location						
Hospital	0.35 (0.20–0.47)	$\chi^2 = 14.23$	0.35 (0.19–0.51)	$\chi^2 = 22.41$	0.29 (0.14–0.48)	$\chi^2 = 3.30$
Nursing home	0.31 (0.23–0.42)	<i>P</i> = 0.003	0.32 (0.20–0.49)	<i>P</i> < 0.001	0.29 (0.15–0.42)	<i>P</i> = 0.35
Home care	0.36 (0.25–0.48)		0.42 (0.28–0.57)		0.29 (0.12–0.47)	
Education	0.30 (0.18–0.41)		0.33 (0.15–0.46)		0.25 (0.10–0.43)	

κ = Cohen's kappa; IQR, interquartile range.

*The four pressure ulcer grades were considered as 'pressure ulcers'. The different grades were not taken in account.

substantial level ($0.60 < \kappa \leq 0.80$). An overview of the results is presented in Table 5.

Inter-observer reliability was higher in more experienced nurses when assessing the differential diagnosis between moisture lesions and pressure ulcers (Table 4).

Differences were statistically significant between the classification skills of nurses working in a hospital environment, home care, a nursing home and an educational setting (Table 4). Those who worked in an educational setting reached a statistically significant lower inter-observer agreement ($\kappa = 0.30$) than those who worked in a clinical setting (hospital environment, home care and nursing home) ($\kappa = 0.35$, IQR = 0.22–0.47) when classifying the total set of photographs (Mann–Whitney's *U*-test = -2.037 , *P* = 0.04).

Differences were found between level of basic nursing education and classification skills (Table 4). Nurses with an undergraduate degree ($\kappa = 0.32$) achieved a statistically significant lower inter-observer agreement than those with

a Master's degree ($\kappa = 0.39$) (Mann–Whitney's *U*-test = -2.334 , *P* = 0.02).

Nurses who stated that they were experts reached a median kappa value of 0.47 (Table 4). Those who said that they had basic experience obtained a statistically significant lower median kappa value of 0.33 (Mann–Whitney's *U*-test = -5.464 ; *P* < 0.001).

When making the differential diagnosis between moisture lesions and pressure ulcers, nurses who had attended training in wound care reached a slightly higher median Cohen's kappa than those nurses who had not attended this specific training ($\kappa = 0.37$ vs. $\kappa = 0.34$; Mann–Whitney's *U*-test = -2.877 , *P* = 0.004). No correlation was found between duration of education and nurses' classification skills ($r_s = 0.005$, *P* = 0.88). The classification skills of those who frequently (at least once a month) read literature about pressure ulcers were statistically significant and better than those who never read this type of literature ($\kappa = 0.36$ vs. $\kappa = 0.28$; Mann–Whitney's *U*-test = -3.551 , *P* < 0.001).

Table 5 The inter-observer agreement (Landis & Koch 1977) presented for the total group by country, experience, level of education, self-attributed expertise in wound care and work location

<i>n</i> (%)	Poor ($\kappa < 0$)	Slight ($0 \leq \kappa \leq 0.20$)	Fair ($0.20 < \kappa \leq 0.40$)	Moderate ($0.40 < \kappa \leq 0.60$)	Substantial ($0.60 < \kappa \leq 0.80$)	Almost perfect ($0.80 < \kappa \leq 1.00$)
Total group	29 (2.0)	324 (22.3)	541 (37.3)	484 (33.3)	72 (5.0)	2 (0.1)
Country						
Belgium	5 (0.8)	143 (21.5)	236 (35.4)	233 (35.0)	48 (7.2)	1 (0.2)
The Netherlands	6 (1.5)	63 (15.3)	149 (36.3)	175 (42.6)	17 (4.1)	1 (0.2)
Portugal	0 (0.0)	5 (10.6)	21 (44.7)	20 (42.6)	1 (2.1)	0 (0.0)
Sweden	5 (4.7)	43 (40.2)	45 (42.1)	13 (12.1)	1 (0.9)	0 (0.0)
United Kingdom	13 (5.9)	70 (31.7)	90 (40.7)	43 (19.5)	5 (2.3)	0 (0.0)
Experience						
< 5 years	5 (1.0)	71 (29.1)	80 (32.8)	78 (32.0)	10 (4.1)	0 (0.0)
5–10 years	7 (3.1)	51 (22.4)	91 (39.9)	67 (29.4)	12 (5.3)	0 (0.0)
10–20 years	7 (1.3)	103 (19.7)	201 (38.4)	188 (35.9)	22 (4.2)	2 (0.4)
> 20 years	9 (2.1)	94 (22.0)	158 (36.9)	139 (32.5)	28 (6.5)	0 (0.0)
Education						
Undergraduate	12 (1.7)	172 (24.2)	270 (38.0)	229 (32.2)	27 (3.8)	1 (0.1)
Bachelor	13 (1.9)	137 (20.5)	252 (37.8)	229 (34.3)	35 (5.2)	1 (0.1)
Master	3 (4.2)	14 (19.7)	19 (26.8)	25 (35.2)	10 (14.1)	0 (0.0)
Expertise						
Expert	1 (1.8)	2 (3.5)	13 (22.8)	32 (56.1)	9 (15.8)	0 (0.0)
Extensive	5 (1.3)	61 (16.4)	143 (38.4)	139 (37.4)	24 (6.5)	0 (0.0)
Basic	16 (2.0)	189 (23.9)	289 (36.5)	263 (33.2)	34 (4.3)	1 (0.1)
Limited	7 (3.7)	63 (33.2)	76 (40.0)	39 (20.5)	4 (2.1)	1 (0.5)
Work location						
Hospital	16 (2.2)	165 (22.7)	255 (35.1)	256 (35.2)	34 (4.7)	1 (0.1)
Nursing home	0 (0.0)	54 (22.0)	111 (45.3)	70 (28.6)	9 (3.7)	1 (0.4)
Home care	5 (1.7)	47 (16.4)	104 (36.4)	108 (37.8)	22 (7.7)	0 (0.0)
Education	3 (4.8)	16 (25.4)	27 (42.9)	14 (22.2)	3 (4.8)	0 (0.0)

If only pressure ulcer photographs were considered, the median Cohen's kappa was 0.29 (Table 4). Approximately, one-third of the photographs was scored one grade too low. Grade 3 was most frequently classified incorrectly (64.5%; 2717/4211). In 33.5% of the observations, nurses classified a grade 3 lesion as grade 2 (blister). Non-blanchable erythema (grade 1) was assessed incorrectly in 39.9% (1694/4273) of the observations. In approximately 40% of the observations, grade 1 lesions were confused with blanchable erythema (Figure 1).

In 72.7% (12300/16913) of the observations of pressure ulcer photographs, the lesions were assessed correctly as a pressure ulcer. Only 22.0% (932/4231) of the observations of moisture lesion photographs were assessed correctly. In 22.0% of the observations, they were seen as a combined lesion, in 19.9% as grade 2, in 16.2% as grade 3 and in 10.2% as a grade 1 pressure ulcer.

Discussion

Inter-observer reliability of the EPUAP classification system was found to be low. Pressure ulcer photographs were often

classified erroneously and only a minority of nurses was able to reach a substantial level of agreement. Similar results were found for the differential diagnosis between moisture lesions and pressure ulcers. The discussion here will focus on three hypotheses for debate. A first hypothesis will focus on the clarity of the EPUAP classification system. A second will concentrate on the complexity of the system. In a third hypothesis, the familiarity of the nurses with the use of the EPUAP classification system will be considered.

The first hypothesis is based on the common confusion between reactive hyperaemia (blanchable erythema) and non-blanchable erythema (grade 1) and on the confusion between moisture lesions and pressure ulcers. This confusion might be caused by unclear definitions of blanchable erythema and grade 1 pressure ulcers given in the actual classification system.

The distinction between a grade 1 pressure ulcer and blanchable erythema is based on the reaction of the tissue to pressure and shearing forces. Blanchable erythema is defined as a normal reactive hyperemic response of the skin after an arterial occlusion. Microcirculation stays intact and tissue damage has not yet occurred (Collier 1999). On the contrary, a grade 1 pressure ulcer indicates clinically visible damage

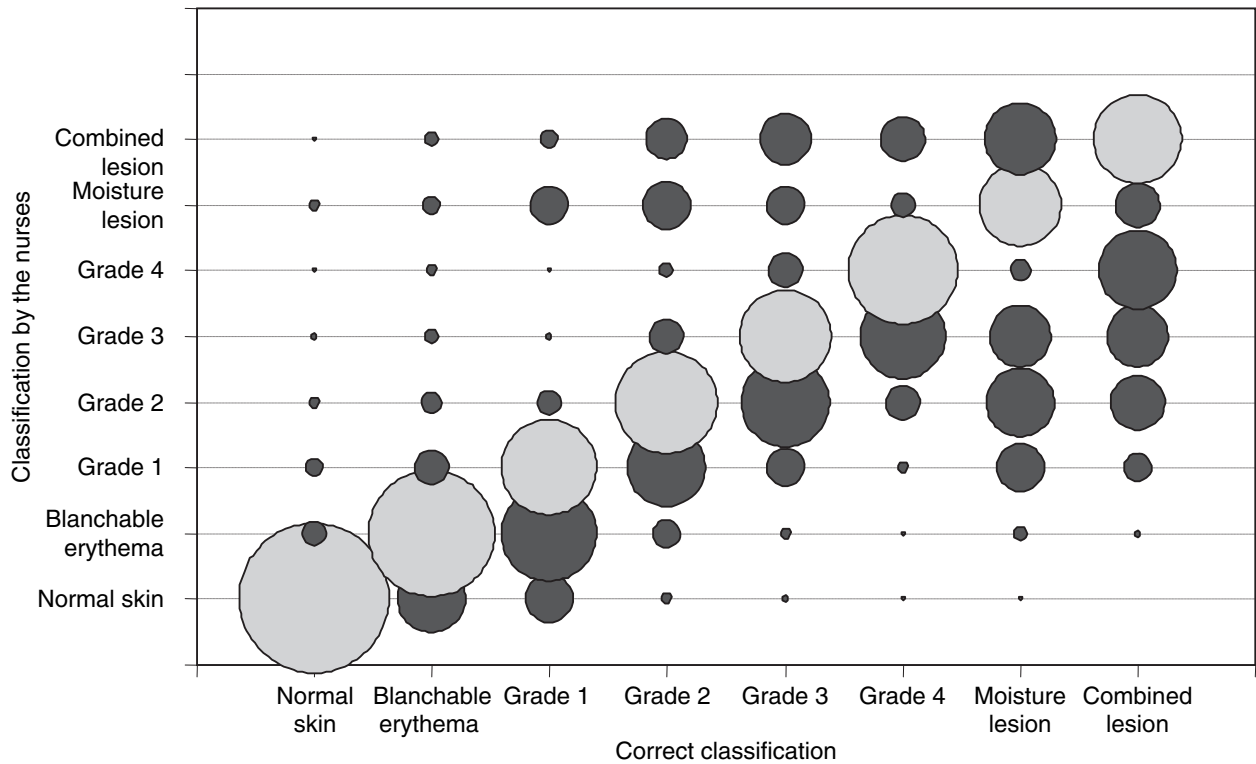


Figure 1 Classification by the nurses ($n = 1452$) compared with the correct classification. The size of the grey circles represents the number of correct classifications; the size of the black circles represents the number of incorrect classifications.

because of pressure and shearing forces and is defined as an abnormal response, presenting as a ‘persistent redness’ of the intact skin. Warmth, oedema, induration or hardness may also be used as indicators, particularly in individuals with a dark skin [Derre *et al.* 1999, European Pressure Ulcer Advisory Panel (EPUAP) 1999].

Non-blanchable erythema is statistically and significantly associated with the development of pressure ulcers (Allman *et al.* 1995). As reported in a study by Vanderwee *et al.* (2007), preventive measures must be taken as soon as non-blanchable erythema occurs. Vanderwee *et al.* concluded that statistically significant fewer patients needed preventive measures when prevention was postponed until non-blanchable erythema appeared, and those patients did not develop more pressure ulcers than patients who received prevention based on the standard risk assessment method (Braden < 17). Using the appearance of non-blanchable erythema to allocate preventive measures led to a considerable reduction of patients in need of prevention without resulting in an increase in pressure ulcers. Prevention must be predominantly aimed at protection or repair of the oxygen supply to the tissue by reducing the intensity and/or duration of pressure and shearing forces. Confusion between a grade 1 pressure ulcer and blanchable erythema might result in a delayed application of preventive interventions.

A moisture lesion is characterized by erosion of the epidermis and a macerated appearance of the skin. It is caused by the sustained presence of urine, faeces, perspiration or wound fluid, and not by a deficiency of oxygen within the tissue. A correct distinction between pressure ulcers and moisture lesions is important in practice because the preventive measures to be taken are different. Skin protection, hygiene and micturition training are indicated for moisture lesions (Maklebust & Sieggreen 1995, Bennett *et al.* 1998). As mentioned above, protection or repair of oxygen supply to the tissue is indicated for the prevention of pressure ulcers. Unambiguous clinical descriptors of the distinction between moisture lesions and pressure ulcers will probably avoid the inadequate application of preventive interventions. Yet, those descriptors are not provided within the current classification system.

The second hypothesis concerns the complexity of the EPUAP system. This hypothesis is based on the confusion between grades 3 and 2 pressure ulcers.

Grade 3 pressure ulcers were often classified as grade 2. The distinction between these grades is based on the type of skin loss: partial- and full-thickness skin loss. Partial-thickness skin loss is defined as a shallow crater involving a loss of the epidermis and/or dermis, and includes grade 2 pressure ulcers (EPUAP 1999). Full-thickness skin loss involves all

tissue layers, and includes grades 3 and 4 pressure ulcers (EPUAP 1999). Observation of the different tissue layers involved appeared to be difficult.

The complexity of the current classification system is an important topic in an international pressure ulcer debate (Donnelly 2005, National Pressure Ulcer Advisory Panel 2007). Both the EPUAP and the NPAUP take a different course.

Discussion within EPUAP concerns reduction of the number of pressure ulcer grades. The distinction between grades 2 and 4 pressure ulcers is of little relevance for their treatment. Suggestions about possible treatment approaches should be more defined. For the prevention and timely detection of pressure ulcers, it would be preferable to use a less complex, 3-grade classification system, which makes the distinction between non-blanchable erythema, a superficial and a deep pressure ulcer. The identification of non-blanchable erythema is particularly critical in differentiating early pressure-induced damage from a normal response to external pressure and for starting timely prevention. If a pressure ulcer develops, wound assessment and evaluation tools, such as the 'TIME'-framework (Fletcher 2005), the 'MEASURE'-framework (Keast *et al.* 2004) or the Pressure Sore Status Tool (Bates-Jensen 1997), can be used. By means of these tools, the characteristics of a wound can be assessed and treatment determined.

National Pressure Ulcer Advisory Panel has increased the number of pressure ulcer grades by adding two more: 'deep tissue injury' and 'cannot be staged'. Deep tissue injury is defined as a purple or maroon intact skin area or a blood-blisters. The lesion is characterized as firm, gentle, gelatinous, warmer, colder or more painful than the surrounding tissue. The surrounding tissue can be damaged rapidly, even if the treatment is optimal. 'Cannot be staged' is defined as a pressure ulcer, which is impossible to assess because of the presence of softened necrosis (yellow, beige, grey, green or brown) and/or a necrotic crust (beige, brown or black) in the wound bed (Black *et al.* 2007, National Pressure Ulcer Advisory Panel 2007). The addition of these two stages is a result of the statement that re-classification is not accepted, even when tissue damage appears to be more extensive than initially thought. Reduction of the risk of receiving no re-imburement or being litigated if a pressure ulcer deteriorates, despite optimal care, should probably also be taken into account when healthcare systems are based on insurance and/or self-payment schemes.

Nurses' familiarity with use of the EPUAP classification system will be considered as a third hypothesis. The impact of basic nursing education and additional training will be discussed in turn.

Poor inter-observer agreement was found for all levels of basic nursing education. Although inter-observer agreement in nurses with a Master's degree was higher than in those with an undergraduate degree, the results were anything but optimal. The slightly higher inter-observer agreement might result from this group possibly being more stimulated by their educational background to read supplementary evidence-based literature and to reflect more thoroughly on daily practice. The development of a positive attitude towards life-long learning seems to be important and needs to be fully supported. Creating high quality educational programmes, allowing nurses to learn how to classify pressure ulcers and how to differentiate other lesions, requires extended knowledge and experience in the field of pressure ulcers. Nurse educators should be encouraged to design such programmes. In this respect, the statistically significant lower inter-observer agreement of nurses working in the educational field is rather worrying.

Although there was greater inter-observer agreement in nurses who identified themselves as expert in wound care, the results were not optimal. Expertise can be obtained by training, which was defined as reading evidence-based literature and following courses about wound care. Both reading evidence-based publications and following courses resulted in better classification skills, and again a positive attitude towards life-long learning might be important in attaining higher classification skills.

Study limitations

A first limitation of this study is the use of convenience sampling. The nurses all stated that they were familiar with the EPUAP classification system. Therefore, it is reasonable to assume that the results presented in this study are rather too 'positive'.

A second limitation might be the use of photographs. Photographs provide merely a static, two-dimensional image of the wound. The visibility of the different tissue layers might be limited. Whether assessment, in practice, is easier than with photographs is unknown. In practice, more aspects can be involved in the assessment, such as the patient's medical history, wound history, mobility, incontinence status and nutritional condition.

Conclusion

European Pressure Ulcer Advisory Panel appears to be aware of the limitations in the current classification system. Efforts to clarify the difference between moisture lesions and pressure ulcers are being made. In a recent position

What is already known about this topic

- The European Pressure Ulcer Advisory Panel (EPUAP) classification system is widely used to determine the severity of pressure ulcers.
- The purpose of a classification system is to standardize record-keeping and provide a common description for the purposes of clinical practice, audit and research.
- The distinction between pressure ulcers and moisture lesions is important in clinical practice because the preventive measures to be taken are different.

What this paper adds

- Inter-observer reliability of the EPUAP classification system was low.
- Moisture lesions were often mistaken for pressure ulcers and blanchable erythema and grade 1 pressure ulcers were frequently confused.
- A thorough evaluation of the classification system is necessary and needs to focus on the clarity and complexity of the system.

statement, EPUAP defined wound-related characteristics (causes, location, shape, depth, edges and colour) and patient-related characteristics to clarify the difference between a pressure ulcer and a moisture lesions. In addition, an e-learning programme has been developed to disseminate information about the system (<http://www.epuap.org/epuap>). However, much more work is needed to reduce the difficulties experienced with the present classification system.

Author contributions

SL and DT were responsible for the study conception and design and BD and DT were responsible for the drafting of the manuscript. BD, SL, FJ, FK, GL, HH, LC, PL and DT performed the data collection and BD and DT performed the data analysis. BD, SL, FJ, FK, GL, HH, LC, PL, VJ and DT made critical revisions to the paper. DT supervised the study.

References

Allman R.M., Goode P.S., Patrick M.M., Burst N. & Bartolucci A.A. (1995) Pressure ulcer risk factors among hospitalized patients with activity limitation. *Journal of the American Medical Association* 273, 865–870.

Bates-Jensen B.M. (1997) Pressure ulcer assessment and documentation: the pressure sore status tool. In *Chronic Wound Care:*

A Clinical Source Book for Healthcare Professionals, 2nd edn (Krasner D.L. & Kane D., eds), Health Management Publications, Inc., Wayne, PA, pp. 37–48.

Bennett R.G., Baran P.J., DeVonde L., Bacetti H., Kristo B., Tayback M. & Greenough W.B. (1998) Low airloss hydrotherapy versus standard care for incontinent hospitalized patients. *Journal of the American Geriatrics Society* 46, 569–576.

Black J., Bahrestani M., Dorner B., Edsberg L., Langemo D., Taler G. & Zulkowski K. (2007) Proposed staging system and deep tissue injury definitions with descriptions. In *Charting the Course for Pressure Ulcer Prevention and Treatment* (Saunders D.A. & Knapp S., eds). 10th National NPUAP Biennial Conference & 20th Anniversary Celebration, San Antonio, TX, USA.

Bours G.J., Halfens R.J., Lubbers M. & Haalboom J.R. (1999) The development of a national registration form to measure the prevalence of pressure ulcers in The Netherlands. *Ostomy/Wound Management* 45, 28–40.

Collier M. (1999) Blanching and non-blanching hyperaemia. *Journal of Wound Care* 8, 63–66.

Dealey C. & Lindholm C. (2006) Pressure ulcer classification. In *Science and Practice of Pressure Ulcer Management* (Romanelli M., Clark M., Colin D. & Defloor T., eds). European Pressure Ulcer Advisory Panel & Springer-Verlag, London, pp. 37–41.

Defloor T. & Schoonhoven L. (2004) Inter-rater reliability of the EPUAP pressure ulcer classification system using photographs. *Journal of Clinical Nursing* 13, 952–959.

Defloor T., Schoonhoven L., Vanderwee K., Weststrate J. & Myny D. (2006) Reliability of the European Pressure Ulcer Advisory Panel classification system. *Journal of Advanced Nursing* 54, 189–198.

Derre B., Grypdonck M. & Defloor T. (1999) *The Development of Nonblanchable Erythema in Intensive Care Patients*. Paper presented at the meeting of Sigma Theta Tau International, 11th International Nursing Research Conference, London.

Donnelly J. (2005) Should we include deep tissue injury in pressure ulcer staging systems? The NPUAP debate. *Journal of Wound Care* 14, 207–210.

European Pressure Ulcer Advisory Panel (EPUAP) (1999) Guidelines on treatment of pressure ulcers. *EPUAP Review* 1, 31–33.

Fletcher J. (2005) Wound bed preparation and the TIME principles. *Nursing Standard* 20, 57–65.

Guggenmoos-Holzmann I. (1993) How reliable are chance-corrected measures of agreement? *Statistics in Medicine* 12, 2191–2205.

Keast D.H., Bowering C.K., Evans A.W., Mackean G.L., Burrows C. & D'Souza L. (2004) MEASURE: a proposed assessment framework for developing best practice recommendations for wound assessment. *Wound Repair and Regeneration* 12, S1–S17.

Landis J.R. & Koch G.G. (1977) The measurement of observer agreement for categorical data. *Biometrics* 33, 159–174.

Maklebust J. & Sieggreen M. (1995) *Pressure Ulcers: Guidelines for Prevention and Nursing Management*. Springhouse Corporation, Pennsylvania.

National Pressure Ulcer Advisory Panel. (2007a) *10th National NPUAP Biennial Conference & 20th Anniversary Celebration*. NPUAP, San Antonio, TX, USA.

National Pressure Ulcer Advisory Panel. (2007b) Pressure ulcers: definitions and stages. NPUAP, Washington DC. Retrieved April 4, 2007, from http://www.npuap.org/documents/PU_Definition_Stages.pdf.

- Nixon J., Thorpe H., Barrow H., Phillips A., Nelson E.A., Mason S.A. & Cullum N. (2005) Reliability of pressure ulcer classification and diagnosis. *Journal of Advanced Nursing* **50**, 613–623.
- Pedley G. (2004) Comparison of pressure ulcer grading scales: a study of clinical utility and inter-rater reliability. *International Journal of Nursing Studies* **41**, 129–140.
- Polit D.F. & Beck C.T. (2003) *Nursing Research: Principles and Methods*. Lippincott Williams & Wilkins, Philadelphia.
- Russell L. & Reynolds T.M. (2001) How accurate are pressure ulcer grades? An image-based survey of nurse performance *Journal of Tissue Viability* **11**, 67–75.
- Sharp A. (2004) Pressure ulcer grading tools: how reliable are they? *Journal of Wound Care* **13**, 75–77.
- Shea J. (1975) Pressure sores: classification and management. *Clinical Orthopaedics and Related Research* **112**, 89–100.
- Vanderwee K., Defloor T. & Grypdonck M. (2007) Non-blanchable erythema as an indicator for the need for pressure ulcer prevention: a randomised controlled trial. *Journal of Clinical Nursing* **16**, 325–335.
- Witkowski J.A. & Parish L.C. (1981) Histopathology of the decubitus ulcer. *Journal of the American Academy of Dermatology* **6**, 1014–1021.
- Wound, Ostomy, and Continence Nurses Society (WOCN). (2003) *Guideline for Prevention and Management of Pressure Ulcers*. Wound, Ostomy, and Continence Nurses Society (WOCN), Glenview (IL).