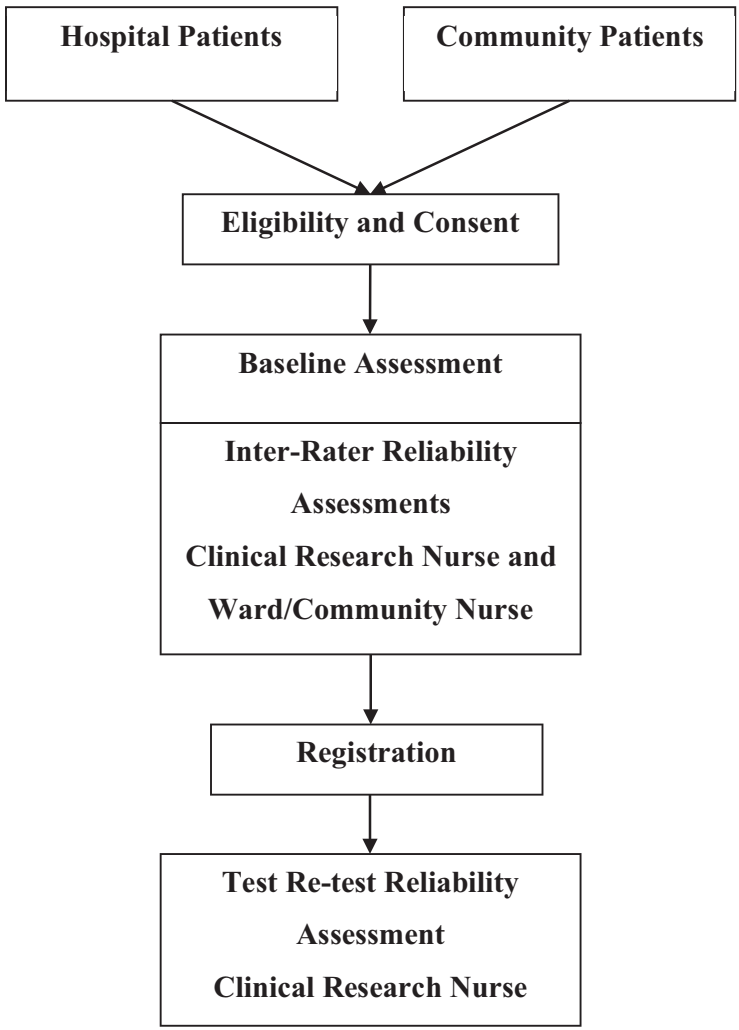
NB: This study protocol (version 3, dated 30 Oct 2012) is in a reduced format including only the study aims, methods and ethical considerations. Sections pertaining to study background have been removed as they are included as a chapter section. Information pertaining to serious adverse events, data monitoring, quality assurance, confidentiality, archiving, statement of indemnity, study organisational structure, funding, and publication policy are available upon request

4 FLOW DIAGRAM PURAF FIELD TEST 1



6 AIM AND OBJECTIVES

The aims of Field Test 1 are to:

1. assess the inter- rater and test-retest reliability of the PURPOSE T
2. assess the convergent validity, known groups validity, data completeness and clinical usability of the PURPOSE T.

7 FIELD TEST 1 METHODS

7.1 Design

The PURPOSE T will be evaluated through field testing using observational descriptive methods. The Field Test will evaluate the PURPOSE T in relation to its inter-rater reliability, test re-test reliability, data completeness, convergent validity, known group differences and clinical usability. Appendix 2 presents full details of the tests and criteria used in the instrument evaluation.

In-patients and community nursing patients will be invited to participate. Demographic characteristics and pressure ulcer risk will be assessed for all patients. Paired assessments will be undertaken using the PURPOSE T, one by a ward/community nurse and one by a nurse from the Tissue Viability Team (TVT; Tissue Viability Nurse Consultant/Specialist/Research Nurse) with specialist tissue viability knowledge. To minimise patient burden the clinical skin assessment component of the PURPOSE T assessment will be undertaken simultaneously by both assessors, but recorded separately with blinding maintained. The other components of the assessment will be undertaken separately and each nurse will remain blind to the corresponding assessment. Finally the 'TVT nurse will reassess the patient using the PURPOSE T at a clinically appropriate timeline determined on an individual patient basis but broadly 1-7 days after the first assessment.

7.2 Description of pressure ulcer risk primary or secondary evaluation tool (PURPOSE T)

The PURPOSE T instrument has been developed to identify whether patients are 'not at risk' or 'at risk' of pressure ulcer development. It consists of 22 data items for the assessment of 6 risk factor domains (mobility, skin, nutrition, perfusion, moisture and sensory percention). 19 items are a yes/no response and 3 items are a 3 point categorical sub-scale. Completion of the

assessment framework leads to a decision about risk status. Nurses using the PURPOSE T will have completed standard training in its use. A draft of the provisional PURPOSE T which is currently being developed by a graphic designer prior to the pre-test stage (April-May 2012) is detailed in Appendix 1.

7.3 Patient eligibility

7.3.1 Inclusion criteria

Patients who meet the following inclusion criteria:

- Aged ≥ 18years
- An inpatient in the acute setting or community nursing patient in the community setting
- Give their written informed consent/verbal witnessed consent/consultee agreement
- Expected to be available for the PURPOSE T re-test

7.3.2 Exclusion criteria

- Patients in obstetric, paediatric, day case surgery or psychiatric settings(acute or community)
- Unable to provide consent/consultee agreement
- Deemed by the attending healthcare professional to be too unwell to be approached and/or complete the study assessment schedule

7.3.3 Sampling strategy

An approximate sample of 230 patients will be purposively sampled ensuring a similar number of hospital and community patients and representation of patients across 4 broad levels of risk (as defined by their mobility and PU status) as follows:

- No mobility restrictions
- Some mobility/ activity limitations
- Bedfast/chairfast
- PU category ≥1

Each ward/community nurse will identify patients on their caseload who have the above characteristics.

We will monitor patient characteristics for other key risk factors including micro and macro circulatory disease, diabetes, nutritional deficits and moisture problems and target sampling if required.

In the hospital setting, specialties (vascular, elderly, medicine, orthopaedics, oncology, surgery) and acute/elective wards will be mapped and ward nurses will be identified in all these areas, ensuring balanced representation of patients.

7.4 Recruitment and consent

Ward/community based nurses will identify suitable patients from their area of practice. A full verbal explanation of the study Patient Information Leaflet will be provided by the attending clinical staff or a member of the Tissue Viability Team (TVT; Tissue Viability Nurse Consultant/Specialist/Research Nurse) for the patient to consider. This will include detailed information about the rationale, design, and personal implications of the study. Following information provision, patients will have as long as they need to consider participation and will be given the opportunity to discuss the study with their family and other healthcare professionals before they are asked whether they would be willing to take part in the study. Assenting patients will then be invited to provide informed, written consent. Should the patient be capable of giving consent but physically unable to complete the written aspects of the consent form, witnessed consent should be obtained using the Witnessed Consent Form. An appropriate witness would be a family member or friend of the patient, or another member of the patient's healthcare team who is not directly involved in the research study.

A record of the patient involvement in the study and consent/assent process detailing the date of consent will be documented in the patient healthcare records.

Assessment of eligibility and informed consent will be undertaken by a member of the TVT The right of the patient to refuse consent without giving reasons will be respected. Further, the patient will remain free to withdraw from the study at any time without giving reasons and without prejudicing any further treatment.

The original consent/assent form will be retained in the Investigator Site File, a copy of the consent form will be given to the patient, a second copy filed in the patient's healthcare records and a third copy will be sent to CTRU.

7.4.1 Consultee agreement

A large proportion of patients suffering from pressure ulcers/at risk of pressure ulcers have receptive or comprehension or language difficulties. They may also have general cognitive impairment affecting their understanding and/or dementia. To ensure that the study population is representative of the clinical population assessed in the course of usual care, recruitment procedures will facilitate consultee agreement. This is important because the nature of a pressure ulcer risk assessment includes history taking and also clinical examination, both of which are impacted when patients have cognitive impairment or language difficulties. In order to assess the reliability and validity of the PURPOSE T as the basis for use in clinical practice, it is important that the study population is a representative patient population.

The assessment of capacity will relate specifically to decisions pertaining to this particular research project. Each patient will be assumed to have capacity unless it is established that they lack capacity. Ward/community based nurses identifying patients for study participation, will be asked to consider aspects of capacity before any approach to patients is made and during the information giving stage prior to consent. The TVT member will assess the patient's ability to understand what decisions they need to make and why; the consequences of the decision to participate; their ability to understand, use and retain the information related to the decision to participate and be able to communicate their decisions effectively (as specified in the Mental Capacity Act 2005). If there is any concern about capacity the ward/community based nurses/TVT member will consult further with other members of the attending clinical team and/or relative/carer/friend (as appropriate) and a decision will be made with the relative/carer/friend as to whether the patient is able to provide written consent. Where the patient is thought not to have capacity to consent, a relative, carer or friend who is interested in the patient's welfare will act as a personal consultee.

The relative/carer/friend will be involved in the information and decision making process with the patient and will advise the TVT member on their presumed wishes and feelings and Consultee Assent will be obtained on behalf of the patient. The relative, carer or friend will

be advised to set aside their own views and provide advice on the participation of the patient in the research, taking into consideration the patient's wishes and interests. Research participants will not be required to do anything which is contrary to any advance decisions or statements that have been made by them in relation to their treatment or any other matter. Advance decisions made by the patient about their preferences and wishes will always take precedence.

If, despite taking all reasonable steps, a personal consultee cannot be identified and contacted then a nominated consultee would be approached. This person would have no connection with the research project. They would be nominated by the TVT member; they would most likely be the participant's lead clinician, their GP or a member of the care team. The consultee would be provided with the information leaflet describing the research study and the role of the consultee and it would be emphasised that they are being asked to act on behalf of the participant, rather than any personal views or feelings.

It is unlikely to place a major burden on consultees as the research is a non-invasive study that has minimal burden on the participant. There are no changes in treatment relating to the study.

7.5 Registration

Patients who are both eligible for study participation and provide informed consent/consultee agreement will be registered. Informed consent for entry into the study must be obtained prior to registration. Following confirmation of informed consent/consultee agreement and eligibility patients will be registered into the study by an authorised member of staff at the study research site.

Registration will be performed centrally using the CTRU automated 24-hour telephone registration system. Authorisation codes and PINs, provided by the CTRU, will be required to access the registration system.

The following information will be required at registration:

- Patient details, including initials, gender and date of birth
- Site code for research site

- Name of person making the registration
- Confirmation of eligibility
- Confirmation of informed consent/consultee agreement

Direct line for registration +44 (0)113 343 3377

8 ASSESSMENT AND DATA COLLECTION

Assessments will be undertaken as follows:

- Baseline
  - Demographics
  - Clinical assessment
  - PURPOSE T assessment (Ward/Community Nurse) and PURPOSE T (member of TVT)
- Test-retest
  - PURPOSE T (same member of TVT who undertook the PURPOSE T assessment at baseline) at a clinically appropriate timeline determined on individual patient basis but broadly 1-3 days for hospital patients and 1-7 days community patients

8.1 Research Assessments

8.1.1 Baseline demographics

A member of the TVT will record baseline demographics information including:

- Name of NHS Trust
- NHS Facility/Service name (name of hospital/intermediate care/community nursing team)
- Type of admission/referral (e.g. elective/acute)
- *Hospital patients only* - speciality
- Initials
- Date of birth
- Gender
- Ethnicity

To enable the test-retest follow-up the TVT member will record and destroy after the visit the following information:

- Patient's NHS ID
- Patient's Hospital/Trust number (if applicable)
- *Hospital patients only* – ward number/name
- *Community patients only* – place of residence

8.1.2 Baseline clinical assessment

A member of the TVT will record baseline clinical assessment including:

- Date and time of assessments
- Braden Score (Braden and Bergstrom 1987) (Appendix 3)
- Waterlow Scale (Waterlow 1985) (Appendix 4)

Table 1 NPUAP/EPUAP Pressure Ulcer Classification System (2009)

Category	Description
Category I Non-blanchable redness of intact skin	Intact skin with non-blanchable erythema of a localised area usually over a bony prominence. Discolouration of the skin, warmth, oedema, hardness or pain may also be present. Darkly pigmented skin may not have visible blanching.
Category II Partial thickness skin loss or blister	Partial thickness loss of dermis presenting as a shallow open ulcer with a red pink wound bed, without slough. May also present as an intact or open/ruptured serum-filled or sero-sanginous-filled blister.
Category III Full thickness skin loss (fat visible)	Full thickness tissue loss. Subcutaneous fat may be visible but bone, tendon or muscle are *not* exposed. Some slough may be present. *May* include undermining and tunnelling.
Category IV Full thickness tissue loss (muscle/bone visible)	Full thickness tissue loss with exposed bone, tendon or muscle. Slough or eschar may be present. Often includes undermining and tunnelling.
Category U (Unstageable/ Unclassified) Full thickness skin or tissue loss – depth unknown	Full thickness tissue loss in which actual depth of the ulcer is completely obscured by slough (yellow, tan, grey, green, or brown) and/or eschar (tan, brown, or black) in the wound bed.

8.1.3 PURPOSE T assessment

- Ward/Community Nurse
    - Date and time of assessment
    - PURPOSE T including skin assessment (sacrum, buttocks, heels, hips and other) (Appendix 1)using the skin classification scale (Table 1)
- Member of TVT
    - Date and time of assessment
    - PURPOSE T including skin assessment (sacrum, buttocks, heels, hips and other) (Appendix 1) using the skin classification scale (Table 1)

8.1.4 Test re-test risk assessments

- Same TVT member that undertook the first PURPOSE T assessment (8.1.4)
    - Date and time of assessment
    - PURPOSE T including skin assessment (sacrum, buttocks, heels, hips and other) (Appendix 1)using the skin classification scale (Table 1)
    - Clinically relevant changes to condition since baseline assessment

8.2 Data collection Procedures

8.2.1 Baseline assessments

Following informed consent/relative assent and at a time convenient to the patient the TVT member will complete demographic, clinical assessments and all components of the PURPOSE T apart from the skin assessment component. This baseline assessment will involve general observation (for example of spontaneous movement), history taking, and consulting relevant sections of the medical/nursing records.

A paired ward/community nurse PURPOSE T assessment will be undertaken separately at a time convenient to the patient and close enough in time to the TVT assessment to avoid any change in clinical condition. A PURPOSE T assessment proforma will be provided to the ward/community nurse with pre-populated standard header details including patient initials, date of birth and study ID. The ward/community nurse will complete all components of the PURPOSE T apart from the skin assessment component.

To minimise patient burden and due to the transient nature of alterations to intact skin which impacts upon the reliability of the skin assessment component (Nixon et al 2005a) the Stage 2 clinical skin assessment component of the PURPOSE T assessment will be undertaken simultaneously by both the ward/community nurse and TVT member, but recorded on separate PURPOSE T proformas, with blinding maintained. This method has been successfully adopted previously in an inter-rater reliability skin assessment study (Nixon et al 2005, Nixon et al 2006).

Following blinded completion of the PURPOSE T proforma the TVT member and the ward/community nurse will separately fold and seal their copies of the completed pro-forma. The TVT member will return the sealed proforma's with the other study documentation in a sealed envelope to the Clinical Trials Research Unit and the other sealed carbonated copies of the PURPOSE T will be kept in the site file.

8.2.2 Test re-test

The TVT member who undertook the initial PURPOSE T assessment will undertake a second PURPOSE T assessment, without access to the first assessment. The length of the test re-test interval must be short enough to ensure that clinical change in the PU is unlikely to occur, but sufficiently long to ensure that respondents do not recall their responses from the first assessment. A short test re-test interval is necessary to ensure that stability per se is being evaluated, rather than clinical change in the PU during the test re-test interval, which will underestimate reliability. We anticipate that the re-test will be undertaken between 1-3 days in hospital patients and 1-7 days in community patients, after the first assessment depending upon the anticipated recovery/deterioration/stability of the patients' condition and for hospital patients, length of stay. The assessing nurse will determine the re-test date and time, with the patient at the end of the baseline assessment visit.

8.2.3 PURPOSE T Field Notes

The TVT members involved in data collection will keep field notes of their experience of using PURPOSE T in clinical practice. The field notes will be summarised and used to inform design amendments and issues of importance for implementation.

9 STATISTICAL CONSIDERATIONS

9.1 Sample size

9.1.1 Inter-rater reliability

In the study population we expect approximately 25% will be 'not at risk' and 75% 'at risk'. In a 2-rater study, the numbers of subjects required to detect a statistically significant κ (2-sided p-value ≤ 0.05) with 90% power and 75% assessed as being 'at risk', assuming a null hypothesis value for κ are given in Table 2. To establish whether the tool gives a high degree of beyond-chance agreement, we will test against a null value of 0.6. With 90% power, 199 patients will be required. To allow for withdrawal/non-compliance in paired ward/community nurse assessments of 15% we will aim to recruit 230 patients.

9.1.2 Validity assessment

No examples of formal sample size estimation methods for evaluation of screening instruments were identified in the literature. Therefore, literature relating to the psychometric evaluation of rating scales were considered. The 'rule of thumb' recommendation of 5-10 patients for every item in a questionnaire has been used to estimate the sample size of 115-230 patients (Nunnally and Bernstein 1994, Blazeby et al 2002). The proposed sample size of 230 to assess the inter-rater reliability of the instrument, with >95% TVT data compliance (based upon previous research experience), will therefore provide sufficient numbers of patients to assess the validity of the risk assessment instrurment.

Table 2 Inter-rater reliability sample size estimates

Kappa to detect	Null value	N required patients (90% Power)
0.7	0.4	114
0.7	0.5	231
0.7	0.6	793
0.8	0.4	64
0.8	0.5	103
0.8	0.6	199
0.8	0.7	536
0.9	0.4	41
0.9	0.5	58
0.9	0.6	89
0.9	0.7	159

9.2 Analysis methods

The analysis plan outlined in this section will be reviewed and a final statistical analysis plan will be written before any data summaries or analyses are performed. The analysis plan will be written in accordance with current CTRU Standard Operating Procedures. Any changes to the final analysis plan and reasons for change will be documented.

9.2.1 Inter-rater and test re-test reliability

Kappa is a statistic that is used to measure agreement beyond that expected by chance, and thus is a measure of "true agreement". It indicates the proportion of agreement beyond that expected by chance (Cohen 1960). Thus kappa is the achieved beyond-chance agreement as a proportion of the possible beyond-chance agreement (Sim and Wright, 2005). The simplest use of kappa is in the situation in which two clinicians each provide an assessment of presence or absence of a characteristic representing inter-rater reliability or when a clinician provides two assessments of the same patient in relation to the presence or absence of a characteristic, representing intra-rater (test re-test) reliability. The concern is how well the ratings agree, not with how well the ratings agree with some "gold standard" or "true" diagnosis.

The range of possible values for kappa is from -1 to 1, though it usually falls between 0 and 1. Unity represents perfect agreement, whereas zero represents agreement expected by chance. Although kappa represents the proportion of agreement greater than that expected by chance, its interpretation is not so straightforward, as there are other factors that can influence the magnitude of the coefficient or the interpretation that can be placed on a given magnitude. Among the factors that can influence the magnitude of kappa are prevalence, bias and non-independence of ratings.

Kappa can be adjusted for prevalence and bias with the resulting kappa coefficient is referred to as a PABAK (prevalence-adjusted bias-adjusted kappa). It is recommended that PABAK is presented in addition to, rather than instead of kappa. Interpretation guidelines have been proposed as standard strengths of agreement for kappa and are detail in Table 3 (Landis and Koch, 1977):

Table 3 The Kappa statistic

Kappa	Strength of agreement
≤0	Poor
0.01-0.20	Slight
0.21-0.40	Fair
0.41-0.59	Moderate
0.60-0.79	Substantial
0.81-1	Almost perfect

It has also been suggested that the interpretation of kappa could be assisted by reporting the maximum value is could attain for the set of data concerned. To calculate the maximum value of kappa (κ_{max}) the proportion of positive and negative judgements by each clinician are taken as fixed and the distribution of paired ratings is adjusted so as to represent the greatest possible agreement. In contrast to PABAK, κ_{max} serves to gauge the strength of agreement while preserving the proportion of positive ratings demonstrated by each clinician. Finally a 95% confidence interval can be constructed around kappa (Bland 2008).

We will undertake kappa (with 95% CIs), PABAK and κ_{max} to assess the inter-rater reliability for agreement of risk status overall (at risk/not at risk). To further ensure the reliability of any findings we will also examine the extent of agreement for individual PURPOSE T items.

In order to assess the test-retest reliability of PURPOSE T, the same approach of using kappa statistics and their variants PABAK and κ_{max} will be employed, except that, rather than two independent raters assessing the risk status of the patient, it will be the same rater carrying out the assessment. In order to preserve the independence of the two assessments, the two assessments will need to be far enough apart in time for the rater not to remember their original assessment (we judge an appropriate time period to be at least 1 days), but also not so far apart that the patient's condition will have altered (we judge an appropriate time period to be no more than 3-7 days).

9.3 Acceptability and data quality

Acceptability will be determined by data quality; assessed by completeness of item-level data (percent of missing data for items) and completeness of confirmation of risk status (percent of people for whom it is possible to assess risk).

9.4 Convergent validity

Convergent validity assesses the degree to which constructs (or scores on a measure) expected to be related are, in fact, related. The degree to which assessment of 'at risk' and 'not at risk' are related to risk assessment status as assessed using the Braden (Braden and Bergstrom 1987) and Waterlow (Waterlow 1985) risk assessment scales will be determined.

A Spearman Rank correlation coefficient will be calculated between PURPOSE T and Braden and Waterlow risk status. In addition, where there are corresponding items between PURPOSE T and Braden and/or Waterlow (e.g. mobility), correlations will be performed to determine how closely PURPOSE T items are related to other risk screening items. For exploratory purposes, the following hypotheses will be proposed as guides to the magnitude of correlations, as opposed to pass/fail benchmarks (high correlation $r > 0.7$; moderate correlation $r = 0.3 - 0.7$; low correlation < 0.3) (Burnand, 1990; Cohen 1960). Moderate to high correlations ($r= > 0.3$) are predicted.

9.5 Known groups validity

Known-group comparisons are used to evaluate the clinical utility of instruments or assessment tools. This method assesses the extent to which the overall assessment or items are able to discriminate between subgroups of patients known to differ in terms of clinical presentations (Kerlinger, 1973).

A chi-square test for independence (used to compare the frequencies of cases found in the various categories of one variable across the different categories of another variable) will be used to determine whether type of hospital patient (e.g. elective vs. acute) is related to risk group (e.g. at-risk vs. not at-risk).We anticipate that there will be a significantly lower proportion of elective surgical patients assessed as 'not at risk', compared to acute patients.

12.2 Ethical considerations

This study will include elderly and highly dependent patients considered as vulnerable. Ethical issues are largely related to the involvement of vulnerable adults/elderly patients with

high levels of co-morbidity including acute and chronic illness. A large proportion of patients suffering from pressure ulcers/at risk of pressure ulcers have receptive or comprehension or language difficulties. They may also have general cognitive impairment affecting their understanding and/or dementia. To ensure that the study population is representative of the clinical population assessed in the course of usual care, recruitment procedures will facilitate consultee agreement. This is important because the nature of a pressure ulcer risk assessment includes history taking and also clinical examination, both of which are impacted when patients have cognitive impairment or language difficulties. In order to assess the reliability and validity of the PURPOSE T as the basis for use in clinical practice, it is important that the study population is a representative patient population.

The ethical issues surrounding these potentially vulnerable patients have been addressed through the study design and the use of local staff including experienced clinical nurses, that is, members of the local TVT to assess patients. In line with good clinical research practice, if a patient is clearly at risk or has an existing pressure ulcer and this is not reported in the patients' healthcare notes, then it will be documented in the patients' healthcare notes and reported to the ward/community nurse responsible for the patients care.

The study will be performed in accordance with the recommendations guiding physicians in biomedical research involving human subjects adopted by the 18th World Medical Assembly, Helsinki, Finland, 1964, amended at the 52nd World Medical Association General Assembly, Edinburgh, Scotland, October 2000. Informed written consent/witnessed verbal consent/consultee agreement will be obtained prior to involvement into the study. The right of a patient to refuse participation without giving reasons will be respected. The patient will remain free to withdraw at any time from the study without giving reasons and without prejudicing his/her further treatment. If a participant withdraws consent from further study participation their data will remain on file and will be included in the final study analysis.

18 REFERENCES

Bennet G, Dealey C, Posnett J (2004) The cost of pressure ulcers in the UK. Age and Ageing. 33(3): 230-235

Bland JM (2008). Measuring health and disease: Cohen's kappa. University of York Department of Health Sciences.

Blazeby, J., Sprangers, M., Cull, A., Groenvold, M., & Bottomley, A. EORTC Quality of Life Group: Guidelines for Developing Questionnaire Modules. 3rd edition revised. 2002.

Braden, B. and Bergstrom, N. (1987) A conceptual schema for the study of the aetiology of pressure sores. Rehabilitation Nursing, Vol. 12 (1) 8-16.

Burnand, B,. Kernan, WN. & Feinstein, AR. (1990). Indexes and boundaries for "quantitative significance" in statistical decisions. *Journal of Clinical Epidemiology*, 1278-84

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 1, 37-46

Collins, D. (2003). Pretesting survey instruments: an overview of cognitive methods. *Quality of Life Research*, 12, 229-238

Cullum N, Deeks J, Fletcher A, Long A, Mouneimne H, Sheldon T, et al. (1995). The prevention and treatment of pressure sores: How effective are pressure-relieving interventions and risk assessment for the prevention and treatment of pressure sores? *Effective Health Care Bulletin.* 2(1) pp1-16.

DeCoster, J. (2004) Data analysis in SPSS. Retrieved: http://www.stat-help.com/notes.html

Department for Education and Skills *(2006) Safeguarding Vulnerable Groups Bill,* London, The Stationery Office.

Department of Health (2000) *No Secrets: Guidance on developing and implementing multi-agency policies and procedures to protect vulnerable adults from abuse* [online] London. Department of Health. Available from: www.doh.gov.uk/scg/nosecrets.html

Department of Health. (2001a) *National Service Framework for Older People*. London, The Stationery Office.

Department of Health (2001b*) Essence of Care: Patient Focused Benchmarking for Health Care Professionals*. London, The Stationery Office

Department of Health (2008) Using the Commissioning for quality and Innovation (CQUIN) Payment Framework for the NHS England. London.

Department of Health (2009) NHS 2010-2015, Good to Great, Preventative People Centred, Productive. The Stationary Office, London

Department of Health (2010) Clinical Governance and Adult Safeguarding. London

European Pressure Ulcer Advisory Panel and National Pressure Ulcer Advisory Panel. Treatment of pressure ulcers: quick reference guide. Washington DC: National Pressure Ulcer Advisory Panel. 2009. Available from: http://www.epuap.org/guidelines/Final_Quick_Treatment.pdf

Gorecki C (2011). The development and validation of a patient reported outcome measure of health related quality of life for patients with PU's: PUQOL Study. PHD Thesis, University of Leeds.

Gorecki C, Brown JM, Nelson A, Briggs M, Schoonhoven L, Dealey C, Defloor T and Nixon J. (2009). Impact of pressure ulcers on quality of life in older patients: a systematic review. *JAGS.* Vol 57, No7. Pp1175-1183

Feinstein, AR & Cicchetti, DV (1990) High agreement but low kappa: I. The problems of two paradoxes. Journal of Clinical Epidemiology; 43:9

Hartmann, DP (1977) Considerations in the choice of interobserver reliability estimates. Journal of Applied behavioural Analysis; 10:103-16

Institute for Innovation and Improvement (2009) High Impact Actions for Nursing and Midwifery. London

Kaltenthaler E, Whitfield MD, Walters SJ, Akehurst RL, Paisley S. (2001). UK, USA and Canada: how do their pressure ulcer prevalence and incidence data compare? *Journal of*

Wound Care. 10(1), pp530-535.

Kerlinger, FN. (1973). Foundations of behavioural research. New York: Holt, Rinehart and Winston

Kitzinger J. (1995). Qualitative Research: introducing focus groups. *BMJ.* 311:299

Landis, J. Richard & Koch, Gary G. (1977) The Measurement of Observer Agreement for Categorical Data. *Biometrics*;1:159-174

McColl E, Jacoby A, /Thomas L, Soutter J, Bamford C, Steen N, et al. (2001). Design and use of questionnaires: a review of best practice applicable to surveys of health service staff and patients. Health Technology Assessment. 5 (31)

McColl, E. (2005). "Developing questionnaires," *In Assessing quality of life in clinical trials*, 2nd ed. P. Fayers & R. Hays, eds., New York: Oxford University Press, pp. 9-23.

The Mental Capacity Act (2005). MC Act, London, HMSO.

National Institute for Health and Clinical Excellence (NICE). (2005). *Pressure ulcer management: The Management of pressure ulcers in primary and secondary care: A clinical Practice Guideline.* RCN.

Nixon J, Coleman S, Gorecki C, Nelson EA, Closs J, Halfens R, Worthy G, Schoonhoven L, Brown JM. Pressure Ulcer risk factors: a systematic review. Unpublished

Nixon J, McGough A. (2001). Principles of Patient Assessment: Screening for Pressure Ulcers and Potential Risk. In Morison M, ed. The Prevention and Treatment of Pressure Ulcers. Edinburgh: Mosby, pp55-74.

Nixon, J., Cranny, G. and Bond, S. (2005a) Pathology, diagnosis and classification of pressure ulcers: comparing clinical and imaging techniques. Wound Repair and Regeneration Vol 13, p.365-372

Nixon, J., Thorpe, H., Barrow, H., Phillips, A., Nelson, E.A., Mason, S., Cullum, N (2005b) The reliability of pressure ulcer classification and diagnosis. Journal of Advanced Nursing Vol 50 (6) p.103-106

Nixon, J., Nelson, E. A., Cranny, G., Iglesias, C., Hawkins, K., Cullum, N., et al on behalf of the Pressure Trial Group. (2006) Pressure Trial: Pressure RElieving Support SUrfaces: a Randomised Evaluation. Health Technol Assess Vol 10 (22).

Nunnally, J. & Bernstein, I. (1994). *Psychometric Theory*, 3rd ed. New York, McGraw-Hill.

Safeguarding Vulnerable Groups Bill (2006), HL Bill 79.

Scott EA and Black. (1991). Appropriateness of Cholcystectomy in the UK – A Consensus Panel Approach. Gut 32:1066-1070

Sim J, Wright CC. (2005). The Kappa statistic in reliability studies: Use, interpretation, and sample size requirements. Physical Therapy, 85 (3).

Sopher R, Nixon J, Gorecki C, Gefen A. (2010) Exposure to Internal Muscle Tissue Loads under the Ischial Tuberosities during Sitting is Elevated at Abnormally High or Low Body Mass Indices. Journal of Biomechanics 43(2):280-286.

Waterlow J. (1985). Pressure Sores: a Risk Assessment Card. Nursing Times 81 (48) 49-55

Willis GB. (2005). Cognitive interviewing: a tool for improving questionnaire design. Thousand Oaks, California. Sage Publications, Inc

Appendix 1: Confidential DRAFT preliminary PURPOSE T

[NB: Data for the preliminary PURPOSE T was collected however the scale is omitted due to copyright. The final PURPOSE T can be obtained from URL:http://ctru.leeds.ac.uk/purpose].

Appendix 2: Reliability and Validity Tests and Criteria

Test Property	Definition/Test	Criteria (Traditional)
Data Quality Acceptability/Data completeness	The extent to which PURPOSE T items are completed and used to allocate a risk category; quality of data is assessed by data completeness for each element of the PURPOSE T and a risk category.	-% item level data missing -% of risk categories allocated -% of items missing where a risk category has been allocated
Inter-rater reliability **Test Re-Test Reliability**	Inter-rater reliability assesses the extent to which the PURPOSE T results obtained by two or more raters agree for the same population. Test re-test reliability assesses the stability of the PURPOSE T over a period of time in which the patient's condition is not expected to change.	- The kappa statistic is a measure of true agreement and indicates the proportion of agreement beyond that expected by chance, that is the achieved beyond-chance agreement as a proportion of the possible beyond-chance agreement.
Content Validity	The extent to which a scale measures what it intends to measure.	-Qualitative evidence from the PU risk factor systematic review and PU-MDS and PURPOSE T consensus study that items in the scale are representative of the construct being measured.
Convergent Validity (Between Scale analysis – analyses against external criteria) **Known group differences**	Evidence that PURPOSE T constructs are correlated with other measures of the same or similar constructs; assessed on the basis of correlations between the measure and other similar measures (Braden Scale and Waterlow Score). The ability of PURPOSE T risk categories to differentiate known groups; assessed by	-Correlations are expected to vary according to the degree of similarity between the constructs being measured by each instrument. Specific hypotheses are formulated and predictions tested on the basis of correlations.

	comparing PURPOSE T risk categories for subgroups who are expected to differ on the construct being measured (significant differences between known group or difference of expected magnitude) (e.g. elective/acute patients).	

Appendix 3: Braden Score

[NB: Data for the Braden scale was collected however the scale is omitted due to copyright. The Braden scale can be obtained from URL: http://bradenscale.com/].

Appendix 4: Waterlow scale

[NB: Data for the Waterlow scale was collected however the scale is omitted due to copyright. The Waterlow scale can be obtained from URL: http://www.judy-waterlow.co.uk/index.htm].