Risk assessment scales for pressure ulcers: 
A methodological review

Panos Papanikolaou\textsuperscript{a,}\textsuperscript{*}, Patricia Lyne\textsuperscript{a}, Denis Anthony\textsuperscript{b}

\textsuperscript{a}Nursing, Health and Social Care Research Centre, School of Nursing and Midwifery Studies, University of Cardiff, 35-43 Newport Road, Cardiff CF24 0AB, UK

\textsuperscript{b}The Mary Seacole Research Centre, School of Nursing and Midwifery, De Montfort University, Charles Frears Campus, 266 London Road, Leicester LE2 1RQ, UK

Received 17 January 2005; received in revised form 19 December 2005; accepted 10 January 2006

Abstract

Much is written about risk-assessment scales (RASs) for pressure ulcers (PU) and their properties demonstrating that they are of limited value. Less is known about the reasons for these limitations and the scope for improvement. This review examines issues such as structure and scoring for the Norton, Waterlow and Braden scales, showing that the equal-weighting technique behind the current RASs is too simplistic and leads to limitations. It concludes that properly trained, experienced nurses should conduct PU risk assessments, whilst more robust data-driven RASs should be developed using the differential weighting scoring method together with advanced statistical techniques.

Keywords: Risk assessment scales; Validity; Reliability; Pressure ulcer

What is already known about the topic?

\begin{itemize}
  \item The most commonly used risk assessment scales, the Norton, the Waterlow and the Braden scale, have insufficient predictive validity and poor reliability.
  \item The Norton and Waterlow risk assessment scales contain some risk components, for which the description is unclear or ambiguous and difficult to understand.
  \item The critical cut-off score for the Norton, the Waterlow and the Braden scale is disputable.
\end{itemize}

What this paper adds

\begin{itemize}
  \item Formal mathematical language is used to show the limitations of the structures underlying the Norton, the Waterlow and the Braden scales.
  \item The equal-weighting scoring technique behind these (and other) scales has limited utility.
  \item The differential weighting scoring technique together with advanced statistical methods and large data-sets should be used to develop data-driven and more robust risk assessment scales. Until this happens, clinical judgement should be used by experienced nurses to assess the risk of pressure ulcer occurrence.
\end{itemize}

1. Background and Introduction

Current estimates of pressure ulcer (PU) epidemiology vary considerably. The estimated point prevalence rate...
among hospitalised adult patients in the UK ranges from 9.6% (Bond 1993) to 11.90% (Birchall 2001), with an incidence rate between 12.0% for surgical patients (Kemp et al., 1990) and 22.07% for elderly people (Papanikolaou et al., 2002). For the community setting, point estimates of the prevalence rate are within the region 4.4% (Hallett, 1996) and 6.8% (Preston, 1991) for adults and children. PUs still remain a common complication throughout the world. A recently completed survey for USA and Canada (Kaltenthaler et al., 2001) showed that prevalence and incidence rates have varied significantly across different care environments. In acute care, the prevalence rate ranged between 4.7% and 29.7%, whilst the incidence rate was between 8.5% over a 1–4-week period to 13.4% for a maximum of 2 weeks. For the community setting, the prevalence rate ranged from 19.2–29.0%, and the incidence rate varied from 0% over a six-month period to 16.5% (the researchers did not state the time period). For nursing homes, the prevalence rate was within the region 15.33–20.7%, whilst the incidence rate varied from 6.2% over 6 months to 13.2% over 12 months. This condition is of considerable significance for those affected, who are often older and suffering severe ill health. It causes distress, which can be severe, to patients and their families.

Pressure ulcers demand the use of resources (i.e. disposables, equipment and nursing time) from the health care system. In the UK National Health Service (NHS) the annual cost of treating PU was computed to range between £1.4 and £2.1 billion at 2000 prices, whilst the treatment cost per patient, per care-episode was estimated to increase from £1064 for a PU grade 1, to £7750 for a PU grade 4, assuming that in both cases healing is uncomplicated by adverse events such as wound colonisation (Bennett et al., 2004). Furthermore, healthcare organisations are exposed to the risk of litigation in cases where pressure damage occurs supposedly as a result of clinical negligence. Tingle (1997) found that settlement values ranging between £4500 and £12,500 had been awarded to claims associated with PU development due to insufficient care. Present figures are likely to be much higher.

For these and other reasons, PU prevention is of great importance. An essential component of preventive strategies is the assessment of risk of PU development in the individual. This forms the subject of the present paper. Whilst several authors have investigated the properties of risk assessment tools, this is, to the best of our knowledge, the first attempt to use a formal mathematical approach for this purpose. This enables us to be very precise about the limitations and scope for improvement. The paper arose during a study of decision-making by nurses responsible for the planning of PU prevention in the community (Papanikolaou et al., 2004), which required a review of the literature relating to PU risk assessment.

1.1. Literature review

If PUs are to be prevented, risk assessment of an individual’s chances of PU occurrence must be undertaken before any prevention plans are put in place. A comprehensive and detailed risk-assessment based on the essential principles of wound healing for every single patient would involve bringing together a considerable quantity of knowledge, and would be virtually impossible in practice. Hence, the interest in a risk assessment scale (RAS) that can be employed to produce a quick assessment, thus acting as a ‘shorthand’ approach to assessment. Current guidelines (NICE, 2005) emphasise that RASs should be used as an adjunct to clinical judgement and should not replace it. Nevertheless, if they are to be used at all, it is important that the scores produced should be as reliable as possible.

Such a scale produces assessments of a range of extrinsic and intrinsic factors (for example, mobility, incontinence) generally believed to contribute to pressure ulcer occurrence. A numerical value is assigned to each of those components according to the patient’s condition. The values achieved within each risk factor are summed to derive the total score. This is usually compared to a critical threshold value, thereby indicating the level of vulnerability to PU development. The relationship between the total score obtained by the RAS and the risk-level of ulcer occurrence can be positive or negative, depending on the scoring system. Thus for some scales, greater vulnerability is indicated by high scores and for others by low scores.

The total score is frequently used (in conjunction with clinical judgement) to:

1. ensure that limited resources (such as nursing labour, dressings, and pressure-relieving equipment) are allocated in an effective and cost-effective manner (Cullum et al., 1995);
2. support clinical decisions (Maylor and Roberts, 1999);
3. act as case-mix adjuster to assist in making objective comparisons of PU incidence measurements between cross-sectional units (e.g. wards, clinics, hospitals) or over time (e.g. years). Use of the score provides the common denominator to allow comparisons of PU incidence between clinical areas or time units (Cullum et al., 1995; Maylor and Roberts, 1999);
4. contribute to the development of risk assessment protocols (Deeks and Dealey, 1996);
5. argue against litigation (Flanagan, 1993, Smith et al., 1995).

Since the early 1960s a variety of RAS have been developed with over 20 such scales currently described in the world-wide literature (Torra, 1998). These have evolved from the observations of health carers about the
kind of person appearing to be vulnerable to pressure damage, or are modifications of existing scales in order to deal with the needs of specific patient-groups. For example, McCormack developed a risk scale particularly designed to handle the assessment of older people (Gould, 2001), whilst Perneger et al. (2002) combined multivariate analysis with the Norton–Braden risk factors to develop a risk calculator for use in the acute hospital setting.

Numerous studies have reviewed these scales. The bulk of this work has looked at the predictive validity (i.e. sensitivity, specificity and predictive values) of these scales (Jalali and Rezaie, 2005; Defloor and Grypdonck, 2004; Brown, 2004, Rycroft-Malone, 2000; Pang and Wong, 1998; Cullum et al., 1995), and has revealed that the estimated predictive validity varies considerably within the same scale and across different scales, when used across different clinical care settings and/or patient populations. Other studies have assessed the reliability of one or another risk instrument (Kelly, 2005, Rycroft-Malone, 2000; Hale et al., 2000, Edwards, 1996) and found that the available risk assessment scales suffer from poor reliability (Rycroft-Malone, 2000). Yet little work has been undertaken to investigate the underlying structure of the scales and the appropriateness of the scoring method. This review aims to enrich our understanding of this issue and ways in which improvement might be secured. It focuses upon the most commonly used general purpose RAS; the Norton, Waterlow and Braden scales.

1.2. Aims and objectives

The overarching aim of this paper is to demonstrate the current state of knowledge of the RAS used to identify those patients at PU risk. Its specific aims are:

- to critically review the most commonly used PU risk instruments in UK and elsewhere: (1) The Norton score; (2) The Waterlow scale; (3) The Braden scale;
- to present the structure of these scales using formal mathematical language;
- to assess the appropriateness of the underlying scoring technique used to compare the total risk score;
- to examine recent research attempts to build sound and empirically driven RAS using rigorous scientific methods.

2. The Norton scale

Norton and her colleagues were pioneers in the largely unexplored area of PU formation (Flanagan, 1997). In the early 1960s they devised the first recognised PU risk-assessment instrument based on their research. In this scale, the total risk score is a linear function of five risk factors considered important by Norton et al. More formally,

\[ S_N = X_1 + X_2 + X_3 + X_4 + X_5, \]

where \( S_N \) is the Norton total score indicating the risk of PU development; \( X_1 \) is the general physical condition score; \( X_2 \) is the activity score; \( X_3 \) is the mobility score; \( X_4 \) is the incontinence score; and \( X_5 \) is the mental state score.

All the risk factors, \( X_1 \) through \( X_5 \), take values between four (most favourable) and one (least favourable). For this reason, a lower total score is suggestive of a higher risk of ulcer development, in other words, risk level and total Norton score are inversely related. In the seventies, the scoring system was reversed (Smith et al., 1995) so that a higher score indicates an increase in the risk of PU formation.

Norton et al. (Goldstone and Goldstone, 1982) initially considered 14 or below to be the critical cut-off point (\( S_N \leq 14 \)) at which an individual would be at risk of ulcer formation. Subsequently, Norton (1996) revised this score to 16 or below (\( S_N \leq 16 \)) as the onset of risk. The scale is easy and convenient to use, yet according to Ek and Bjurlf (1987) its use should be confined to those nurses with good understanding of its parameters and training in its use. The scale has received the following criticisms.

(1) Nutrition is not included among the risk factors (Flanagan, 1997). Its absence has been justified on the grounds that it is implicit as part of the patient’s general condition (Norton, 1996).

(2) The use of the term ‘incontinence’ to describe ‘the ability to be continent’ is likely to produce errors in assessment. Instead, using the direct term ‘continence’ and re-phrasing its constituents would have increased its clarity (Norton, 1996).

(3) A scoring classification mechanism was missing in the original version. As a result, all patients classed at risk (i.e. \( S_N \leq 14 \)) but allocated different scores (e.g. \( S_{N1} = 10, S_{N2} = 8, S_{N3} = 6 \)) are put on the same footing, thus restraining the health-care professionals’ capability to design and implement patient-specific care plans. A classification guide was added to the modified version (Gosnell, 1973).

(4) The influence of modern medicines is omitted from its structure, although certain medicines (e.g. anti-inflammatory, cytotoxics) are likely to make the tissues more friable, thus accelerating the process of PU occurrence. This information was not available to the team which constructed the original scale.

3. The Waterlow scale

The Waterlow scoring system, which emerged from a large prevalence study carried out in Southwest England in 1984, was designed to serve three purposes (Waterlow, 1988). The first was to provide an assessment of the
risk of PU damage; the second was to make recommenda-
dations for preventive measures across different risk ciri
stances and the last was to deliver a PU classification sys-
tem (Flanagan, 1997). The Waterlow Pressure Ulcer Scoring System may therefore be viewed as one that encompasses three tasks of basic nursing care: standard risk assessment, PU grading and planning preventive and treatment actions. It thus could be used across various clinical settings. The Waterlow total risk score is made up of a normal and a special risk section. It is an additive function of two sets of risk factors. More formally,

\[
S_W = Q_1 + Q_2 + Q_3 + Q_4 + Q_5 + Q_6 + Q_7 + d_1 W_1 + d_2 W_2 + d_3 W_3 + d_4 W_4,
\]

where \(S_W\) is the Waterlow total score indicating the risk of PU development; \(Q_1\) is the build/weight for height (a body mass index assessment) score; \(Q_2\) is the continence score; \(Q_3\) is the skin type score; \(Q_4\) is the mobility score; \(Q_5\) is the gender score; \(Q_6\) is the age score; \(Q_7\) is the appetite score; \(W_1\) is the tissue malnutrition score (condition-specific); \(W_2\) is the neurological deficit score (condition-specific); \(W_3\) is the major trauma/surgery score (condition-specific); \(W_4\) is the medication score (treatment-specific), and \(d_j\) is the indicates the presence or absence of the special risk in consideration. Hence, it takes on the value one if the particular risk is present \((d_j = 1)\) and zero if that risk is not there \((d_j = 0)\).

In the normal risk section, the values of factors \(Q_1, Q_2, Q_3\) and \(Q_4\) range from zero (most favourable) to three (least favourable), whilst others appear with different scoring levels. In the special risk section, all the factors are assigned different values. A positive relationship holds between total score and risk of PU formation. A total score of at least 10 \((i.e. S_W \geq 10)\) is suggestive of a patient at risk of PU occurrence. Risk categories are specified within the scale thus allowing the nursing staff to establish priorities concerning the frequency and level of care required and to allocate resources accordingly.

The scale is relatively easy to use taking up to 2 min to complete, once the users have become accustomed to its use (Waterlow, 1988). However, this has been questioned by independent studies. Banks and Bale (1992) report that community nurses felt that the scale was time-consuming and added unnecessary paperwork, thereby reducing the amount of time available for the patient’s care. Maylor and Roberts (1999) suggest that its use relies upon direct observation of the skin and the availability of information about the patient’s condition. As a result, its use becomes more difficult in the nursing home care setting, where some data may be unavailable. The following remarks have been made.

1. Pain is not treated as a separate risk factor (Flanagan, 1993). This has been defended on the basis that the patient’s experience of pain is a distinctly important facet of PU prevention, thus requiring an independent assessment. The results could then be added in the patient’s preventive care plan.

2. The ‘skin type’ risk factor is subjected to multiple scoring. The conditions ‘tissue paper’ and ‘dry skin’ are awarded the same value so that both make an equal contribution to the cumulative patient’s score. Yet the former represents a more vulnerable condition than the latter, which contradicts the logic underlying the scale’s configuration (Lyne et al., 1999).

3. The ‘build/weight for height’ risk factor, assessment for the body mass index, imposes an unnecessary upper limit to the involvement that body configuration makes to the overall score (Lyne et al., 1999). For example, a score of 3 is assigned to encompass all those with below average configuration, ranging from the slightly underweight to the extremely lean.

4. The ‘continence’ risk component fails to assign a score to the catheterised patient, who may well have experienced a degree of incontinence prior to catheterisation and therefore have suffered tissue damage. On the other hand, it is not clear whether ‘occasionally incontinent’ refers to urine or faeces and if this is accompanied by catheterisation, in which case partially continent patients can be placed at higher-than-necessary risk level, hence unnecessarily increasing the demand for preventive scarce resources (Lyne et al., 1999).

5. The ‘mobility’ risk sub-scale is subjected to a ceiling effect, inasmuch as the highest score of 5 is represented by ‘chair-bound’. This term can cover a range of situations, from the lightweight person who can readily be assisted from bed to chair, to the grossly obese person with congestive cardiac failure who habitually sleeps upright in a chair only and changes position with great difficulty (Lyne et al., 1999).

6. The medication subscale does not take into account the impact of the interaction between sedative and other medicines upon the healing process (Cook et al., 1999).

7. The ‘gender’ sub-scale places females at higher risk than males. Yet this is an indefensible and ungrounded argument. Evidence drawn from national and international work suggests that gender is not important compared with other prognostic factors, other things being equal. Margolis et al., (2002) found that the likelihood of PU occurrence increased considerably with advancing patient age among British elderly in the primary care area. However, the study failed to detect any significant link between gender and PU development. Maklebst and Mag-
nan (1994) found that there was no statistically important relationship between gender and PU risk. The study concluded that greater disease burden (as proxied by co-morbidity variables) rather than
gender per se is more likely to shape the chances of 
PU occurrence. Therefore ‘gender’ itself does not 
predispose to PU risk formation.

(8) The presence of neurological deficit attracts a score 
ranging between 4 and 6, yet different nurses may 
arbitrarily attach any of these points to the subscale, 
and thus to the total risk-score. Furthermore, 
reliability studies have shown that raters may value 
its importance differently, thus adversely affecting 
the estimated reliability coefficient. Inter rater 
reliability was found to be lowest in those sub-scales 
where there was the most scope for varied inter-
pretation (Cook et al., 1999).

The Waterlow scale can be seen as a more advanced 
and comprehensive risk-scoring system than the Norton 
scale. It includes the ‘disease and treatment specific risks’ 
section and differentiates across disparate levels of risk 
status. This allows the user to quantify the previously 
obscure concept of risk status, and furnishes preventive 
and treatment guidelines. However, the effort to increase 
the sophistication and scope of the instrument has 
introduced its own burden of ambiguity.

4. The Braden scale

The Braden PU risk-assessment tool was developed in 
the 1980s following an observation made by Braden 
(Flanagan, 1993) that while nursing staff were careful in 
turning, repositioning and caring for the skin of patients 
in nursing homes in USA, poor nutritional status was 
the principal contributing factor to ulcer formation. 
Like its predecessors, the Braden risk instrument is also 
a linear transformation of its risk factors.

\[ S_B = Z_1 + Z_2 + Z_3 + Z_4 + Z_5 + Z_6, \]

where \( S_B \) is the Braden total score indicating the risk of 
PU development; \( Z_1 \) is the Sensory perception; \( Z_2 \) is the 
Activity score; \( Z_3 \) is the Mobility score; \( Z_4 \) is the 
Nutrition score; \( Z_5 \) is the Moisture score; and \( Z_6 \) is the 
Friction and shear score.

These risk-components (i.e. the \( Z \) variables) tackle the 
two primary etiologic factors in ulcer formation:

(I) Intensity and duration of pressure; 
(II) Tissue tolerance for pressure.

The first three sub-scales: sensory perception, activity, 
and mobility, capture clinical situations predisposing the 
patient to intense and prolonged pressure. The remain-
ing three sub-scales: nutrition, moisture, and friction/ 
 shear, reflect factors adversely affecting the tissue 
tolerance for pressure (Ayello and Braden, 2001).

Like the Waterlow scale, the Braden scale contains a 
risk classification so that nursing preventive interven-
tions can be adjusted to the individual’s needs. It is 
relatively self-explanatory and convenient to use, 
provided nurses have received the necessary training 
(Smith et al., 1995). The subsequent criticisms have been 
offered.

(1) It needs detailed knowledge of the patient’s dietary 
intake, the ability to move or seek assistance and the 
frequency of the skin being moist. Accordingly, the 
patient’s assessment immediately following admis-
sion may not be highly accurate.

(2) The cut-off score has been questioned. Studies have 
been conducted across disparate care settings (acute 
care, home care, long-term institutional care) to 
determine the optimal risk-predicting score. Some 
have recommended a lower cut-off score (Langemo et 
al., 1991); others have suggested a higher cut-off point 
(Capobianco and McDonald, 1996). The debate 
about the critical cut-off score highlights the fact that 
the same critical score, in all likelihood, would not be 
applicable to every single patient care setting.

(3) Recent evidence indicates that certain Braden sub-
scales (or sub-scores) may be more important than 
others for predicting risk. For example, Baldwin and 
Ziegler (1998) demonstrated that the mobility sub-
scale was the only subscale directly related to PU 
ocurrence and a better determinant than the overall 
total score in critically injured trauma patients.

5. Pressure ulcer risk assessment: unresolved issues

5.1. Need for training

Whilst it is frequently claimed that these PU risk 
scales are simple and easy-to-use, there is some 
controversy on this point. Evidence suggests (Ek and 
Bjurulf, 1987; Banks and Bale, 1992; Smith et al., 1995; 
Maylor and Roberts, 1999) that these scales only 
provide valid measurements of the PU risk occurrence 
in the hands of adequately trained nurses.

5.2. Need for clarity

The Norton and the Waterlow scales do not provide 
descriptions of their risk components, although this need 
has been recognised by Norton (1996). The absence of 
explanatory comments carries the risk that users will 
make misjudgements when assigning numerical values to 
the sub-scales. For this reason, use of these scales 
requires informed clinical judgement and thus needs to 
be restricted to the better-trained health-care profes-
sionals.

5.3. Predictive performance

Rycroft-Malone (2000) reviewed the evidence of each 
scale’s reliability and found that the scales have not been
properly assessed so that they all suffer from poor levels of demonstrated reliability. A recent study examined the Waterlow risk instrument’s reliability and found that it suffers from poor reliability (Kelly, 2005).

In regard to predictive validity, the existing work has indicated that the scales are over-sensitive to the number of patients allocated to the ‘at-risk’ group of PU occurrence and under-sensitive to the number of patients assigned to the ‘risk-free’ group. A recent RCN publication (Bick and Stephens, 2004) has suggested that formal risk assessment activity should be used in conjunction with clinical judgement rather than in isolation, and this is reflected in the UK by the national guidelines (NICE, 2003, 2005).

5.4. Determining the critical threshold score

The initially suggested critical score for each of these risk instruments (i.e., \( S_N \leq 10, S_W \geq 10, S_B \leq 16 \)) at which tissue breakdown commences has been disputed. Clark and Farrar (1992) analysed the predictive abilities of the Norton, Waterlow and Braden scale and found that the score discriminating between risk and risk-free status did not match with the one indicated in the literature. Subsequent studies undertaken for one or another RAS have added further support. Anthony and Unsworth (1998) found that the Waterlow risk instrument would perform well for wheelchair inpatient users if the threshold was set at 15 or above. Bergquist and Frantz (2001) suggested that the Braden scale would achieve its best predictive ability for elderly adults receiving home health care, if the cut-off score was set at 19. On the basis of this evidence it is clear that the optimal cut-off point differentiating between risk and non-risk status has yet to be determined for different instruments.

Current disagreements about the appropriate cut-off score have important implications for clinical effectiveness. Health-care professionals receive confusing signals about the best available evidence when applying PU risk-assessment strategies and this may well produce variations in the appropriateness of the resultant healthcare preventive plans.

One solution would be to allow each healthcare delivery unit to set its critical threshold, based on empirical information about the needs and characteristics of the patient group in consideration. The development of context-specific critical cut-off points for different risk scales would offer greater flexibility to healthcare decision makers, thus leading to enhanced clinical effectiveness.

5.5. Scoring method

The scoring mechanism employed to discriminate between risk and non-risk status regards every single risk factor as making an equal contribution to the overall risk score. This is precisely true for the Norton scale where its sub-scales are equally valued but the same approach also applies, in principle, to the Waterlow and Braden instruments.

Scrutiny of each instrument’s configuration reveals that for each risk-component, the rate of change is constant. For example, in the Waterlow RAS, if the level of continence deteriorates from occasionally incontinent to catheterised/incontinent of faeces the numerical value attached to each level of continence would be one and two, respectively. As a result, the increase in the overall risk score would be one unit, other things being equal. On the other hand, if the level of skin type is classed as ‘tissue’ and subsequently as ‘oedematous’ this re-classification would leave the total score unchanged. Similar comments hold for the Norton and Braden scales, although the rate of change within their sub-scales would be unchanging rather than non-existent.

The equal-weighting technique is the simplest approach to scale scoring. However, it fails to consider that some risk factors may be more important to the construct underlying the scale than others and should therefore contribute more to the overall risk score (Bowling, 1991). This amounts to saying that differential valuing should be applied to dissimilar risk factors based on their empirically demonstrated importance. Failure to use this mechanism means that the non-weighted total risk score may diverge from the true, yet unknown value, possibly affecting the clinical effectiveness of the planned nursing interventions and distorting the allocation of resources.

5.6. Towards a solution

It is clear that a PU risk instrument of high quality and practical utility must possess reasonably high predictive performance and convenience in use. In order to progress towards the ideal, several authors (Cullum et al., 1995, Deeks and Dealey, 1996) have called for the use of statistical modelling to determine and adequately weight the risk factors. This is because multivariable estimation methods make it possible to capture the quantitative impact of several prognostic variables on PU development ceteris paribus (Papanikolaou et al., 2002). In response to this, several investigators have embarked on the use of multivariate techniques to improve the available risk assessment methods. The next section provides an account of these research attempts.

6. Multivariate modelling of pressure ulcer risk—a new approach to an old problem

Table 1, adapted and updated from Nixon and McGough (2001), presents some characteristics of the
### Table 1
Pressure ulcer prospective cohort and cohort studies using multivariate methods

<table>
<thead>
<tr>
<th>Study detail</th>
<th>Sample information</th>
<th>Significantly found risk factors</th>
<th>Statistical methods</th>
</tr>
</thead>
</table>
| **Guralnik et al., 1998**  
Objective: To ascertain the risk factors affecting the PU occurrence. | 5193, US nationwide cohort 55–75 years, 10 years follow-up | 1. Appetite  
2. Skin condition  
3. Frequency of care  
4. Norton score  
5. Age  
6. Diagnosis | Logistic analysis |
| **Berlowitz and Wilking (1989)**  
Objective: To identify prospectively risk factors for PU and compare these results with a cross-sectional analysis in the same population. | Cohort data 185 Chronic medical records | Cohort results 1. Cerebrovascular accident  
2. Bed or chair bound  
3. Impaired nutritional intake | Logistic analysis |
| **Ek et al. (1991)**  
Objective: To investigate the relationship between nutritional state and the pressure ulcer development. | 495 Long-term medical patients LOS > 3 weeks | 1. Albumin  
2. Mobility  
3. Activity  
4. Food intake | Multiple regression |
| **Hoshowsky and Schramm (1994)**  
Objective: To determine the key variables affecting the PU development in operative patients. | 505 Surgical patients | 1. Time on operating table  
2. Vascular disease  
3. Age over 40 years  
4. Pre-operative Hemphill scale | Logistic analysis |
| **Allman et al. (1995)**  
Objective: To identify prospectively specific demographic, medical, functional status, and nutritional characteristics of bedridden and chairbound hospitalised patients that predict the in-hospital development of stage 2 and greater PU. | 286 In-patients > 55 years old Bed/chair > 4 days Hip fracture LOS > 5 days | 1. Non-blanchable erythema of intact sacral skin  
2. Lymphopenia  
3. Immobility  
4. Dry sacral skin  
5. Decreased body weight | Multivariate Cox regression |
| **Perneger et al. (1998)**  
Objective: To identify risk factors for PU in adult patients in a Swiss general hospital. | 2,373 hospital patients Age > 15 years old | 1. Age  
2. Total Norton score  
3. Hospitalisation for fracture  
4. Surgical intervention  
5. Nasogastric tube or Intravenous line | Proportional hazards model (A time-failure analytic method) |
| **Haalboom et al. (1999)**  
Objective: | 113 hospitalised patients exc. ICU, all ages | 1. Incontinence for urine  
2. Neurological disturbances  
3. Friction forces | Logistic analysis |
<table>
<thead>
<tr>
<th>Study detail</th>
<th>Sample information</th>
<th>Significantly found risk factors variables</th>
<th>Statistical methods</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Objective:</strong> To identify variables associated with post-operative PU incidence.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Van Marum et al. (2000)</strong></td>
<td>220 Nursing home patients Age &gt; 64 years old Not admitted for psycho-geriatric care Examined for PU within 48 hours of admission</td>
<td>1. Mobility 2. Mental health</td>
<td>Logistic analysis</td>
</tr>
<tr>
<td><strong>Objective:</strong> To investigate the contribution of the risk factors of the Dutch PU RAS in the detection of nursing home patients at risk of PU development.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Anthony et al. (2000a)</strong></td>
<td>777 elderly in-patients Age &gt; 64 years old LOS &gt; 6 days Excluded cases with an on-admission PU</td>
<td>1. Waterlow total score 2. Serum albumin</td>
<td>Logistic analysis</td>
</tr>
<tr>
<td><strong>Objective:</strong> To ascertain the relevance of nutritional substances such as serum albumin as predictor to PU in addition to the Waterlow score.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Objective:</strong> To determine those risk factors that best optimise the predictive ability of the Waterlow scale.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Objective:</strong> To ascertain the relative contribution of a range of risk factors to the risk of pressure ulcer. The Waterlow risk sub-scales were redefined and better organised.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Anthony et al. (2002)</strong></td>
<td>All admission over a 5-year period, N = 45,735.</td>
<td>1. Age</td>
<td>Logistic analysis</td>
</tr>
<tr>
<td><strong>Objective:</strong> To explore the relevance of ethnicity in pressure ulcers.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Perneger et al. (2002)</strong></td>
<td>1,190 in-patients drawn from internal medicine, orthopaedics, abdominal surgery, intensive care, and dermatology. Age &gt; 15 years old</td>
<td>1. Age 2. Friction 3. Mobility 4. Mental Status</td>
<td>Proportional hazards model</td>
</tr>
<tr>
<td><strong>Objective:</strong> To derive a PU risk-calculator for patients admitted to acute care hospitals.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Papanikolaou et al. (2003)</strong></td>
<td>498 Medical in-patients Age &gt; 64 years old LOS &gt; 4 days</td>
<td>1. Appetite 2. Inter-hospital transfer 3. Skin condition 4. Mobility 5. Gender 6. Incontinence</td>
<td>Logistic analysis</td>
</tr>
<tr>
<td><strong>Objective:</strong> To ascertain the relative contribution of a range of risk factors to the risk of pressure ulcer. The Waterlow risk sub-scales were redefined and better organised.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** This list is representative and not intended to be exhaustive. It merely provides examples and serves to highlight the scope of the issue in discussion. Studies using discriminant analysis are not included because it is not considered a reliable analytic method due to strong distributional conditions being assumed for the disturbances.
studies which have been carried out in order to
formulate empirically derived risk scores using multi-
variate methods. The following review deals with a
sample drawn from this list.

Alman et al. (1995) used multivariate analysis to
examine the main factors affecting PU development
among hospitalised bedridden and chairbound patients.
The study showed that five risk factors (non-blanchable
erythema of intact skin, immobility, dry skin, decreased
body weight) accounted for the PU development. Its
merit lies in the fact that advanced regression methods
were used to explore the relationship between PU
occurrence and those risk factors.

Anthony and his associates, employed regression
analysis to investigate the possibility of improving the
predictive performance of the Waterlow scale. One study
indicated that a shorter version of the Waterlow score
had better classification ability in wheelchair patients
than in general inpatients (Anthony and Unsworth,
1998). Another suggested that a short version of
Waterlow RAS was almost as effective as other more
complex alternatives (Anthony et al., 2000a). Two
additional studies demonstrated that albumin
was a significant predictor of pressure damage (Anthony
et al 2000b, 2002).

Perneger et al. (2002) applied multivariate techniques
to determine the key prognostic factors of hospital-
acquired PU so as to derive a simple, yet effective, PU
risk scale. The analysis indicated that the risk factors
‘friction/shear’, ‘age’, ‘mobility’ and ‘mental health’ were
significant and used to form ‘Fragment’, a new risk
instrument, whose predictive performance was found to
be better than that of the Norton and Braden scale.
However, an assessment of its reliability has yet to be
undertaken. This study demonstrated that it is possible
to build an empirically driven PU risk scoring system.

Papanikolaou et al. (2002, 2003) used logistic analysis
to ascertain the possibility of developing an improved
risk-instrument based on the Waterlow RAS. The
dependent variable was hospital-acquired PU. The
covariates were entered in the model as a series of dummy
variables, thereby capturing the quantitative
threshold effect of each one (Harrell et al., 1996). Owing
to built-in limitations in the Waterlow scale, a point
demonstrated earlier (see Section 3), it was imperative to
re-define and re-group the categories on some of the
risk-items. The regression analysis demonstrated that a
simplified version of the Waterlow RAS achieved good
predictive performance and had the potential for further
development. The contribution of this research is
twofold. First, re-grouping the categories on some of the
sub-scales of the Waterlow RAS has made it possible
to get around the limitations inherent in this tool.
Second, using multivariable techniques has allowed
assessment of the relative contribution of the risk factors
used in this popular RAS more accurately than
previously. The main limitations (acknowledged by the
authors) are the use of retrospective data and lack of
control for any preventive strategies employed by the
carers.

From the studies described above and Table 1, the
following points may be drawn:

(1) Although there are differences in the study design
(prospective vs. retrospective), the estimation tech-
nique used to fit the stochastic equation (logistic vs.
proportional hazards), the sample-specific features
(adults vs. elderly people), the clinical setting
(nursing vs. hospital care), a few risk-factors have
consistently and repeatedly emerged to be significant
predictors of PU occurrence: mobility, perfusion,
age, nutrition, skin assessment, medical condition
and mental health.

(2) Whilst multivariate techniques make it possible to
quantify the effect of risk factors upon PU
occurrence, their utility is limited. This is because
the results of any individual regression analysis are
related to a particular sample (however, small or
large this might be), and set of circumstances. Hence,
the results cannot be generalised without further
replication studies.

(3) There is a need to account for context-specific
aspects such as inter-hospital transfer (Papanikolaou
et al., 2002, 2003) and to differentiate between
patient and complication-specific risk factors (Ber-
lowitz et al., 1989).

(4) Large datasets with clear definitions of pressure
damage and its characteristics and data on the
significant risk factors are required in order to
permit the application of regression techniques
leading to the development of both a basic nation-
wide data-driven risk assessment scale and associ-
ated context specific modules.

7. Concluding remarks

This paper has shown that the Norton and Waterlow
RASs all contain some risk factors, for which the
definition is unclear or ambiguous and difficult to
understand. This introduces error into the derived
scores.

Also the review has suggested that these two scales
together with the Braden scale have inadequate reli-
bility and insufficient predictive performance, whilst the
critical threshold point for each risk instrument varies
across different clinical care settings and/or patient
populations. To deal efficiently with this issue, health-
care units should determine their own cut-off point,
after taking into account the special needs of their
patient populations and/or local clinical settings.
Furthermore, for the first time, it has shown that it is too simplistic to use the equal-weighting scoring technique to capture empirically and adequately the underlying PU risk score. This method has several built-in limitations, as demonstrated above (see Section 5.5).

The solution that has been adopted in the practical setting is to follow available guidelines, which draw attention to some of these limitations and advise caution in the use of RAS by combining them with clinical judgement. We question the utility of this approach. If RAS have such limitations, what contribution can they make to our confidence in clinical judgement, other than prompting us about the items, which should be considered when making such judgements? If they do not make such a contribution, how can we justify the time spent in scoring these factors, computing risk scores, and processing and storing the resultant data?

We suggest that a more efficient use of nursing time would be to ensure that all who conduct PU risk assessments are properly trained in the factors to be considered and give particular attention to those which are of most relevance in a given clinical setting, without having to assign scores of doubtful value. Furthermore, that the data to be gathered should relate to the characteristics of patients who do develop PU in the various settings and that this should be collected in a form which allows the accumulation of large, comparable data sets which can be used to derive more robust risk assessment tools, using the differential weighting scoring approach and advanced statistical methods. In other words, strengthen the reliance we place on the clinical judgement of trained and experienced nurses until we can produce tools which actually enhance it further and enable less experienced people to use them with confidence.

Acknowledgement

We would like to thank the following people: Ms. Lisa Franklin and Mr. Philip Satherley for providing us with helpful comments during the development of this study; Prof. Davina Allen, Director of the Centre for Nursing, Health and Social Care Research for her support and encouragement; and lastly Mrs. Faye Hurn for her generous secretarial support. We are grateful to Dr. Dimitri Parthimos, a lecturer at the Department of Diagnostic Radiology and Dr. Gianni Zois, a mathematician for their constructive comments. We would also like to thank two anonymous referees and the editor for their assistance for improving the paper.

References


Smith, L.N., Booth, N., Douglas, D., Robertson, W.R., Walker, A., Durie, M., Fraser, A., Hillan, E.H., Swaffield,


