Australian Critical Care 36 (2023) 179–185

Contents lists available at ScienceDirect

Australian Critical Care

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

Research paper

Inter-rater reliability of descriptors for the classification of mucosal pressure injury: A prospective cross-sectional study

Simone Fitzgerald, RN, MN [a, b], Lauren McTier, RN, PhD [a, c], Christina Whitehead, RN, BN, MBioethics [d], Kristy Masters, RN, Grad Cert CritCare [d], Rochelle Wynne, RN, PhD [a, e, *]

[a] School of Nursing & Midwifery, Deakin University, 1 Geringhap Street, Geelong, Victoria, Australia; [b] Intensive Care Unit, Austin Health, Studley Road, Heidelberg, Victoria, Australia; [c] Centre for Quality & Patient Safety Research, Institute for Health Transformation, Deakin University, Geelong, Australia; [d] Intensive Care Unit, Nepean Hospital, Penrith, NSW, Australia; [e] Western Sydney Nursing & Midwifery Research Centre, Blacktown Clinical & Research School, Western Sydney University & Western Sydney Local Health District, Blacktown Hospital, Marcel Crescent Blacktown, NSW, Australia

ARTICLE INFORMATION

Article history:
Received 4 September 2021
Received in revised form
6 December 2021
Accepted 6 December 2021

Keywords:
Critical care
Critical care nursing
Nursing research
Pressure ulcer classification
Pressure ulcer nursing
Pressure ulcer prevention and control
Quality of health care

ABSTRACT

Background: Mucosal pressure injuries (PIs) are usually caused by pressure from essential medical devices. There is no universally accepted criterion for assessment, monitoring, or reporting mucosal PI. Reliable descriptors are vital to benchmark the frequency and severity of this hospital-acquired complication.

Objectives: The objective of this study was to determine whether modified Reaper Oral Mucosa Pressure Injury Scale (ROMPIS) descriptors improved the reliability of mucosal PI assessment. Secondary aims were to explore nurses' knowledge of and attitudes toward mucosal PI.

Methods: A prospective cross-sectional survey was distributed to nurses from two tertiary affiliated intensive care units via REDCap® to capture demographic data, knowledge, attitudes, and inter-rater reliability (IRR) measures. Nurses were randomised at a 1:1 ratio to original or modified ROMPIS descriptors and classified 12 images of mucosal PI. IRR was assessed using percentage agreement, Fleiss' kappa, and intraclass correlation coefficients.

Results: The survey response rate was 20.9% (n = 98/468), with 73.5% (n = 72/98) completing IRR measures. Agreement was higher with modified (75%) than original ROMPIS descriptors (69.4%). IRR was fair for the original ($\kappa = 0.30$, 95% confidence interval [CI] [0.28, 0.33], z 26.5, $p < 0.001$) and modified ROMPIS ($\kappa = 0.29$, 95% CI [0.26, 0.31], z 25.0, $p < 0.001$). Intraclass correlation coefficient findings indicated ratings were inconsistent for the original (0.33, 95% CI [0.18, 0.59], F 18.8 (11 df), $p < 0.001$) and modified ROMPIS (0.31, 95% CI [0.17, 0.57], F 17.6 (11 df), $p < 0.001$). PI-specific education and risk factor recognition were common.

Conclusion: Modified descriptors had marginally better agreement. Participants understand management and prevention but need to strengthen their perceived capacity for mucosal PI risk assessment. This work provides a foundation for future benchmarking and a platform from which further research to refine and test descriptors specific to mucosal PI can be generated.

© 2021 Australian College of Critical Care Nurses Ltd. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Pressure injuries (PIs) are considered an indicator of nursing care quality both in Australia and worldwide.[1] Consensus exists that most PIs are preventable, particularly if reliable intervention programs are implemented in healthcare settings.[2] Consequently, minimising harm from PI is linked to most hospital risk registers[3] and reliable Australian data on PI identification, prevention, and reporting are a priority.[4] Accurate descriptors for PI classification are required to allow for a common description of PI severity for the purposes of clinical audit, quality monitoring, and research, as well as optimisation of strategies for prevention and early treatment interventions.[5] Classifying PI severity involves the use of staging, or grading tools. The National Pressure Injury Advisory Panel grades PIs into four stages of severity and is the most common tool globally.[6] However, it cannot be used to classify PI occurring on areas of the body that are structurally different to human skin, such as the mucous membranes.[7] The National Pressure Injury Advisory Panel

* Corresponding author at: Westerns Sydney Nursing & Midwifery Research Centre, Western Sydney University & Western Sydney Local Health District, Australia.
E-mail address: R.Wynne@westernsydney.edu.au (R. Wynne).

https://doi.org/10.1016/j.aucc.2021.12.004
1036-7314/© 2021 Australian College of Critical Care Nurses Ltd. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

do recommend mucosal PI are included in PI prevalence and incidence studies despite the absence of a specific classification category and noted as a PI without a stage identified.

Mucosal PIs predominantly occur in patients admitted to intensive care units (ICUs) and are caused by medical devices such as endotracheal or enteral feeding tubes.[8] In areas other than the ICU, medical device use is less common, so estimations of mucosal PI are often very low. Analysis of data collected over 2 years from 18 Australian hospitals in Queensland showed hospital-acquired PI prevalence was 3.4% (n = 7291). ICU patients in this cohort were 3.8 times more likely to acquire a PI than those not in the ICU, and the incidence of ICU medical device–related PI was 22% in contrast to 2% for patients not in the ICU.[9] In a secondary analysis of data collected for eight quarterly PI incidence and prevalence studies, the overall rate of hospital-acquired PI was 5.4% (n = 113) and 34.5% (n = 39) of these were medical device related PI. When patients had a medical device, they were 2.4 times more likely to develop a PI.[10] A prospective observational study of consecutive patients (n = 483) admitted to six ICUs and followed up for 7 days identified 61 hospital-acquired PIs in 48 (9.9%) patients, and 15 (3.1%) of these patients had 20 medical device–related PIs.[11] A recent systematic review found few studies intentionally report mucosal PI incidence or prevalence, despite data indicating that more than one-third of medical device–related PIs may be mucosal.[12] There is minimal literature focused on reliable descriptors for classification of mucosal PI that may compound problems with reporting. Reaper et al[13] developed the Reaper Oral Mucosa Pressure Injury Scale (ROMPIS) that, when tested in a cohort of ICU nurses, had fair inter-rater reliability (IRR) ($\alpha = 0.307$, 95% confidence interval [CI] [0.20, 0.41]), a finding attributed to nurses' lack of confidence with mucosal pressure injury (MPI) assessment. In the context of limited options for mucosal PI assessment, the ROMPIS provides a benchmark and warrants further validation in the ICU setting.

Nurses' prevention knowledge and attitudes have been extensively explored in the context of PI on human skin;[14] less is known regarding their approach to the prevention of mucosal PI. Nurses lead PI prevention[14] through identification of patients at risk and monitoring and maintaining skin integrity with appropriate prevention strategies.[15] Given most mucosal PIs occur in critical care areas, understanding the knowledge and attitudes of ICU nurses towards mucosal PI is vital to identify appropriate preventative strategies and opportunities to improve education and practice. The aim of this study was to determine whether modified ROMPIS (M-ROMPIS) descriptors improved IRR in mucosal PI assessment. Secondary aims were to explore nurses' knowledge of, and attitudes towards, mucosal PI.

2. Methods

2.1. Design

A prospective cross-sectional survey was created in Research Electronic Data Capture (REDCap®) to test IRR of the standard[13] and modified mucosal PI descriptors and nurses' perceptions and knowledge of mucosal PIs and their prevention. Human Research and Ethics Committee (HREC) approval was granted under the National Mutual Acceptance Scheme and site-specific approvals as required (HREC/60121-Austin-2019; 2020/STE04211; DUHREC 2020–234).

2.2. Setting & sample

The population consisted of ICU nurses from two public tertiary teaching hospitals, one in NSW and one in Victoria, Australia. The Victorian ICU had 23 beds, treating approximately 2200 patients per year, and 318 nursing staff members at the time of the survey. The NSW site had 22 beds, treating approximately 2300 patients per year, with 150 nursing staff members. The invitation to participate was distributed to all ICU nursing staff members in both units, in July 2020 via staff email distribution lists, followed by a reminder email invitation 2 weeks later. The invitation contained the Participant Information Sheet and a link to the REDCap® survey; consent was implied by participation. To maintain study integrity, participants were asked not to discuss their responses with others. The survey remained open for 6 weeks, from the 3rd of July to the 16th of August 2021, to give participants sufficient time to complete the survey and submit their responses. Strategies, such as sending a reminder email 2 weeks after the initial invitation, including the current response rate in the reminder email to potentially motivate individuals to complete the survey as recommended by McPeake et al,[16] and having ICU champions reminding staff about the survey during handover, double staffing, and periods of lower acuity, were used to optimise response rates.

2.3. Instruments

The standard set of mucosal PI descriptors were those designed by Reaper et al.,[13] in the original ROMPIS tool. The ROMPIS has a three-stage classification system of severity; stage 1 represented the least severe, and Stage 3 the most severe, as well as the option of classifying a mucosal PI 'unstageable'. To construct the modified descriptors, an expert panel of ICU nurses revised the standard ROMPIS tool descriptors with language consistent with commonly used PI assessment nomenclature while maintaining clarity of differences between mucosal PI and skin PI. The panel was comprised of six critical care nurses, four in clinical roles and two academics from different universities. Two critical care nurses were from the same healthcare service. No members of the panel completed the survey. The lead investigator drafted changes which the panel reviewed independently and then discussed to achieve consensus. It was anticipated that familiarity with the language in the revised tool, the M-ROMPIS, would result in users finding this tool simpler to interpret and use. Additionally, the option 'unstageable' was omitted from response options for both tools in this study. Reaper et al[13] found limited consensus of participants that deemed images as 'unstageable', and IRR improved when the 'unstageable' results were omitted. It may be that availability of options such as 'unstageable' allows for generation of coefficients not specifically valid to the context and thus may limit IRR estimations in classification systems.[17] Standard and modified descriptors of the ROMPIS tool are listed in Table 1.

The REDCap® survey consisted of three sections. Section 1 contained participant demographic information. Section 2 comprised questions relevant to the secondary aim of exploring ICU nurses' perceptions and knowledge of mucosal PIs and their prevention. These questions were based on the Moore and Price[18] (2004) Staff Attitude Scale towards PI prevention and were adapted to fit the mucosal PI context. The modifications only included the addition of the word mucosal to PI statements. Modifications were reviewed and agreed upon by the convened expert panel of ICU nurses. Instrument responses were in the form of a Likert scale, free text, or checkboxes. Once section 2 of the survey was complete, section 3 became available and participants were randomised by an automated rule in REDCap® to complete the IRR assessment for the ROMPIS or M-ROMPIS. Randomisation followed a 1:1 allocation sequence to ensure each tool was used a similar number of times. Section 3 included 12 images of mucosal PI, from each of the three grades of severity described in the tools. Participants were asked to classify each mucosal PI into one of the three stages of severity according to the tool descriptors provided. Images were sourced from clinical photographs used with permission as well as from

Table 1
ROMPIS and M-ROMPIS tool criterion.

Stage	Original ROMPIS tool	M-ROMPIS tool
Stage 1	Redness and demarcation of the lip and buccal mucosa, with no visible destruction or loss of epithelial tissue, ulceration, or blisters. Nonblanchable erythema on the corners of the mouth.	Intact lip and mucosal tissue. Redness and/or bruising on the lip or mucosal tissue. No visible skin loss, destruction, ulceration, or blisters.
Stage 2	Destruction and differentiation of buccal mucosa, as manifested by blisters, soft coagulum, or clotting on mucosal tissue; superficial loss of nonkeratinised epithelial tissue; or damage to epidermal and dermal layers of the corners of the mouth, without evidence of damage to underlying fascia.	Visible superficial skin loss to the lips and/or mucosal tissue. Injury may have a blister-like appearance.
Stage 3	Loss of mucosa and submucosal tissue as evidenced by damage to/exposure of the fascia and underlying muscle in the lips or corners of the mouth.	Full-thickness loss of tissue in the lip or mucosal tissue. This may present as an open wound or scab. OR damage to both the inner and outer mucosal tissue in the same anatomical location.

Note. The original ROMPIS tool is reprinted from the study by Reaper et al. (2017).
M-ROMPIS, modified Reaper Oral Mucosa Pressure Injury Scale; ROMPIS, Reaper Oral Mucosa Pressure Injury Scale.

published studies with Creative Commons licensing that allowed reproduction for research purposes. The selected images were sourced by the lead investigator and reviewed by the expert ICU panel of nurses. When the final 12 images were selected, they were independently reviewed by two clinical nurse consultants with expertise in the management of acute and chronic wounds to ensure they accurately represented the relevant assigned mucosal PI category aligned with the assessment tools. There were two images of stage 1, six of stage 2, and four of stage 3. Participants were not informed that there were two versions of the mucosal pressure injury assessment tool.

2.4. Data analysis

Survey data were extracted from REDCap® in Excel® format and imported into IBM SPSS® Statistics, version 27.0 (Armonk, NY: IBM Corp.). Continuous data were analysed using descriptive statistics: mean (*M*) and standard deviation (*SD*) or median (*Med*) and interquartile range (IQR) according to whether distributions were normal. Frequency (n) and proportion (%) were calculated for categorical variables and each response category: strongly agree, agree, neither agree nor disagree, disagree, strongly disagree. Qualitative open-ended response data were analysed using the "Framework" method of qualitative data analysis.[19]

To assess the IRR of the ROMPIS and M-ROMPIS, a variable was created to note if the participant response was correct or incorrect for each image. A rating was correct if the PI was assigned the same severity criterion assigned by the panel of expert ICU nurses and incorrect if it did not. The decision was made a *priori* to exclude cases with incomplete response data from IRR analyses as it would not be possible to determine if the response was missing because of an interruption or because a decision could not be made regarding staging. The sum of correct responses for the complete set of images was used as a continuous dependent variable in analyses to describe overall scores for each tool using a two-tailed independent-group t-test or nonparametric alternative. The maximum possible score was 12. Percent agreement, the reliability statistic obtained by dividing the number of correct ratings by the total number of ratings,[20] was calculated for each tool. Gold standard assessment would equate to 100% of responses being correct, so percent agreement provided an additional reference point for the comparison of accuracy.

Fleiss' kappa (κ) is a measure of inter-rater agreement used in situations where there are two or more raters and the response variable is a categorical variable.[21] Cohen's kappa coefficient was used to assess how strong the level of agreement was between raters, where a value of <0.20 is poor, 0.21 to 0.40 is fair, 0.41 to 0.60 is moderate, 0.61 to 0.80 is good, and 0.81 to 1.00 is very good.[22] There are six assumptions required to be met when using Fleiss' kappa: (i) a categorical response variable; (ii) response categories are mutually exclusive; (iii) all raters assess the same number of categories; (iv) the raters are similar; (v) the raters are independent; and (vi) targets are randomly sampled from the population. In this study, assumption 6 was violated because good-quality images of mucosal PIs were challenging to locate. In addition, there was a need to retain a finite set of images to assess IRR as a larger number of images would potentially reduce power to detect agreement and conversely increase sample size requirements.

To strengthen confidence in the findings in the context of not meeting all six assumptions for Fleiss' kappa, intraclass correlation coefficients (ICCs) were calculated. Findings from this test provide a composite of intrarater and inter-rater variability. This study used a two-way mixed-effects ICC model where participant effects (rating applied to images) were random and measures effects (images and scale items) were fixed. 'Single measures' results provide an indication of intrarater reliability that enables results to reflect reliability of specific raters, and 'Average measures' test reliability between raters and thus reflect inter-rater reliability. Absolute agreement, which focuses on raters assigning the same score to the same subject[23] or in this instance image, was selected as the basis of model output. Cronbach's alpha was used to assess the strength of the consistency in the findings. A statistician was consulted to ensure data analyses were appropriate and accurate.

3. Results

The survey was distributed to 468 (100%) ICU nurses at both sites with a response rate of 20.9% (n = 98). Of the respondents, 88.8% (n = 87/98) completed section 1 and 2 and 73.5% (n = 72/98) completed section 3 IRR measures. Nurses were on average 38.3 (standard deviation [SD] = 11.3) years of age; most were female (n = 72/87, 82.8%), had been a registered nurse for 12 years (IQR = 8–18), and worked in the ICU for 9 years (IQR = 8–18). The majority had completed a postgraduate qualification (n = 68/87, 78.2%), and almost a third reported no additional PI-specific education (n = 24/87, 27.5%). Of the 87 (88.8%) participants who were randomised to either the ROMPIS or M-ROMPIS, there were 36 (36.7%) complete responses for the ROMPIS and 36 (36.7%) for the M-ROMPIS. There were no differences between nurses who did and did not complete IRR assessments in gender (χ^2 0.33 (2), $p = 0.85$), age (Z −0.61, $p = 0.54$), years as a registered nurse (Z −0.72, $p = 0.47$), years practicing in the ICU (Z −0.74, $p = 0.46$), or level of education (χ^2 2.31 (5), $p = 0.81$).

Similarly, when comparing nurses who completed the ROMPIS (n = 36) and M-ROMPIS (n = 36), there were no differences in gender (χ^2 1.48 (2), $p = 0.48$), age (Z −1.06, $p = 0.29$), years as a registered nurse (Z −1.19, $p = 0.23$), years practicing in the ICU (Z −1.51, $p = 0.13$), or level of education (χ^2 4.61 (5), $p = 0.46$). Total

scores for each tool were similar between participants completing the ROMPIS with an average of 8.4 (SD = 1.7, IQR = 6) correct responses and those completing the M-ROMPIS 8.7 (SD = 1.5, IQR = 6; t 0.89 [70], $p = 0.37$) from a total possible score of 12. The lowest number of correct responses was 5 (n = 3, 4.2%), and the highest was 11 (n = 7, 9.7%). No rater provided a correct response to all 12 questions. When comparing responses according to stages, stage 3 had the highest rate of correct responses, followed by stage 2 and then stage 1 as illustrated in Table 2.

Percentage agreement was calculated for each tool. A total of 36 participants categorising 12 images provided 432 possible response options for the ROMPIS. Of these, 300 were correctly identified; thus, the percentage agreement was 69.4% (300/432). Fleiss' kappa ($\kappa = 0.30$, 95% CI [0.28, 0.33], z 26.5, $p < 0.001$) indicated overall agreement was fair. Correct staging had a good level of agreement ($\kappa = 0.79$) in contrast to incorrect ($\kappa = 0.52$) that had moderate agreement. The stronger the level of agreement for the correct staging category was, the weaker the level of agreement became for incorrect staging (Table 2). The ICC calculation based on single measures (intrarater reliability) and absolute agreement, in a two-way mixed-effects model, was also fair for the ROMPIS (0.33, 95% CI [0.18, 0.59], F 18.8 (11 df), $p < 0.001$) that indicates individual raters results were inconsistent. Output for the average measures was 0.94 (95% CI [0.89, 0.98], F 18.8 (11 df), $p < 0.001$) that suggests that raters in this sample were consistently inconsistent in their ratings. Cronbach's alpha of 0.95 indicated a high degree of internal consistency for the ICC results.

In contrast, findings related to the M-ROMPIS IRR testing were superior, but still only acceptable, with 324 of 432 responses being correct and thus a percentage agreement of 75.0%. Fleiss' kappa ($\kappa = 0.29$, 95% CI [0.26, 0.31], z 25.0, $p < 0.001$) indicated overall agreement was again fair. Correct staging had a very good level of agreement ($\kappa = 0.82$) in contrast to incorrect ($\kappa = 0.47$). Not unlike the ROMPIS, the stronger the level of agreement for the correct staging category was, the weaker the level of agreement became for incorrect staging (Table 2). The ICC calculation based on single measures (intrarater reliability) and absolute agreement, in a two-way mixed-effects model, was also fair (0.31, 95% CI [0.17, 0.57], F 17.6 (11 df), $p < 0.001$) for the M-ROMPIS. Output for the average measures was 0.94 (95% CI [0.88, 0.98], F 17.6 (11 df), $p < 0.001$) that suggests that M-ROMPIS raters were also consistently inconsistent in their ratings. Cronbach's alpha for the M-ROMPIS ICC results was also 0.94, indicating a high degree of internal consistency.

Nurses agreed or strongly agreed that all patients are at a risk of developing mucosal PI (n = 61, 70.1%), that ICU patients are at a greater risk than those in lower acuity care (n = 79, 90.8%), and that continuous nursing assessment provides an accurate account of mucosal PI risk (n = 72, 82.8%). Interestingly, just over a third of nurses (n = 29, 33.3%) believed their clinical judgement was more effective than a risk assessment tool, and 64.1% (n = 41) agreed mucosal PI are avoidable, 81.6% agreed mucosal PI prevention is a routine part of their practice (n = 71), 62% agreed mucosal PI risk assessment should be done more than once per shift (n = 54), 63.2% agreed mucosal PI prevention is a high priority in ICU (n = 55), and 71.3% agreed mucosal PI is a serious hospital-acquired complication (n = 62). Only 10 (11.5%) nurses thought mucosal PI prevention was time-consuming, and 21 (24.1%) thought it was not as common as it used to be.

Participants were asked to select from a list of known risk factors for PI, those factors that they thought increased risk of mucosal PI in ICU patients (Table 3). Medical devices (n = 85, 97.7%) and poor nutritional status (n = 80, 92%) were factors with the highest level of agreement. In contrast, nurses perceived low (n = 34, 39.1%) or high (n = 9, 10.3%) blood albumin level and anaemia (n = 34, 39.1%) as risk factors. Additional risk factors identified by participants were related to the securement strategy (n = 5, 5.7%), oral hygiene (n = 8, 9.2%), and side effects of toxic medications, for example, mouth ulcers (n = 6, 6.9%). The frequency of selected prevention strategies is illustrated in Table 4. Repositioning of securement devices (n = 81, 93.1%), regular skin assessment (n = 81, 93.1%), and removal of unnecessary medical devices (n = 79, 90.8%) were the most common strategies selected for mucosal PI prevention.

Most (n = 68, 78.2%) participants provided one or more responses to a free-text question regarding barriers to mucosal PI prevention. Three themes were identified from qualitative analysis: first, staffing, workload, and organisational; second, education and knowledge; and lastly, difficulty or inability to reposition medical devices (Table 5). The majority of participant responses were related to the staffing, workload, and organisational theme. The most frequent response in this theme was lack of time and workload (n = 47, 54% of total responses). Similarly, most (n = 65, 75%) participants responded to the free-text question regarding enabling factors for improved mucosal PI prevention. Analysis of these responses identified three themes: first, measures to aid communication and documentation; second, improved staffing; and finally, additional education (Table 5).

4. Discussion

The primary aim of this study was to determine if modified descriptors for mucosal PI improved inter-rater agreement in mucosal PI assessment. The original ROMPIS descriptors by Reaper et al.[13] indicated IRR ($\propto = 0.31$; 95% CI [0.20–0.40]) was fair when tested on a population of ICU nurses. In this study, agreement was

Table 2
Level of agreement for each mucosal pressure injury stage according to the ROMPIS and M-ROMPIS.

Stage	Image	Correct	ROMPIS (N = 36)							
		n	%	κ	κy	κn	95% CI		z	p
1	4, 10	31/72	43.1	0.26	0.58	0.68	0.21	0.32	9.36	<0.001
2	1, 2, 6, 7, 9, 12	131/216	60.6	0.13	0.66	0.47	0.09	0.16	7.81	<0.001
3	3, 5, 8, 11	138/144	95.8	−0.01	0.96	0.04	−0.04	0.03	−0.19	0.85
			M-ROMPIS (N = 36)							
		n	%	κ	κy	κn	95% CI		z	p
1	4, 10	40/72	55.5	0.52	0.78	0.73	0.46	0.57	18.3	<0.001
2	1, 2, 6, 7, 9, 12	144/216	66.7	0.06	0.69	0.37	0.03	0.09	3.67	<0.001
3	3, 5, 8, 11	140/144	97.2	−0.01	0.97	0.01	−0.05	0.02	−0.69	0.48

NB: Data shown as the number (n) and percentage (%);
CI, confidence interval; *κ, overall agreement; κn, agreement incorrect*; *κy, agreement correct*; M-ROMPIS, modified Reaper Oral Mucosa Pressure Injury Scale; ROMPIS, Reaper Oral Mucosa Pressure Injury Scale.

Table 3
Risk factors for MPI.

Risk factor	n	%
Medical devices	85	97.7
Duration of pressure on mucosal membranes	83	95.4
Poor nutrition	80	92.0
Dry mucosa	76	87.4
Mechanical ventilation	72	82.8
Comorbidities	71	81.6
Oedema	70	80.5
Friction	67	77.0
Sedative medications	64	73.6
Poor sensory perception	63	72.4
Vasoactive medications	61	70.1
Immobility	59	67.8
Haemodynamically unstable	56	64.4
Noninvasive ventilation	50	57.5
Low body mass index	47	54.0
Anaemia	34	39.1
Low blood albumin	34	39.1
Analgesia	26	29.9
High blood albumin	9	10.3

NB: Data shown as the number (n) and percentage (%). Percentages do not add up to 100%.

similar for both the ROMPIS ($\kappa = 0.30$) and the M-ROMPIS ($\kappa = 0.29$). Overall agreement for each tool was fair because of moderate agreement in correct and incorrect ratings for stage 1 and stage 2 images. Correct ratings were very good for stage 3 images irrespective of the tool used. Intraclass correlations to assess the reliability of agreement indicated raters were reliably inconsistent in their ratings. These findings align with literature examining PI using the standard international classification system[6] where IRR is commonly reported to be between 0.39[24] and 0.59.[25] Despite this, the accepted international PI staging system has been widely adopted and provides global gold standard measures for PI classification, reporting, and monitoring. Percentage agreement for the international staging system is on average 68.5%[26] which is comparable with ROMPIS percentage agreement (69.7%). Percentage agreement for the M-ROMPIS was 76%, and for nine of 10 images used in this study, correct classification was higher when M-ROMPIS descriptors were used than when original ROMPIS descriptors were used. Classification was identical for two of the 10 images: one, a stage 2, and one, a stage 3 mucosal PI. Findings from this study indicate that the M-ROMPIS has marginally better agreement than the original ROMPIS, but neither set of descriptors is particularly reliable unless there is a stage 3 mucosal PI evident.

Nurses' perceptions of mucosal PI were consistent with findings from previous research conducted in the context of PI occurring on human skin, which predominantly reports nurses having proactive attitudes towards PI prevention and management.[15,27] Participants cited unavoidable reasons for their inability to reposition medical devices exerting pressure on mucosal membranes, such as facial oedema, and avoidable reasons such as nasogastric tubes secured with stitches. This indicates that although nurses are the primary providers of mucosal PI prevention, education regarding prevention and treatment strategies could involve the wider multidisciplinary team. It appears that there is also a need to improve ICU nurses' knowledge of mucosal PI prevention and management, as most nurses who undertook this survey had not participated in mucosal PI–specific education. Additional focused education may in turn lead to the implementation of appropriate strategies to manage factors considered avoidable causes of mucosal PI.

Table 4
Prevention strategies for MPI.

Prevention strategy	n	%
Repositioning securing devices	81	93.1
Regular skin assessment	81	93.1
Removal of unnecessary medical devices	79	90.8
Nutritional support	72	82.8
Skin moisturising/hydration	67	77.0
Risk assessment	67	77.0
Padding under medical devices	59	67.8
Massage	26	29.9

NB: Data shown as number (n) and percentage (%). Percentages do not add up to 100%.

There are several limitations associated with the design of this study. Electronic surveys generally have a low response rate.[28] This was anticipated, and steps were taken to overcome this by including reminders, site champions, and release at two sites. It is likely that the impact of COVID-19 may have also influenced response rate as release of the survey coincided with a period of higher ICU patient acuity. Response rate impacts sample size, so there is a degree of uncertainty regarding whether a higher number of participants would change level of agreement or the inter-rater or intrarater findings. Findings from this work do, however, provide a basis from which sample size for future research can be calculated.

Nurses self-selected to participate, so this may not constitute a representative sample of the population.[29] Nurses who chose to participate may have had more interest in mucosal PI prevention and management. It is, however, unlikely that a greater interest would have affected internal validity as almost a third of participants indicated they had no further education in PI prevention and management and the survey was available to any nurse in the ICU rather than targeting those with mucosal PI expertise. The structure of the survey may have influenced the rate of IRR completion as the grading tools appeared last. Rotating the order of sections 2 and 3 of the survey might have increased response and completion rates. Reaper et al.[13] reported limited consensus with the 'unstageable' category and postulated that this category provided a mechanism to avoid or negate having to determine PI staging. In this study, unstageable was removed from both the ROMPIS and M-ROMPIS and images were specifically chosen from each of the three stages to focus responses on categorisation of mucosal PI. This approach does limit assessment outcome options, justified by the intent of staging selection for the purposes of IRR and measurement of agreement levels.

Although the IRR results are consistent with previous research, poor levels of agreement are not unusual in research that has explored PI risk assessment and staging. The CIs in this study were relatively narrow, which is indicative of precisely incorrect agreement. Including multiple images for each mucosal PI stage may have contributed to the variability in participant responses, irrespective of which tool was used to assess mucosal PI in this study. When the study was designed, the intent was to emulate real-life assessment that rarely involves patients presenting with PI that are similar in appearance. It was anticipated that descriptors would provide adequate information to correctly stage each image. The findings, however, have revealed that unless the mucosal PI was most severe and consistent with a stage 3 injury, raters were neither consistently correct nor incorrect. We recommend refining future study design by systematically testing instruments with fewer or even only a single image in each category and having a set number of raters use the same instruments for risk assessment. This will simplify analyses, the interpretation of findings, and the ability to determine optimal descriptors for mucosal PI.

As with the original study by Reaper et al.,[13] images were used, and this may have influenced individual assessment, particularly in the context of almost two-thirds of participants thinking their own judgement is less effective than a risk assessment tool. Evidence indicates that clinical judgement is as useful as risk assessment tools to prevent PI,[30] but the impact of assessment using imagery

Table 5
Barriers and enablers to MPI prevention.

Barriers		
Staffing, workload, and organisational (n = 47, 54%)	Education and knowledge (n = 21, 24%)	Difficulty or inability to reposition medical devices (n = 19, 22%)
Time Insufficient staffing Skill mix Lack of resources Lack of compliance with protocol, poor nursing care, effort	Lack of education and experience Conflicting information about what MPI entails MPI prevention not a high priority	Nasogastric tube attached to patient with stitches Patient in prone position Significant facial oedema Difficult airway leading to apprehension about repositioning Patient agitation or noncompliance with care Coagulopathy where device movement may disturb clots and cause blood loss Haemodynamically unstable Difficult to open/see inside patient's mouth
Enablers		
Assessment, communication, and documentation (n = 38, 44%)	**Staffing, workload, and organisational (n = 31, 36%)**	**Education and knowledge (n = 17, 20%)**
Regular oral care, assessment and repositioning of medical devices Standardised approach with clear guidance and requirements for MPI prevention and documentation Risk assessment tools Ongoing evaluation of practice	Sufficient staffing More time/less time-consuming MPI prevention strategies Improved budget (unspecified) Lower acuity of patients Regular review of the requirement for medical devices Leadership from senior staff highlighting MPI importance and assisting with implementing preventative strategies More or better access to resources for MPI prevention	Improved education In-service Clinical champions

NB: Data shown as the number (n) and percentage (%).

rather than real patients is not well understood despite this approach being common.[9] Prospective studies to explore the superiority of expert clinical judgement in contrast to risk assessment tools and determine which characteristics impact on expert judgement are warranted.

The M-ROMPIS may have slightly better agreement than the ROMPIS, but whether it is a reliable tool for assessing the severity of mucosal PI requires additional exploration. A systematic review and meta-analysis of PI assessment scales for ICU patients confirmed variability in evaluation criteria impacts the clinical utility of available options. In addition, despite widespread use, there is a lack of verification of tools used for PI risk assessment in the ICU setting.[31] Future research must incorporate robust design to revise and refine existing mucosal PI classification options. Using instruments in real time also accommodates the assessment of content and construct validity. In this study, mucosal PI descriptors were modified in response to a panel of experts agreeing that the existing terminology did not align with vernacular of ICU clinicians. Systematic sampling, which involves the selection of cases drawn from a population at fixed intervals,[32] is a sampling technique that would minimise the potential for selection bias providing a representative sample. Recruitment strategies should also accommodate all healthcare disciplines responsible for the ICU patient. Prospective cohort studies would be beneficial for two key reasons: the capacity to incorporate participant feedback on the nature of descriptors and their relationship to mucosal PI severity in real time and the opportunity to monitor the evolution of mucosal PIs as they emerge. Evidence that describes the trajectory of mucosal PI from initial insult to resolution has not been located.

5. Conclusion

ICU patients are frequently at a high risk of developing mucosal PI from essential therapeutic equipment. This study built on the work of Reaper et al.[13] by testing the reliability of the original ROMPIS descriptors in a different population of end users. In addition, the M-ROMPIS tool was developed and tested to determine if using language consistent with commonly used PI assessment terms would improve reliability. Although there are limitations, findings indicate the M-ROMPIS may have marginally better agreement than the ROMPIS. Insights are also provided regarding ICU nurses' perceptions of mucosal PI and its prevention. Participants claim a proactive attitude towards mucosal PI management and its prevention but require further education to strengthen their perceived capacity for risk assessment and clinical judgement. This work provides a foundation for future benchmarking and a platform from which further research to refine and test descriptors specific to mucosal PI can be generated.

Conflict of interest

The authors declare no conflict of interest.

Funding

None declared.

CRediT authorship contribution statement

Simone Fitzgerald: Conceptualisation, Methodology, Investigation, Data curation, Writing – original draft. **Lauren McTier:** Conceptualisation, Methodology, Writing – review & editing of draft. **Christina Whitehead:** Methodology, Investigation, Data curation. **Kristy Masters:** Methodology, Investigation, Data curation. **Rochelle Wynne:** Conceptualisation, Methodology, Investigation, Data curation, Formal analysis, Writing – original draft.

References

[1] Nguyen K, Chaboyer W, Whitty J. Pressure injury in Australian public hospitals: a cost-of-illness study. Aust Health Rev 2015;39:329–36.

[2] Gaspar S, Peralta M, Marques A, Budri A, Gaspar de Matos M. Effectiveness of hospital-acquired pressure ulcers prevention: a systematic review. Int Wound J 2019;16:1087–102.
[3] Team V, Tuck M, Reeves J, Way M, Enticott J, Evans S, et al. Pressure injury data in Australian acute care settings: a comparison of three data sets. Int Wound J 2020;17:578–86.
[4] Haesler E, Carville K, Haesler P. Priority issues for pressure injury research: an Australian consensus study. Res Nurs Health 2018;41:355–68.
[5] Australian Wound Management Association. Pan Pacific clinical practice guideline for the prevention and management of pressure injury. Osborne Park, WA: Cambridge Media; 2012.
[6] Edsberg LE, Black JM, Goldberg M, McNichol L, Moore L, Sieggreen M. Revised national pressure ulcer advisory panel pressure injury staging system: revised pressure injury staging system. J Wound, Ostomy Cont Nurs 2016;43(6):585–97. https://doi.org/10.1097/won.0000000000000281.
[7] National Pressure Injury Advisory Panel. Mucosal pressure ulcers: a NPUAP position statement. 2008. Retrieved from, https://www.npuap.org/wp-content/uploads/2012/01/Mucosal_Pressure_Ulcer_Position_Statement_final.pdf.
[8] McCann S, Huether S. Structure, function, and disorders of the integument. In: McCance K, Huether S, editors. Pathophysiology: the biologic basis for disease in adults and children. St. Louis: Elsevier; 2019. p. 1496–529.
[9] Coyer F, Miles S, Gosley S, Fulbrook P, Sketcher-Baker K, Cook JL, et al. Pressure injury prevalence in intensive care versus non-intensive care patients: a state-wide comparison. Aust Crit Care 2017;30(5):244–50.
[10] Black JM, Cuddigan JE, Walko MA, Didier LA, Lander MJ, Kelpe MR. Medical device related pressure ulcers in hospitalized patients. Int Wound J 2010;7:358–65.
[11] Coyer F, Stotts NA, Schmied Blackman V. A prospective window into medical device-related pressure ulcers in intensive care. Int Wound J 2014;11:656–64.
[12] Fullbrook P, Lovegrove J, Miles S, Isaqi B. Systematic review: incidence and prevalence of mucous membrane pressure injury in adults admitted to acute hospital settings. Int Wound J 2021. https://doi.org/10.1111/iwj.13629. Epub.
[13] Reaper S, Green C, Gupta S, Tiruvoipati R. Inter-rater reliability of the Reaper Oral Mucosa Pressure Injury Scale (ROMPIS): a novel scale for the assessment of the severity of pressure injuries to the mouth and oral mucosa. Aust Crit Care 2017;30:167–71.
[14] Avsar P, Patton D, O'Connor T, Moore Z. Do we still need to assess nurses' attitudes towards pressure ulcer prevention? A systematic review. J Wound Care 2019;28:795–806.
[15] Barakat-Johnson M, Barnett C, Wand T, White K. Knowledge and attitudes of nurses towards pressure injury prevention. J Wound, Ostomy Cont Nurs 2018;45:233–7.
[16] McPeake J, Bateson M, O'Neill A. Electronic surveys: how to maximise success. Nurse Res 2014;21(3):24–6.
[17] Kottner J, Raeder K, Halfens R, Dassen T. A systematic review of interrater reliability of pressure ulcer classification systems. J Clin Nurs 2010;18:315–36.
[18] Moore Z, Price P. Nurses' attitudes, behaviours and perceived barriers towards pressure ulcer prevention. J Clin Nurs 2004;13:942–51.
[19] Ritchie J, Spencer L. Qualitative data research for applied policy research. In: Bryman A, Burgess R, editors. Analyzing qualitative data. London: Routledge; 1994. p. 173–94.
[20] Bryer J. Relationship between intraclass correlation (ICC) and percent agreement. 23rd August, 2021. Retrieved from, irrsim.bryer.org/articles/IRRsim.html.
[21] Laerd Statistics. Fleiss' kappa using SPSS statistics. Available from: https://statistics.laerd.com/spss-tutorials/fleiss-kappa-in-spss-statistics.php.
[22] Landis J, Koch G. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.
[23] Koo T, Li M. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine 2016;15:155–63.
[24] Jesada E, Warren J, Goodman D, Iliuta R, Thurkauf G, McLaughlin M, et al. Staging and defining characteristics of pressure ulcers using photographs by staff nurses in acute care settings. J Wound, Ostomy Cont Nurs 2013;40:150–6.
[25] Bergquist-Beringer S, Gajewski B, Dunlon N, Klaus S. The reliability of the national database of nursing quality indicators: a triangulation approach. J Nurs Care Qual 2011;26:292–301.
[26] Ham W, Schoonhoven L, Schuurmans M, Veugelers R, Leenon L. Pressure area education improves interrater reliability, identification, and classification skills by emergency nurses and physicians. J Emerg Nurs 2015;41:43–51.
[27] Coyer F, Cook JL, Doubrovsky A, Campbell J, Vann A, McNamara G. Understanding contextual barriers and enablers to pressure injury prevention practice in an Australian intensive care unit: an exploratory study. Aust Crit Care 2019;32(2):122–30.
[28] Cope D. Using electronic surveys in nursing research. Oncol Nurs Forum 2014;41:681–2.
[29] Osborne S, Schneider Z. Quantitative data collection and study validity. In: Schneider Z, Whitehead D, LoBiondo-Wood G, Haber J, editors. Nursing and midwifery research: methods and appraisal for evidence-based practice. Chatswood: Elsevier Australia; 2013. p. 200–17.
[30] Webster J, Coleman K, Mudge A, Marquart L, Gardner G, Stankiewicz M, et al. Pressure ulcers: effectiveness of risk-assessment tools. A randomised controlled trial (the ULCER trial). BMJ Qual Saf 2011;20(4):297–306.
[31] Zhang Y, Zhuang Y, Shen J, Chen X, Wen Q, Jiang Q, et al. Value of pressure injury assessment scales for patient in the intensive care unit: systematic review and diagnostic test accuracy meta-analysis. Intensive Crit Care Nurs 2021;64:103009. https://doi.org/10.1016/j.iccn.2020.103009. Epub.
[32] Shields L, Watson R. Common quantitative methods. In: Schneider Z, Whitehead D, LoBiondo-Wood G, Haber J, editors. Nursing and midwifery research: methods and appraisal for evidence-based practice. Chatswood: Elsevier Australia; 2013. p. 161–83.