Screening for depression in older people in the acute general hospital setting: The performance of the Edinburgh Depression Scale (EDS)

Thesis submitted for the degree of

Master of Philosophy

at the University of Leicester

by

Collins Iheanyichukwu Esiwe MBBS, MRCPsych

Department of Health Sciences

University of Leicester

January 2011
Abstract

Screening for depression in older people in the acute general hospital setting: the performance of the Edinburgh Depression Scale (EDS)
C I Esiwe

Introduction
Depression affects approximately 25% of older people in general hospitals, yet it is often undiagnosed and under treated leading to extended lengths of stay and increased morbidity and mortality. Recent UK depression guidelines recommend screening at risk populations such as the physically ill. Several depression screening tools exist, but they include somatic symptoms that are common to both physically ill and depression. Many practitioners also find them cumbersome and time consuming for routine use.

Aim
The aim of this study is to assess the performance of the EDS as a brief screening tool for depression in older adults in the acute general hospital setting, and compare its performance with the Geriatric Depression Scale (GDS,4-item and 15-item) and the two questions suggested in the NICE depression guidelines (NICE-2; NICE, 2009).

Methods
118 eligible older adults under the care of 4 acute medical teams at the Leicester General Hospital between October 2006 and June 2010 were assessed for depression using the Present State Examination – Schedules for Clinical Assessment in Neuropsychiatry (PSE-SCAN; WHO, 1982). Participants were also screened for depression with the EDS, GDS and the two NICE questions .The PSE-SCAN determined ICD-10 diagnosis of depressive episode acted as ‘gold standard’. The test characteristics of the screening instruments were calculated at various cut-offs. Receiver Operating Characteristic (ROC) curves enabled comparisons of the various scales and determined the most appropriate cut-off threshold.

Results
The estimated prevalence of ICD-10 depressive episode was 22%. All scales had comparable sensitivity (80 – 100%), specificity (70 – 86%), and negative predictive values (93-100%). Areas under the ROC curves for the scales were comparable (range 0.85 -0.92) with no statistically significant difference between them.

Conclusion
The EDS and the two NICE questions performed just as well as other established screening tools and may be utilized in screening for depression in this population.
Acknowledgements

I would like to thank everyone who was involved in this study: Professor James Lindesay and Dr Michael Dennis for their advice and support during the supervision of this project and writing up of the thesis; Sarah Baillon for administering the screening tools to participants, assisting with statistical analysis and for advice on computational analysis of the data; Dr Aniruddha Rajkonwar for helping with recruitment of participants and data analysis; Dr Nelson Lo and all other medical colleagues and staff on medical wards of the Leicester General Hospital, for helping in identifying eligible patients for the study; my colleagues and staff at Lincolnshire Partnership NHS Foundation Trust including those at the audit department for their patience and support throughout this study.

I am also grateful to all my family and friends for encouraging me all the way.

Without you this research would not have been possible.
6.3 Detection of depression using the screening instruments
1. The Edinburgh Depression Scale
2. The 15-item Geriatric Depression Scale
3. The 4-item Geriatric Depression Scale
4. The two questions suggested by NICE depression guidelines
5. Test characteristics of the 4 instruments compared
6. Statistical analysis comparing AUCs of ROC curves

6.4 Detection of milder forms of depression using the screening instruments

SECTION FOUR: DISCUSSION

7. Discussion
7.1 Overview
7.2 Assessments and screening interviews
7.3 The setting of the study
7.4 Patients who did not take part in the study
7.5 Performance of the screening instruments
1. The Geriatric Depression Scale
2. The Edinburgh Depression Scale
3. The two questions suggested by NICE depression guidelines
4. The 4-item Geriatric Depression Scale
5. The four screening instruments compared
7.6 Subsyndromal or sub threshold depression
7.7 Strengths and weakness of this study
7.8 Conclusions and Clinical implications
7.9 Suggestions for further research

SECTION FIVE: APPENDICES

Appendix 1 DSM-IV criteria for major depressive episode
Appendix 2 ICD-10 criteria for depressive episode
Appendix 3 The Edinburgh Depression Scale
Appendix 4 The Geriatric Depression Scale
Appendix 5 NICE depression screening questions
Appendix 6 Consent form for study
Appendix 7 Confirmation of ethical opinion letter from LREC
Appendix 8 Patient information sheet

SECTION SIX: REFERENCES
List of Tables and Figures

Table

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Comparison of symptoms of depression in ICD-10 and DSM-IV</td>
<td>12</td>
</tr>
<tr>
<td>1.2</td>
<td>Prevalence estimates of depressive disorders among older adults in the community</td>
<td>14</td>
</tr>
<tr>
<td>1.3</td>
<td>Classification of antidepressants currently used in older adults</td>
<td>29</td>
</tr>
<tr>
<td>2.1</td>
<td>Prevalence of depression in older adults in selected treatment facilities</td>
<td>33</td>
</tr>
<tr>
<td>2.2</td>
<td>Medical conditions and central acting drugs that may cause organic depressive disorder</td>
<td>36</td>
</tr>
<tr>
<td>3.1</td>
<td>UK National Screening Committee criteria for appraising the viability, effectiveness and appropriateness of a screening programme</td>
<td>44</td>
</tr>
<tr>
<td>3.2</td>
<td>Screening instruments for detecting major depression in older adults in primary care</td>
<td>49</td>
</tr>
<tr>
<td>6.1</td>
<td>Demographic and characteristics of study sample</td>
<td>71</td>
</tr>
<tr>
<td>6.2</td>
<td>Age and sex specific prevalence of ICD-10 depressive episode</td>
<td>72</td>
</tr>
<tr>
<td>6.3</td>
<td>Performance of EDS at various cut-offs</td>
<td>73</td>
</tr>
<tr>
<td>6.4</td>
<td>Proportion of participants scoring at various cut-off points on the EDS</td>
<td>74</td>
</tr>
<tr>
<td>6.5</td>
<td>Performance of GDS at various cut-offs</td>
<td>76</td>
</tr>
<tr>
<td>6.6</td>
<td>Proportion of participants scoring at various cut-off points on the GDS</td>
<td>76</td>
</tr>
<tr>
<td>6.7</td>
<td>Performance of GDS-4 at various cut-offs</td>
<td>78</td>
</tr>
<tr>
<td>6.8</td>
<td>Proportion of participants scoring at various cut-off points on the GDS-4</td>
<td>78</td>
</tr>
<tr>
<td>6.9</td>
<td>Performance of NICE-2 at various cut-offs</td>
<td>80</td>
</tr>
<tr>
<td>6.10</td>
<td>Proportion of participants scoring at various cut-off points on the NICE-2</td>
<td>80</td>
</tr>
<tr>
<td>6.11</td>
<td>Comparison of performance of the 4 screening instruments</td>
<td>82</td>
</tr>
<tr>
<td>6.12</td>
<td>Cases of mild and moderate depression identified by the screening instruments using their optimum threshold</td>
<td>85</td>
</tr>
<tr>
<td>7.1</td>
<td>Comparison of the performance of the GDS-15 in a number of published validation studies in inpatient settings</td>
<td>95</td>
</tr>
<tr>
<td>7.2</td>
<td>Comparison of the performance of the EDS in a number of published validation studies</td>
<td>100</td>
</tr>
<tr>
<td>7.3</td>
<td>Comparison of the performance of the GDS-4 in a number of published validation studies</td>
<td>108</td>
</tr>
</tbody>
</table>

Figure

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Receiver operating characteristics curve for the EDS</td>
<td>75</td>
</tr>
<tr>
<td>6.2</td>
<td>Receiver operating characteristics curve for the GDS</td>
<td>77</td>
</tr>
<tr>
<td>6.3</td>
<td>Receiver operating characteristics curve for the GDS-4</td>
<td>79</td>
</tr>
<tr>
<td>6.4</td>
<td>Receiver operating characteristics curve for the NICE-2</td>
<td>81</td>
</tr>
<tr>
<td>6.5</td>
<td>Receiver operating characteristics curve for the GDS-15, GDS-4, NICE-2 and EDS</td>
<td>83</td>
</tr>
</tbody>
</table>
SECTION ONE: INTRODUCTION
1. DEPRESSION IN OLDER ADULTS

1.1 Introduction

According to the 1999 report of the World Health Organisation, Depression accounts for the loss of an average of 4 years of active life for every person in the world. It is projected to be the second leading cause of disability worldwide by 2020 (WHO, 2001) with substantial economic and public health implications. Depression is the most common psychiatric mood disorder in Older Adults with 10-20% of over 65s living in the community suffering some degree of depression (Beekman et al. 1999) and up to 45% in settings where co-morbid physical illnesses are common e.g., on acute medical wards, residential home residents and users of home care services (Koenig et al., 1988; Ames, 1990).

Depression in older adults is associated with increased mortality over and above age or physical ill-health effects (Koenig & Kuchibhatla, 1998). There is also widely accepted evidence to show that the outcome of a variety of medical conditions is adversely affected by depression increasing the likelihood of prolonged stay, unnecessary investigations and often admission to residential care (Koenig & Kuchibhatla, 1998). Depressive symptoms in older persons are an important risk factor for the development of cognitive impairment as well as functional decline. This in turn causes pain and suffering not just for the patient, but also for those who care for them. Perhaps more important is the fact that depression is a risk factor for suicide in the elderly particularly when associated with physical illness (Caine et al., 1996; Anderson, 2001).

Despite the fact that depression is treatable in 65-75% of elderly patients (Alexopoulos et al., 2001), it remains undiagnosed or undertreated particularly on acute medical wards. Part of the problem is that depression is often co-morbid with other medical conditions,
cognitive impairment as well as various social and economic difficulties, hence it is assumed to be a normal consequence of these problems. Older adults may also deny depression or perceive it as shameful, adding to under-recognition and lack of treatment.

To ensure that depressed patients are identified and receive appropriate treatment, screening at risk populations such as those with physical health problems has been advocated (NICE, 2009). Screening when combined with an integrated approach to the subsequent management of depression achieves desirable outcomes and benefits, not only for the patient, but for the health service as a whole.

1.2 Classification

Classifying a heterogeneous disorder like depression brings up a number of issues relevant to clinicians and researchers.

Firstly, some argue that it is best viewed as a unitary disorder with various manifestations forming a spectrum ranging from no disease through minor or sub-threshold depression to major depressive disorder. The problem with this one-dimensional approach is that it fails to capture the heterogeneous nature of the illness such as duration and course of the disorder.

Another approach adopted by most clinicians is a categorical classification. Categories have been defined in terms of symptom patterns and of the course and outcome of the different disorders. Operational diagnostic criteria such as DSM-IV-TR (American Psychiatric Association 2000) and ICD-10 (World Health Organization 1993) employ this method of classification and they are useful in clinical work and in research. The advantage of operational criteria like the DSM-IV and ICD-10 is that reliability of diagnosis is improved
because they rely on a number of discriminating symptoms to be present before a particular diagnosis can be made. The downside is that in clinical practice considerable overlap exists for these syndromes and boundaries are often not distinct. There is also the problem of failing to take into account co-morbid physical illness or significant cognitive impairment, which is common in older adults and often makes diagnosis of depression difficult.

Finally, a functional approach to classification aims to look at the degree to which the symptoms affects the ability of the individual to cope with activities of daily living, i.e. depressive symptoms that warrants attention would be one in which functional abilities have been significantly impaired, even though they may not meet the criteria for the categories described above.

The recently updated National Institute for Health and Clinical Excellence (NICE) guidelines on depression (NICE, 2009) highlights the importance of recognizing sub threshold depression because it is very common (particularly in older adults) and also causes considerable morbidity and human economic costs.

While being aware of these issues, the categorical approach continues to reflect current nomenclature and works relatively well for classifying depressive disorders in older adults.

**Depressive/Major Depressive Episode**

The criteria for diagnosing depression according to DSM-IV and ICD-10 are similar but not identical.
The DSM-IV, gives nine criteria for depression: depressed mood, sleep disturbance, lack of interest or pleasure in activities, guilt and feelings of worthlessness, lack of energy, loss of concentration and difficulty making decisions, anorexia or weight loss, psychomotor retardation or agitation, and suicidal ideation. The presence of at least 5 (1 or 2 core symptoms plus 4 or more of the others) of these criteria, occurring nearly every day during the same two-week periods supports the diagnosis of depression.

The DSM-IV divides major depression further into:-

**Sub threshold depressive symptoms:** Fewer than 5 symptoms of depression.

**Mild depression:** Few, if any, symptoms in excess of the 5 required to make the diagnosis, and symptoms result in only minor functional impairment.

**Moderate depression:** Symptoms or functional impairment are between ‘mild’ and ‘severe’.

**Severe depression:** Most symptoms and the symptoms markedly interfere with functioning and can occur with or without psychotic symptoms.

The NICE guideline on depression has adopted the DSM-IV definition of severity because it includes both number of symptoms and degree of functional impairment (Appendix 1&2).

The ICD-10 uses an agreed list of 10 depressive symptoms, and divides the common form of major depressive episode into four groups: not depressed (fewer than 4 symptoms), mild (4 symptoms; 2 of: depressed mood, loss of interest and enjoyment, and loss of energy/fatigue), moderately depressed (5-6 symptoms), and severe (7 symptoms or more, with or without psychotic symptoms). To qualify as depressed according to ICD-10, patients must have at least two of the three core symptoms and they must also be present for at least 2 weeks.
The criteria for DSM-IV and ICD-10 are summarized in appendix 1 and 2 respectively but the table below compares the symptoms of depression in both classificatory systems.

### Table 1.1
Comparison of symptoms of depression in ICD-10 and DSM-IV

<table>
<thead>
<tr>
<th>ICD 10</th>
<th>DSM-IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depressed mood</td>
<td>Depressed mood</td>
</tr>
<tr>
<td>Loss of interest and enjoyment</td>
<td>Loss of interest or pleasure</td>
</tr>
<tr>
<td>Reduced energy and decreased activity</td>
<td>Loss of energy/fatigue</td>
</tr>
<tr>
<td>Loss of confidence or self esteem</td>
<td>Worthlessness/excessive or inappropriate guilt</td>
</tr>
<tr>
<td>Ideas of guilt and unworthiness</td>
<td></td>
</tr>
<tr>
<td>Recurrent thoughts of death or self harm</td>
<td>Recurrent thoughts of death, suicidal thoughts or actual suicidal attempts.</td>
</tr>
<tr>
<td>Diminished ability to think or concentrate, or indecisiveness</td>
<td>Diminished ability to think or concentrate, or indecisiveness</td>
</tr>
<tr>
<td>Change in psychomotor activity with agitation or retardation</td>
<td>Psychomotor agitation or retardation</td>
</tr>
<tr>
<td>Sleep disturbance</td>
<td>Insomnia/hypersomnia (sleep disturbance)</td>
</tr>
<tr>
<td>Change in appetite with weight change</td>
<td>Significant appetite and/or weight loss</td>
</tr>
</tbody>
</table>

1, 2, 3 = core symptoms

The ICD-10 therefore requires only 4 symptoms for a diagnosis of mild depressive episode, compared to the DSM-IV which requires 5. In theory this means that the ICD-10 should identify more people as having a depressive episode while the DSM-IV will identify more severe depression.

### Minor, sub syndromal, or sub threshold depression

It is important to be aware of clinically significant depressive symptoms that do not meet the full criteria for the diagnosis of a depressive /major depressive episode because they are more common than major depression in older people. Eight to 20 percent of older adults in the community and up to 37% in primary care settings suffer from depressive symptoms (Alexopoulos, 1997). Minor depression is increasingly recognised as causing considerable morbidity and human and economic costs (NICE, 2009). It has been associated with impairment similar to that of major depression, including impaired physical functioning, disability days, poorer self-rated health and use of psychotropic medications (Beekman et al., 1995). It is also a risk factor for future major depression (Rowe &
Rapaport, 2006) and causes increased use of health services, excess disability, poor health outcomes, including increased mortality (Unutzer et al., 1997).

“Minor depression” is not yet recognised as an official disorder in both classificatory systems. In the DSM-IV research criteria, minor depression is diagnosed by the presence of at least two but less than five symptoms while in the ICD-10, it overlaps with the category of mild depressive episode (i.e. 4 symptoms) [see table 1.1 above].

Dysthymia is another category of subsyndromal depression that is present in both the DSM-IV and ICD-10. Dysthymic disorder is a persistent disturbance of mood, less severe than major depression that lasts for 2 years or longer.

1.3 Epidemiology

Prevalence

Estimates of the prevalence of depression in older adults varies widely, depending on sample selection, diagnostic tool or standardized interview employed, clinical setting of the study and “caseness” criteria.

Epidemiological studies using criteria designed for younger people (such as DSM or ICD) report a much lower prevalence of ‘major’ depression in old age than earlier in life but a relatively high prevalence of ‘minor’ depression or of depressive symptoms (Katona, 1994; Guruland et al., 1996). On the other hand, estimates based on symptom checklists or scales, are consistent with the clinical impression that the prevalence of depression increases with advancing age.

Table 1.2 summarizes the findings from a number of studies across Europe, North and South America, Finland and the Netherlands. What they show is that pervasive depression (clinically significant depression) affects between 10 – 15% of older adults in the
community while the prevalence of “depressive illness” is much lower, 1-5% in most community surveys. While it may appear that the prevalence of depressive illness is lower in older adults compared to young and middle-aged adults, it is possible that the diagnostic criteria (ICD-10 and DSM-IV) and standardized interviews (GMS-AGECAT, CES-D, DIS) provide biased diagnoses with the elderly. The discrepancies described above probably reflects differences in symptom pattern rather than real differences in severity (older people being less likely to report overt depressed mood or suicidal ideation; while being more likely to complain of intractable physical symptoms, poor memory and poor concentration).

In a review of community prevalence of depression in older adults by Beekman et al (1999), prevalence rates varied from 0.4 – 35%. Arranged according to diagnostic category, major depression was relatively rare (1.8%), minor depression more common (9.8%), while all depressive syndromes deemed clinically significant yielded an average prevalence of 13.5%. Higher prevalence rates were also found in women (14.1% for women and 8.6% for men) and among older people living under adverse socio-economic circumstances.

<table>
<thead>
<tr>
<th>Study/site</th>
<th>Sample/location</th>
<th>N subjects</th>
<th>Age (years)</th>
<th>Criteria</th>
<th>Disorder</th>
<th>Prevalence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beekman et al., 1995</td>
<td>Netherlands</td>
<td>3056</td>
<td>55-85</td>
<td>CES-D / DIS</td>
<td>Major Depression</td>
<td>2.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Minor Depression</td>
<td>12.9%</td>
</tr>
<tr>
<td>Pahkala et al., 1995</td>
<td>Finland</td>
<td>1086</td>
<td>65+</td>
<td>DSM</td>
<td>Major Depression</td>
<td>2.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Minor Depression</td>
<td>14.3%</td>
</tr>
<tr>
<td>Copeland et al., 1999 (EURODEP study)</td>
<td>9 European Centres.</td>
<td>13,808</td>
<td>65+</td>
<td>GMS-AGECAT</td>
<td>Depression</td>
<td>12.3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Males (8.6%) Fema</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>les (14.1%)</td>
<td></td>
</tr>
<tr>
<td>Steffens et al., 2000</td>
<td>USA</td>
<td>4559</td>
<td>65+</td>
<td>DIS</td>
<td>Major Depression</td>
<td>Males (2.7%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Females (4.4%)</td>
</tr>
<tr>
<td>Study</td>
<td>Site</td>
<td>N</td>
<td>Age</td>
<td>Instrument</td>
<td>Disorders</td>
<td>Prevalence</td>
</tr>
<tr>
<td>------------------------------------</td>
<td>-----------------</td>
<td>------</td>
<td>-----</td>
<td>--------------------------------------</td>
<td>----------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Ritchie et al., 2004</td>
<td>France</td>
<td>1873</td>
<td>65+</td>
<td>Mini-International Neuropsychiatric Interview DSM-IV criteria</td>
<td>Major Depression</td>
<td>3%</td>
</tr>
<tr>
<td>de Jonge et al., 2006</td>
<td>Spain</td>
<td>4803</td>
<td>55+</td>
<td>GMS-AGECAT</td>
<td>Depression</td>
<td>11.2%</td>
</tr>
<tr>
<td>McDougall et al., 2007 (MRC CFA study)</td>
<td>5 UK sites</td>
<td>2640</td>
<td>65+</td>
<td>GMS-AGECAT</td>
<td>Depression, Severe Depression</td>
<td>8.7%, 2.7%</td>
</tr>
<tr>
<td>Costa et al., 2007</td>
<td>Brazil</td>
<td>392</td>
<td>70+</td>
<td>SCAN/ICD10</td>
<td>Depressive episode</td>
<td>19.2%</td>
</tr>
<tr>
<td>Previle et al., 2008</td>
<td>Quebec</td>
<td>2798</td>
<td>65+</td>
<td>DSM-IV</td>
<td>Depression, Major depression, Minor depression</td>
<td>12.7%, 1.1%, 5.7%</td>
</tr>
</tbody>
</table>

Prevalence rates also vary by location or clinical setting. In medically hospitalised patients, estimates for major depression range between 12 and 58%, with an average of about 25% for depressive episode (Koenig et al. 1988 & 1997; Jackson and Baldwin 1993; McCusker et al. 2005). Other studies have reported rates of 33% in older adults attending their GPs (Mc Donald, 1986), 26% of users of home care services and 35 – 50% of residential home residents (Ames et al. 1988). Many of the studies described above excluded those living in institutions hence the resulting estimates are likely to underestimate the true prevalence of depression in older adults.

**Incidence**

Incidence studies of depression in older adults are rare in literature. Those that rely on DSM criteria reflect a similar pattern of decline in rates of major depression with advancing age. The 13-year follow up of the participants of the Baltimore Epidemiological Catchment Area study revealed, however, that the distribution of incidence of DSM-based major depression across the life span was bimodal, with a primary peak in the fourth decade and a secondary peak in the sixth decade (Eaton et al., 1997). The overall estimated annual incidence was 3.0 per 1000 per year and prodromal symptoms were present many years...
before the full criteria for major depression were met. Incidence studies also revealed an increased risk of depression among women as they age (Eaton et al., 1989).

1.4 Aetiology

The aetiology of depression in older adults is often multifactorial and can be summarised under predisposing, precipitating and perpetuating risk factors. These risk factors also have biologic and psychosocial elements. Developing depression results from an imbalance between predisposing, precipitating and protective factors. It is important however to point out that some factors may be more important than others in the aetiology of depression in older adults.

Predisposing factors

- Increasing levels of disability and chronic disease
- Genetic pre-disposition, early adversity and serious life events (less important)
- Risk factors associated with the stress of growing old including being single (divorced or widowed), living in large cities, poor health, functional limitations, poor social networks, loneliness (Beekman et al, 1995)
- Gender - Female sex (70:30 female: male ratio)
- Previous or family history of depression.
- Personality – “Avoidant” and “dependent” personality dysfunction is associated with depression in older adults.
- Use of certain medications – for example nonselective beta blockers, calcium antagonists, benzodiazepines and systemic corticosteroids.
- High alcohol intake in some older adults.
Vascular risk factors for depression have been suggested from studies and MRI findings suggesting that vascular lesions in the brain e.g. brain hyper intensities may contribute to a unique variety of late-life depression.

Medical illness - severe or chronic diseases associated with high rates of depression include stroke (30 - 60%), coronary heart disease (8 – 44%), cancer (1 – 40%), Parkinson’s disease (40%), Alzheimer’s disease (20 – 40 %), and dementia (17 – 31 %) [Boswell et al, 1996]. Depression in older adults is also associated with pain, urinary incontinence and head injury.

Precipitating factors

Life-events

- Bereavement
- Episode of acute ill health
- Sudden homelessness or having to move into an institution
- Major financial crisis
- Separation
- Medical illness or threat to life of someone close

Chronic stress

- Declining health and mobility; dependence
- Sensory loss, cognitive decline
- Housing problems
- Marital difficulties
- Major problems affecting family member
- Socioeconomic decline
- Problems at work; retirement
• Caring for a chronically ill and dependent family member

It is important to note that adversity alone is not sufficient to explain the development of depression in older adults. Studies looking at adversity as a cause for depression have found that around 25% of older adults exposed to adverse life events do not necessarily develop depression (Murphy, 1982).

**Protective factors**

• Early development of good coping skills
• Achievement
• Socioeconomic status
• Marital status
• Social support

**1.5 Clinical Presentation**

The clinical presentation of depression in older adults tends to vary in pattern when compared with younger adults. Older adults tend to present with somatic complaints, sleep disturbance, greater likelihood of psychotic symptoms (Brodaty et al., 1997; Katona and Shankar, 2005) and agitation leading to misdiagnosis and often unnecessary investigations. Other symptoms of depression that are common in older adults include changes in appetite and sleep patterns, loss of interest, lack of energy, social withdrawal and increased dependency (Reynolds, 1996). Incidentally, these symptoms are also common in the older adult without depression but suffering from chronic medical conditions thus making recognition of depression difficult.
The clinical presentation of depression in older adults is therefore complicated by the overlap of aging, physical illness, psychosocial antecedents such as bereavements and social isolation, and the tendency of older adults to engage in somatisation rather than report symptoms of dysphoria and worthlessness.

The main influences on the presentation of depression in older adults are summarised in the list below.

**Factors which influence how depressive disorder presents in old age (Baldwin, 2002)**

- Overlap of physical disorder with those of the somatic symptoms of depression
- Tendency of older adults to minimize a complaint of sadness and instead become hypochondriacal
- Neurotic symptoms (severe anxiety, obsessive-compulsive symptoms, hysteria) newly presenting which ‘mask’ depression
- Any act of deliberate self-harm, including medically ‘trivial’ overdoses
- ‘Pseudo dementia’ or ‘depression related cognitive impairment’
- Behavioural disturbance such as alcohol abuse or shoplifting

**Depression and dysphoria**

Older adults tend to minimize or deny feelings of depressed mood or sadness. Complaints of not coping, irritability or anxiety are more common. They will often acknowledge a loss of interest or pleasure in their daily activities as well as some sleep or appetite disturbance.

**Anxiety/agitation**

Newly arising neurotic symptoms, irritability, or importuning behaviour in older adults should always arouse a suspicion of depression. Over 80% of elderly patients with depression present with some of these symptoms (Evans & Mottram, 2000).
**Low energy, fatigue**

These complaints are common in older adults but are often attributed to physical ill health resulting in misdiagnosis.

**Cognitive difficulties**

Depression in older adults is frequently associated with subtle cognitive impairments which can persist after the depressive episode has resolved. Cognitive impairment that persists after resolution of depression often indicates a risk factor for the development of dementia in future. It is also important to note that depression can coexist with dementia.

**Somatisation & Hypochondriasis**

Somatic complaints e.g. fatigue, pain, gastrointestinal complaints, concentration difficulties or diminished energy are often features of depression in older adults. These symptoms are often attributed to their physical illnesses and medical help is often sought in the first instance, delaying treatment of the underlying depressive disorder.

In hypochondriasis, normal bodily sensations are attributed serious physical disease (e.g. cancer) with active seeking of medical investigations and reassurance.

**Behavioural disturbance**

Unusual behavioural disturbance may be a leading symptom of depression especially among those highly dependent on carers. Examples include refusing food, inappropriate urinary and faecal incontinence, screaming, and theatrical falls. (Evans & Mottram, 2000)

Late-onset alcohol abuse or, rarely shoplifting may also be indicative of depression in older adults (Baldwin, 2008).
1.6 Management

The principles of managing depression in adults (and older adults) including those with a chronic physical health problem have been updated recently by NICE in their 2009 clinical guideline. Since an increased rate of chronic physical health problems are reported in older adults, the management principles adopted in this thesis reflects those in the updated UK national depression guidelines (NICE, CG91, 2009). NICE sets out a framework within which varying levels of interventions are offered to patients depending on the severity of the depressive illness. This stepped-care model including a summary of the NICE guidelines on the management of depression is covered in detail in the next section of this thesis.

This part of the thesis will cover the essential components of the interventions including, assessment and diagnosis, psychosocial intervention and drug treatment. The clinical evidence for efficacy of the interventions described below is detailed in the recent NICE guidelines on the management of depression (NICE, CG91, 2009).

Assessment and Diagnosis

The assessment and diagnosis of depression in Older Adults requires a high index of suspicion because it is often denied by the current generation of elderly people. Depression remains a pathological process, not a normal reaction to aging, social situation or physical ill effects. Undetected or under treated, it causes prolonged suffering to the patient and increases demand on carers, social and health services.

The basic principles of biopsychosocial assessment of any patient with a psychiatric disorder apply, with attention to the clinical features described above. Depressed older people will consult their general practitioner two or three times more often than non
depressed elders, presenting opportunities to identify and treat depression (Anderson, 2001). In general hospital wards, detection rates are also poor and treatment is offered in a minority of cases.

There are many barriers to the diagnosis of depression in older adults. Signs and symptoms of depression are often attributed to ‘normal aging’, physical illness, dementia, psychosocial antecedents, or any of a number of age-associated afflictions. Other barriers to diagnosis include the fact that depression often amplifies physical symptoms, distracting patients’ and doctors’ attention from the underlying depression; and many older patients deny psychological symptoms of depression or refuse to accept the diagnosis because of stigma (see section above on clinical features).

Assessment usually begins with comprehensive history including a collateral history from relatives or carers. A comprehensive assessment of depression should not rely simply on a symptom count, but should take into account the degree of functional impairment and/or disability (NICE, 2009). Given the increased likelihood of physical problems with advancing age, a good medical and drug history is important in excluding possible organic causes for depression (see table 6). The mental state examination (including cognitive assessment), physical evaluation, and relevant investigations complete the assessment.

In an attempt to exclude organic causes, base-line laboratory tests should at least include ECG, urinalysis, general blood biochemistry screen (urea & electrolytes; liver function test; bone profile; lipid profile and random glucose), full blood count, and determination of thyroid stimulating hormone, vitamin B12 and folate levels. Other investigations that may be required if necessary include neuroimaging (CTSCAN/MRI) to exclude space
occupying lesions or in treatment resistant cases where cerebrovascular disease may be implicated in aetiology and poor response to treatment.

Screening instruments have been recommended by NICE as a means of identifying and monitoring depression in high-risk groups like the elderly. They suggest that anyone suspected of suffering from depression should be screened by two simple questions i.e. During the past month have you been bothered by 1) feeling down, depressed or hopeless, and 2) having little interest or pleasure in doing things. If the answer is ‘yes’ to either question, then they should have a comprehensive assessment by a practitioner competent in mental health assessment. The use of other objective measures such as depression rating scales has been recommended as a means of assessing severity and monitoring progress on any treatment.

Once diagnosed, depression remains treatable in majority of cases and is also the most reversible cause of psychiatric morbidity and mortality in later life (Anderson, 2001).

**Treatment**

The management of depression in older adults requires a holistic approach addressing biologic/physical, psychological and social factors. Treatment should ideally be individualized and targeted at specific problems relevant to the older adult in question. Treatment options can range from low intensity psycho-social interventions for milder forms of depression to high intensity psycho-social intervention and drug therapy for more severe forms of the illness (NICE, 2009). In clinical practice, treatment of depression occurs in various settings depending on severity. More severe forms of depression are managed in secondary care and often involve a multidisciplinary team including the
community psychiatric nurse, occupational therapists, physiotherapists, psychologists and social workers. Within primary care and acute general hospital settings other services including podiatry, speech and language therapists, counsellors and district nurses complement the input from community mental health teams. NICE recommends starting treatment with the least restrictive and most effective intervention. Drug treatment is not recommended routinely to treat sub threshold depressive symptoms or mild depression.

In general treatments that work for younger people also work for older adults except that co morbid physical illness, poly-pharmacy, and altered pharmacokinetics and different pharmacodynamics makes dosing and choice of drugs slightly different in older adults. NICE recommends prescribing at an age-appropriate dose taking into account the effect of general physical health and concomitant medication on pharmacokinetics and pharmacodynamics [(NICE, CG 90(2009)]. Monitoring carefully for side effects, such as increased risk of bleeding with SSRIs in older persons has also been recommended.

**Psychosocial treatments**

Psychosocial treatments for depression are effective and often preferable to antidepressant medication in treatment of milder forms of depression in older adults. In patients without significant cognitive impairments, the response rates to psychosocial therapy are broadly similar to those obtained in adult populations. Psychosocial treatments are broadly divided into two groups

- **Low-intensity psychosocial interventions**: include interventions such as individual guided self help based on CBT principles, computerised CBT, physical activity and peer support programmes. These are suitable for very mild forms of depression including sub-threshold cases.
• **High intensity psychosocial interventions** include individual CBT, group CBT, interpersonal therapy (IPT) or behavioural couples therapy. These are suitable for patients with persistent sub threshold depressive symptoms or those with mild to moderate depression that has not responded to low intensity psychological intervention.

**Definition of interventions (adapted from NICE, CG91 (2009)).**

• **Guided self-help**

Guided self-help (GSH) is a self-administered intervention designed to treat depression, using of a range of books or other self-help manuals based on an evidence-based intervention and designed specifically for the purpose. A healthcare professional (or para-professional) facilitates the use of this material by introducing, monitoring and reviewing the outcome of such treatment.

• **Peer (self-help) support**

Peer (self-help) support is any intervention where an individual (in groups or pairs) with a common condition (e.g. depression) or the relatives or carers of individual with a common condition meet to provide emotional or practical support to each other. Typically there is no direct professional input to the group although there may be some limited psycho-educational input. Support can be individual or group based although most interventions fall into the latter category.

• **Computerised cognitive behaviour therapy**

Computerised cognitive behaviour therapy (CCBT) is a form of cognitive behaviour therapy, which is delivered using a computer (including CD-ROM and the internet). It can be used as the primary treatment intervention, with minimal therapist involvement or as augmentation to a therapist-delivered programme where the introduction of CCBT supplements the work of the therapist.
• **Physical activity**
When used as a treatment for depression, physical activity is defined as a structured, achievable physical activity with a recommended frequency, intensity and duration (NICE CG90, 2009). It can be undertaken individually or in a group. Physical activity may be divided into aerobic forms (training of cardio-respiratory capacity) and anaerobic forms (training of muscular strength/endurance and flexibility/co-ordination/relaxation) (American College of Sports Medicine, 1980). The aerobic forms of physical activity, especially jogging or running, have been most frequently investigated. In addition to the type of physical activity, the frequency, duration and intensity should be described.

• **Cognitive behavioural therapies**
Cognitive behavioural therapies (CBT) are defined as discrete, time limited, structured psychological interventions, derived from the cognitive behavioural model of affective disorders and where the patient:
1) Works collaboratively with the therapist to identify the types and effects of thoughts, beliefs and interpretations on current symptoms, feelings states and/or problem areas
2) Develops skills to identify, monitor and then counteract problematic thoughts, beliefs and interpretations related to the target symptoms/problems
3) Learns a repertoire of coping skills appropriate to the target thoughts, beliefs and/or problem areas.

• **Couple-focused therapies**
Couple-focused therapies are time limited, psychological interventions derived from a model of the interactional processes in relationships where interventions are aimed to help participants understand the effects of their interactions on each other as factors in the development and/or maintenance of symptoms and problems.

The aim is to change the nature of the interactions so that they may develop more supportive and less conflictual relationships.
• **Problem-solving therapy**
Problem-solving therapy (PST) is a discrete, time limited, structured psychological intervention, which focuses on learning to cope with specific problems areas and where therapist and patient work collaboratively to identify and prioritise key problem areas, to break problems down into specific, manageable tasks, problem solve, and develop appropriate coping behaviours for problems.

• **Interpersonal therapy**
Interpersonal therapy (IPT) is a discrete, time limited, structured intervention that focuses on the patient’s relationships in the outside world, aiming to improve them where possible or, if more appropriate, to modify the patient’s reliance on unhelpful relationships (Jacoby et al., Oxford textbook of Old Age Psychiatry(2008). There are four possible foci in IPT: grief and loss (e.g., death of a loved one), interpersonal role disputes (e.g., conflict with adult children), role transitions (e.g., retirement) and interpersonal deficits (e.g., lack of assertiveness skills). The foci relevant to older adults are usually role transitions and role disputes.

• **Psychodynamic interventions**
Psychodynamic interventions seek to reduce symptoms by helping the patient gain insight into the ways the disorder developed in the past. Insight is achieved by bringing to conscious awareness mental contents that were previously outside consciousness, by interpreting their significance, and by linking past experiences with present modes of functioning (Gelder et al., 2006).
Summary of recommendations on the use of psychological therapies

Cognitive behavioural therapy (CBT) and interpersonal therapy (IPT) are recommended as alternatives to antidepressants in acute treatment of major depression of mild to moderate severity. Problem-solving treatment (PST), psychodynamic therapy, reminiscence therapy and bibliotherapy (guided self-help) are other forms that have been found to be effective. Psychological or behavioural therapy is not recommended as sole therapy for severe major depression (BAP guidelines, 2006; NICE, 2009), but can be adjuncts to antidepressant treatment in these cases. In a meta-analysis of studies of cognitive-behavioural, brief psychodynamic, interpersonal, reminiscence, and eclectic therapies for late-life depression, psychotherapy was found to be more effective than placebo and comparable to antidepressants although CBT had the largest evidence base (Pinquart et al., 2006; NICE CG90, 2009).

Psychological treatments are also useful in maintaining remission and preventing relapse. Compared to placebo, response rates of up to 80% was obtained for acute and maintenance treatment with a combination of interpersonal therapy and nortriptyline. Those on placebo and routine clinical visits had a 90% recurrence rate, suggesting that combined treatment using both pharmacotherapy and psychotherapy is better than treatment with either alone (Reynolds et al., 1999). In a later study by the same group addressing prevention of recurrence of depression in older adults aged on average 70 years and older, combined treatment (IPT and paroxetine) did not prevent relapse compared to treatment with either antidepressant or IPT alone (Reynolds et al., 2006).

Drug treatment

Antidepressants remain the mainstay of drug treatment for depression in older adults with response rates of up to 70 percent. They have been shown to have superior efficacy (50%
or greater reduction in symptom severity) over placebo with no particular drug emerging as particularly effective. Effect sizes are similar to those seen in depression associated with physical illness.

Studies comparing the efficacy of various classes of antidepressants including SSRIs, TCAs, MAOIs and other antidepressant agents, have found no particular difference between them. There is also no ideal antidepressant as all are associated with some form of side effect. However, because SSRIs lack significant anticholinergic, cardiovascular and sedative side-effects as well as being less toxic in overdose, they are preferred as first-line agents. A classification of the antidepressants used in treatment of depression in older adults is shown in the table below.

Table 1.3: Classification of antidepressants currently used in older adults

<table>
<thead>
<tr>
<th>Class of drug</th>
<th>Drug</th>
<th>Main side effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tricyclic antidepressants (TCAs)</td>
<td>Amitriptyline, Imipramine,</td>
<td>Sedation, anticholinergic, postural hypotension, weight gain, arrhythmia.</td>
</tr>
<tr>
<td></td>
<td>Nortriptyline</td>
<td><strong>Dosulepin and amitriptyline are most toxic in overdose</strong></td>
</tr>
<tr>
<td></td>
<td>Dosulepin</td>
<td>Lofepramine causes less side effects than older TCAs</td>
</tr>
<tr>
<td></td>
<td>Lofepramine (newer TCA)</td>
<td></td>
</tr>
<tr>
<td>Monoamine oxidase inhibitors (non-</td>
<td>Phelazine, tranylcypromine</td>
<td>Postural hypotension, dizziness, drowsiness, insomnia, headaches, nervousness</td>
</tr>
<tr>
<td>reversible)</td>
<td></td>
<td>(Use is often restricted to patients refractory to other antidepressant drugs)</td>
</tr>
<tr>
<td>Reversible inhibitors of monoamine</td>
<td>Moclobemide</td>
<td>Sleep disturbance, nausea, agitation.</td>
</tr>
<tr>
<td>oxidase A (‘RIMA’ agents)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selective Serotonin Reuptake inhibitors (SSRIs)</td>
<td>Fluoxetine, Fluvoxamine, paroxetine, sertraline, citalopram, escitalopram</td>
<td>Nausea, vomiting, dyspepsia, abdominal pain, diarrhoea, sexual dysfunction, headache. Insomnia, agitation, restlessness and anxiety may also occur.</td>
</tr>
<tr>
<td>Atypical antidepressants</td>
<td>Trazodone, mianserin</td>
<td>Sedation, dizziness &amp; headaches (Trazodone)</td>
</tr>
<tr>
<td>Noradrenaline reuptake inhibitors</td>
<td>Reboxetine</td>
<td></td>
</tr>
<tr>
<td>(NARI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serotonin/noradrenaline reuptake</td>
<td>Venlafaxine, duloxetine</td>
<td>Nausea, insomnia, dizziness (both drugs)</td>
</tr>
<tr>
<td>inhibitors (SNRI)</td>
<td></td>
<td>Somnolence, hyper- and hypotension (venlafaxine)</td>
</tr>
<tr>
<td>Noradrenaline and specific serotonin</td>
<td>Mirtazapine</td>
<td>Increased appetite, weight gain, somnolence, headache.</td>
</tr>
<tr>
<td>enhancers (NASSa)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Although SSRIs are preferable as first line agents they do, however increase the risk of GI bleeding and the updated NICE guidelines on depression recommends prescribing gastro protective agents in people at risk such as older adults. The choice of what antidepressant to use is often governed by the patient’s individual circumstances, clinical presentation and physical co morbidity. The anticholinergic side effects of TCAs such as dry mouth, urinary retention, constipation, confusional states and orthostatic hypotension can lead to severe problems in older adults particularly those with pre-existing cardiac disease. On the other hand, SSRIs and venlafaxine can inhibit the metabolism of other drugs that are co-prescribed increasing the risk of adverse events. Ultimately a balance of risk and benefit of treatment should be presented to the patient.

**Summary of NICE recommendations on the use of antidepressants**

The nature of depression in older adults without physical health problems is fundamentally the same as those with these problems. Recommendations given in this part of the thesis apply to older adults with/without physical illness.

- Antidepressants are not recommended for routine use in the treatment of subthreshold depressive symptoms or mild depression because the risk-benefit ratio is poor.

- NICE recommends the use of antidepressants in people with a past history of moderate or severe depression, sub threshold depressive symptoms or mild depression that persists after other interventions, and those with an initial presentation of sub threshold depressive symptoms that has been present for a long period.

- SSRIs in generic form should always be prescribed first, with citalopram and sertraline preferred due to their low potential for interaction with other drugs.

- Dosulepin should not be prescribed.

- TCAs should generally be avoided due to cardio-toxicity and poor tolerability.
• Other newer non-SSRI antidepressants (e.g. mirtazapine, mianserin, moclobemide, reboxetine and trazodone) can be used in treatment of depression in older adults particularly when SSRIs are not tolerated or effective.

• Antidepressants should be prescribed according to the British National Formulary recommended doses.

• Options to consider if patient fails to respond to treatment include, increasing the level of support, increasing the dose of the antidepressant up to BNF limits, switching to another antidepressant and augmenting the antidepressant with other medication.

• Treatment should continue for at least 6 months after remission. A review of the need for continued treatment should take into account the patients previous history of depression and any concurrent physical and psychosocial difficulties.

**Electroconvulsive treatment (ECT)**

ECT is effective in the treatment of severe and complex depression in older adults and the safety profile is reported as very good (NICE, 2009)

ECT should be considered for acute treatment of severe depression that is life threatening and when a rapid response is required, or when other treatments have failed.

ECT is still not recommended as a routine treatment for moderately severe depression, it is presented as an option in those with moderate depression who have repeatedly not responded to both drug and psychological treatment.

**1.7 Prognosis**

The outcome of depression in older adults has been documented extensively in literature and what is widely reported is that over longer periods of follow up, depression tends to run a chronic remitting course (Blazer, 2003). Results from a meta-analysis of outcomes of depression in older adults in community settings indicated that over a mean follow-up of
12 and 38 months, only 19-34% of depressed older adults remained well with 27% being continuously ill (Cole and Bellavance, 1997). Most of those remaining (21%) had died. In another more recent review comparing the prognosis of depression in old age versus middle age, poorer outcomes were reported in older people in terms of higher chances of relapse during follow up compared to middle aged adults (Mitchell and Subramaniam, 2005). Prognostic factors are generally divided into general factors and those relevant to the characteristics of the depressive illness. General prognostic factors include chronic stress associated with poor environment, crime and poverty, becoming a victim of crime, poor perceived social support and supervening physical illness. Poor outcomes are also related to features of the depressive illness such as slow initial recovery, more severe initial depression, depressive symptoms lasting two years or more, three or more previous episodes of depression, and the presence of organic cerebral pathology (Jacoby et al., 2008).

A positive link has also been reported between depression and non suicide mortality with a trend for the effect to be greater in males than females (Schulz et al., 2002). Factors predicting increased mortality in depressed older adults include duration, chronicity and severity of depressive symptoms (Geerlings et al., 2002). In another study, this link between depression and non suicide mortality disappeared after controlling for a number of confounding factors such as chronic disease, health habits, cognitive impairment, functional impairment and social support (Blazer, 2001). This association between depression and non suicide mortality may therefore be determined by the presence of those variables.

Other adverse outcomes associated with depression include poor physical health, increased use of health services and suicide. The association between depression and physical ill health is discussed in the next section of this thesis.
2. DEPRESSION AND PHYSICAL ILLNESS IN OLDER ADULTS

2.1 Prevalence

Depression is approximately two to three times more common in people with a chronic physical health problem than in people who have good physical health (NICE CG91, 2009). A number of studies have documented the prevalence of major depression in elderly medical inpatients to range from 6-45% (see table 2.1 below); with an additional 15-25% experiencing clinically significant symptoms not captured by the diagnosis of major depression. The same is true for estimates of prevalence in long-term care facilities (major depression [5-15%]; clinically significant depressive symptoms [additional 30%]). The reasons for such variability in prevalence rates is shown in the table below and reflects differences in diagnostic methods used to identify the disorder. Studies using structured and validated diagnostic interviews especially those validated in older people tend to reflect a more accurate indication of the prevalence and this often narrows to 6-25 % (Katona and Shankar, 2004). Despite this high prevalence, detection rates for depression remains very low at between 9 and 55 % (Rapp et al., 1988; Loke et al., 1996) and even when detected, it is often undertreated or not referred to specialist psychiatric services for intervention (Anderson and Philpott, 1991).

Table 2.1
Prevalence of depression in Older Adults in selected treatment/inpatient facilities

<table>
<thead>
<tr>
<th>Study/site</th>
<th>Reference</th>
<th>Sample size</th>
<th>Mean age</th>
<th>Measurement</th>
<th>Disorder</th>
<th>Prevalence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute medical inpatients</td>
<td>Koenig et al., 1986</td>
<td>171</td>
<td>70+</td>
<td>Screening plus modified DIS</td>
<td>Major depression</td>
<td>11.5%</td>
</tr>
<tr>
<td>(USA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Minor depression</td>
<td>23.0%</td>
</tr>
<tr>
<td>Acute medical inpatients</td>
<td>O’Riordan et al., 1989</td>
<td>111</td>
<td>80</td>
<td>GDS &amp; DSM-III</td>
<td>Major depression</td>
<td>4.5%</td>
</tr>
<tr>
<td>(Ireland)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Dysthymic disorder</td>
<td>3.6%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significant depressive symptoms</td>
<td>10.8%</td>
</tr>
<tr>
<td>Setting</td>
<td>Reference</td>
<td>Sample Size</td>
<td>Age</td>
<td>Instrument</td>
<td>Diagnoses</td>
<td>Prevalence</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>--------------------</td>
<td>-------------</td>
<td>-----</td>
<td>---------------------</td>
<td>------------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Long-term care facility</td>
<td>Parmelee et al., 1989</td>
<td>708</td>
<td>84</td>
<td>DSM-III-R checklist</td>
<td>Major depression Minor depression</td>
<td>12.4% 30.5%</td>
</tr>
<tr>
<td>Acute medical inpatients (UK)</td>
<td>Burn et al., 1993</td>
<td>100</td>
<td>82</td>
<td>GMS/AGECAT</td>
<td>Depressive psychosis(major depression) Depressive neurosis</td>
<td>6% 17%</td>
</tr>
<tr>
<td>Medical inpatients (USA)</td>
<td>Blank et al., 1997</td>
<td>150</td>
<td>80</td>
<td>DIS</td>
<td>Major depression + Subsyndromal depression</td>
<td>11%</td>
</tr>
<tr>
<td>Medical inpatients (USA)</td>
<td>Koenig et al., 1997</td>
<td>460</td>
<td>60+</td>
<td>DIS / DSM-IV</td>
<td>Major Depression Minor Depression</td>
<td>10-21% 14-25%</td>
</tr>
<tr>
<td>Hospitalized patients on emergency wards (France)</td>
<td>Clement et al., 1999</td>
<td>120</td>
<td>79.8</td>
<td>ICD-10</td>
<td>Depressive disorder</td>
<td>58%</td>
</tr>
<tr>
<td>Inpatients in General Hospitals (Germany)</td>
<td>Schneider et al., 2000</td>
<td>260</td>
<td>60+</td>
<td>ICD-10</td>
<td>Major depression Minor depression</td>
<td>14.1 – 35.5% 17.6%</td>
</tr>
<tr>
<td>Rehabilitation inpatient, day patient (UK)</td>
<td>Pomeroy et al., 2001</td>
<td>87</td>
<td>78.4</td>
<td>ICD 10</td>
<td>Depression</td>
<td>19.5%</td>
</tr>
<tr>
<td>Acute care facility (Canada)</td>
<td>McCusker et al., 2005</td>
<td>1686</td>
<td>65+</td>
<td>DIS / DSM-IV</td>
<td>Major depression Minor depression</td>
<td>14.2 -44.5% 7.9-9.4%</td>
</tr>
<tr>
<td>Acute medical inpatients(UK)</td>
<td>Cullum et al., 2006</td>
<td>223</td>
<td>65+</td>
<td>ICD-10 GDS</td>
<td>Depressive disorder</td>
<td>17.7%</td>
</tr>
<tr>
<td>Medical inpatients (Norway)</td>
<td>Helvik et al., 2010</td>
<td>484</td>
<td>80.7</td>
<td>HAD</td>
<td>Depression (&gt;8 HAD-D)</td>
<td>10.3%</td>
</tr>
<tr>
<td>Medical inpatients (Greece)</td>
<td>Michopoulos et al., 2010</td>
<td>200</td>
<td>65+</td>
<td>SCID</td>
<td>Major depressive episode</td>
<td>14%</td>
</tr>
</tbody>
</table>

**Measurement:** BAS-DEP, Brief Assessment Schedule Depression Part; DIS, Diagnostic Interview Schedule; GMS, Geriatric Mental State Schedule; GDS, Geriatric Depression Scale; HAD, Hospital Anxiety and Depression Scale; DSM III/IV, Diagnostic and Statistical Manual, 3rd or 4th edition; ICD-10, International Classification of Diseases and Related Health Problems 10th Revision.

### 2.2 Factors contributing to the increased risk of depression in physical illness

The increased risk of depression in physically ill can be explained by the following factors:

- Mood disturbance can arise from structural brain disease which is more common in older adults
• Depression could be caused by drugs prescribed for physical conditions. The elderly are particularly susceptible to side effects of drug therapy as they are subject to polypharmacy (see table 2.2)

• Drug induced depression may precipitate depression by affecting the levels of available neurotransmitters

• The combination of a susceptible patient and a depressogenic drug may precipitate depression severe enough to lead to suicide (see table 2.2) (Abou-Saleh et al., 2011)

• Chronic physical illness, physical function and cognitive function all independently predict depressive morbidity in late life

• Inadequately treated depression can lead to cardiovascular complications from the medication but no benefit to the patient leading to increased morbidity and mortality (Murphy et al., 1988)

• Physical illness can cause impairments in mobility leading to social isolation and difficulty maintaining relationships which can in turn lead to depression

• Depression can exacerbate symptoms of coexisting physical illnesses such as pain from arthritis

• Neurodegenerative disorders that present to other medical specialists (e.g. Parkinson’s disease) have a higher prevalence of depression than other physical illnesses. It is also important to note that the physical appearance of patients with Parkinson’s disease can be confused with depression

2.3 Diagnostic difficulties

Diagnosing depression in older adults with physical illness can be difficult because the clinical presentation of physical illness resembles those of depression. Depression can present with somatic symptoms, the physical illness can cause depression, and both may
even coexist in the same patient. In assessing older adults with physical illness who may be suffering from depression, it is important to focus on the cognitive symptoms of depression rather than somatic symptoms. Other covert manifestations of depression include poor compliance or refusal of essential medical treatment (Mac Hale, 2002).

Having a high index of suspicion is therefore necessary to help improve rates of detection. The National Institute for Health and Clinical Excellence recommends screening for depression with two questions (discussed in detail in another section of this thesis) and referring to specialist mental health services if an affirmative answer to either question is given by the patient. Table 2.2 summarises medical conditions and drugs commonly associated with depression.

<table>
<thead>
<tr>
<th>Table 2.2 Medical conditions and central-acting drugs that may cause organic depressive disorder (Oxford textbook of Old Age Psychiatry, 2008, OUP)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Medical Conditions</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><em>Organic brain disease</em></td>
</tr>
<tr>
<td><em>Occult carcinoma</em></td>
</tr>
<tr>
<td><em>Chronic infections</em></td>
</tr>
<tr>
<td><strong>Central-acting drugs</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
2.4 Management

Principles of management of depression in older adults with physical illness
(Adapted from the recent UK guidelines on the management of depression in people with chronic physical conditions, NICE, 2009)

When depression is co morbid with a physical illness, it can be difficult to recognise because the symptoms of depression resemble those of the physical disorder. This may lead the practitioners involved to look for physical causes and ignore the depressive illness. The older adult who is physically ill is even more difficult to identify. Firstly, they often do not open up to being depressed and even when they do, the symptoms are atypical and sometimes attributed to the physical condition, loneliness or other social circumstances. Older adults now occupy two-thirds of general hospital wards and up to a third of these are depressed. It is therefore important that medical teams looking after older people are able to identify depression quickly and refer on to specialists. It is clear from literature that coexistent depression and physical illness worsens the prognosis of both, and this is particularly pertinent in older adults who often suffer from multiple physical conditions. The recent UK depression guidelines sought to address this issue by issuing separate guidance covering the management of depression in the physically ill. Aspects that are relevant to older adults are summarised below.

Stepped Care Model

NICE proposes a framework where the least restrictive and most effective intervention is offered to a patient first before going on to offer other interventions from other steps if the previous step is found to be unsuitable for the patient. There are four steps in this model covering assessment and diagnosis and treatment options for all the severities of depression as defined by the DSM-IV criteria.
Step 1 (Identifying Depression and initial management in primary care and General Hospital settings)

Focus of intervention:

Anyone with a chronic physical condition suspected of suffering from depression (including those with a past history of depression).

Intervention

- Screen for depression by asking these 2 questions: During the past month have you been bothered by 1) feeling down, depressed or hopeless, 2) having little interest or pleasure in doing things.

- Answering ‘yes’ to either question should lead to further assessment by an appropriate professional.

- Appropriate professional should assess the patient further taking into account the role of the physical illness, the degree of functional impairment and duration of the episode. The professional should not rely on symptom count to make a diagnosis.

- Intervention should include support, psycho education, active monitoring by specialist teams and referral for further intervention in step 2 if necessary.

Step 2 (Recognised depression (milder forms) in primary and general hospital settings)

Focus of intervention:

People with physical health problem who have persistent depressive symptoms not severe enough to be captured by diagnostic criteria (i.e. fewer than 5 symptoms) or mild to moderate depression.
**Intervention**

- **Active monitoring** for those who may recover without any intervention or those with sub threshold or mild depression who do not want any intervention.

- **Low intensity psycho social interventions**

  1. Individual guided self-help based on the principles of cognitive behavioural therapy (CBT).

  2. Computerised CBT (CCBT).

  3. A physical activity programme (modified for the particular physical health problem).

  4. A peer support programme in a group of patients with a shared physical health programme.

- **Advice on sleep hygiene** should be offered

- **Drug treatment** should only be considered if the patient has

  1. Past history of moderate or severe depression, or

  2. Mild depression that complicates the care of the physical health problem, or

  3. Initial presentation of sub threshold depressive symptoms that have been present for at least 2 years, or

  4. Sub threshold depressive symptoms or mild depression persisting after other interventions.
Step 3 (Recognised depression in primary care and general hospital setting)

Focus of intervention:

People with persistent sub threshold depressive symptoms or mild to moderate depression with inadequate response to initial interventions, and moderate to severe depression.

Intervention

- **persistent sub threshold depressive symptoms or mild to moderate depression with inadequate response to initial interventions**
  
  1. an antidepressant (usually an SSRI), or
  
  2. High intensity psychological intervention (group CBT, individual CBT or behavioural couples therapy).

- **moderate depression**
  
  1. High intensity psychological intervention (group CBT, individual CBT or behavioural couples therapy).

- **severe depression and a chronic physical health problem**
  
  1. Combination of individual CBT and an antidepressant.

Step 4 (Severe and complex depression usually in specialist mental health services)

Focus of intervention:

People with severe and complex depression who may be at risk to life or self-neglect

**Intervention** (may include any of the following with close collaboration with services treating the physical health problem).

- **Medication** (Usually antidepressants)

- **High intensity psychological interventions**
• Electro-convulsive therapy and other physical treatments

• Combined treatments

• Crisis resolution and Home treatment

Summary

Research has shown that depression is treatable in older adults including those who are physically ill. The key challenge in the management of depression in these groups remains identifying those who require antidepressant treatment. Many symptoms of depression are also symptoms of a physical illness and patients may not see the relevance of antidepressant treatment to their physical condition. The problem of medical comorbidities leads to poly-pharmacy with the potential for drug interactions. It is also important to note that depression may be a sign of a physical illness or arise secondary to medication prescribed for a physical condition (see table 2.2). Provided all physical causes have been ruled out or managed appropriately, treatment of depression is essentially the same as for any older adult.

2.4 Outcomes

Older adults now occupy a significant proportion of beds in general hospitals. A recent report from the Royal College of Psychiatrists(2005) reports that in a typical district hospital 60% of the beds will be occupied by older people, 29% of whom will have depression. Depression is frequently associated with chronic medical illness and can in turn adversely affect the outcome of co morbid physical problems such as cardiovascular disease, osteoporosis, stroke, cancer, chronic lung disease, dementia and Parkinson’s disease (Krishnan et al., 2002). The association between depression and physical illness is bi-directional with physical illness also capable of provoking depression, leading to
increased disability associated with the original physical disorder (Prince et al., 1998; NICE, 2009).

Several studies report adverse outcomes for older medical inpatients with depression including delayed discharge, greater risk of readmission, increased need for rehabilitation or institutional care and higher mortality. Cole et al carried out a meta-analysis on the outcomes of depression in elderly medical inpatients. The combined results of the studies indicated that over a mean follow up of 1.5 and 18 months, 18% to 19% of patients were well, 29% to 43% remained depressed, and 22% to 53% had died. These outcomes were much worse than those of elderly patients in hospital-based psychiatric services (60% were well or had had relapses with recovery, and 14% to 22% were continuously ill). Factors associated with worse outcome included more severe depression, more serious physical illness and symptoms of depression before admission. Low rates of detection and under treatment may partly explain these poor outcomes (Baldwin, 2000).

The cost of depression to older people and their carers is substantial and the cost to the NHS considerable thus increased attention to detection and treatment of depression in medical settings may improve outcomes and reduce costs to the NHS.
3. SCREENING FOR DEPRESSION

The National Screening Committee (NSC, 1998) defines screening as “the systematic application of a test or inquiry, to identify individuals at sufficient risk of a specific disorder to warrant further investigation or direct preventative action, amongst persons who have not sought medical attention on account of the symptoms of that disorder.”

Screening tests are designed to identify the possibility that disease or disorder might be present and to prompt further evaluation in those who screen positive. A screening test should therefore be regarded as only one possible first stage of the diagnostic sequence and should lead to effective and early intervention such as treatment with antidepressants or psychological therapy in the case of depression.

Screening tests should be easy to administer, acceptable to patients, have high sensitivity (i.e., identify most of the individuals with the disease), identify a treatable disorder and identify a disorder where intervention improves outcome (Wilson & Junger, 1968).

3.1 Requirements of screening

For screening to be useful, the condition must be sufficiently common in the target group to warrant the screening. As depression is a common disorder in older adults and can cause significant morbidity and mortality if missed or left untreated, screening or case-finding scales are useful.

The UK National Screening Committee gives certain criteria for evaluating the viability of a screening tool or programme (table 3.1).
Table 3.1 UK National Screening Committee criteria for appraising the viability, effectiveness and appropriateness of a screening programme

<table>
<thead>
<tr>
<th>The Condition</th>
<th>The screening test</th>
<th>The treatment</th>
<th>The Screening programme</th>
</tr>
</thead>
<tbody>
<tr>
<td>-The condition should be an important health problem.</td>
<td>-should be safe, simple, precise, and validated; suitable cut-off should be defined and agreed.</td>
<td>-an effective treatment should be identified through the screening programme, with evidence that early treatment leads to better outcome.</td>
<td>-High quality RCTs should provide evidence that the screening programme effectively reduces morbidity.</td>
</tr>
<tr>
<td>-The epidemiology and clinical course of the disease should be adequately understood.</td>
<td>-should be acceptable to the population - should have high specificity (low rate of false positives) and a very high sensitivity (very low rate of false negatives), although this is difficult to assess when evaluating a screening tool for depression. - must be relatively cheap or else the cost per case detected is prohibitively expensive.</td>
<td>-Clinical management of the condition and patient's outcomes should be optimised for all healthcare providers before the screening programme is offered.</td>
<td>-the screening programme should be clinically, socially, and ethically acceptable to health professionals and the public.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-the benefit from screening should outweigh the physical and psychological harm</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-the cost of the screening programme should be economically balanced in relation to expenditure on medical care.</td>
</tr>
</tbody>
</table>

3.2 Why screen for depression?

The national service framework for older people (Department of Health UK, 2001) placed emphasis on the detection and treatment of depression in older people. In 2004, NICE released the first edition of its clinical guideline on the management of depression and highlighted the importance of screening high-risk groups in primary care and general hospital settings (NICE, 2004). The recently updated guideline continues to recommend that clinicians screen for depression in people with a past history of depression or a chronic physical health problem with associated functional impairment (NICE, 2009). In the United States, its Preventative Services Task Force (USPSTF) also issued depression screening recommendations based on a systematic review by Pignone and his colleagues in 2002. They encouraged primary care physicians to routinely screen their adult patients for depression “in clinical practices that have systems in place to assure accurate diagnosis, effective treatment, and follow-up of recognised cases” (grade B recommendation). The Canadian Task Force on Preventative Health Care made a similar recommendation in 2005.
after reviewing the available evidence at the time. They found fair evidence to recommend screening adults in the general population for depression in primary care settings with integrated programs for feedback and access to case management or mental health care (grade B recommendation).

In comparison with the aforementioned reviews of depression screening, a Cochrane review in 2005, which excluded all studies that included “complex quality improvement/care management” strategies, concluded that screening programmes were not effective in improving outcomes (Gilbody et al., 2005). Complex enhanced collaborative care, not screening alone, was found to improve outcomes of depression (Gilbody et al., 2006).

Further reviews have followed with divergent views on the benefits of depression screening. A meta-analysis in 2008 that focused on screening alone, found no substantial effect of screening or case-finding instruments on the overall recognition rates of depression, the management of depression by clinicians or on depression outcomes (Gilbody et al., 2008). These findings were true in both primary care and general hospital settings. In 2009, the USPTF updated the 2002 recommendation for screening adults in primary care and found limited evidence to support screening and feedback without further care support such as assessment and monitoring by specialist mental health services (O’Connor et al., 2009).

Although UK guidelines recommend screening of high risk groups, it goes a step further by recommending that it should be the first stage of a process that leads to further assessments by specialists in mental health. This recommendation is consistent with the Canadian and United States recommendation on screening for depression.
Unfortunately, detection rates by non-psychiatric physicians in primary care, nursing homes or general hospitals are low with 35%-50% of cases not recognized (Whooley et al., 1997). Koenig et al., (1988) found that 80% of depression in medically ill elderly patients was not detected by doctors caring for them. The reasons for the low detection rates have been attributed to the fact that many older adults present with somatic complaints and experience symptoms of depression that do not meet the criteria for depressive disorders. There is also a lack of awareness of the manifestations of depression in older adults by non-psychiatric medical staff.

Encouraging the use of screening tests has been advocated as a means of increasing the awareness of depression among medical colleagues and ensuring that depressed patients are identified and receive appropriate treatment (NICE, 2009). Although several screening tests for depression exist, not all are suitable for use in older adults particularly in acute hospital settings where physical co morbidities are present. Identifying a valid screening tool that is easy to administer within a short time and which is acceptable to the older adult is thus a challenge. If there was one screening instrument for depression that was valid in a variety of settings in the general hospital then this may be of great benefit - health professionals are likely to become more familiar and comfortable with one tool thereby increasing their alertness.

3.3 Depression Screening Instruments in Older Adults

Numerous screening instruments have been developed to detect depression but only a few have been specifically designed to detect depression in older adults. The differences in symptom patterns among older adults with depression means that scales designed for adults may not be suitable for detecting depression in this age group.
Depression screening instruments provide an indication of the severity of symptoms and assess the severity within a given period of time (e.g., the past seven to 14 days). Instruments can be self-reports, interviews by trained or non-trained staff and informant-based measures. Each screening instrument usually has a unique scoring system with higher scores consistently reflecting more severe symptoms. All measures have a statistically predetermined cut-off score at which depression symptoms are considered significant. Some measures group scores into different levels of symptom severity (Sharp et al., 2002).

Considerations for selecting a measure include characteristics of the population to be screened, psychometric properties of the instrument, time required to complete the measure, time required to score the measure, ease of use, and cost of obtaining the measure (Sharp and Martin, 2002) (see Table 3.1 above). In choosing an appropriate depression screening measure in older adults, it is important to consider levels of cognitive or sensory impairment because majority of screening tools have not been validated in patients with dementia (Kafonek et al., 1989).

The following questions about the characteristics of any screening instrument should be asked before deciding to use one:

1. How sensitive is the test or screening instrument?
   
   A sensitive test is one in which a positive result identifies most people with the target disorder. The people with the disorder who test positive on the screening test are called true positives.

2. How specific is the test or screening instrument.
A specific test is one in which a negative result identifies most people without the disorder. The people without the disorder who test negative on the screening test are called true negatives.

3. What is the positive predictive value?
Clinicians need to know the significance of a positive test result i.e. what proportion of subjects who test positive actually have the disorder? The proportion of subjects with a positive test result who have the disorder is called the positive predictive value (PPV). The PPV is not a characteristic of the screening instrument and tends to increase when the prevalence of a condition is high.

4. What is the negative predictive value?
Similarly, when a diagnostic test is negative, it is useful clinically to know what proportion of those with a negative test result is free from the disorder. This is called the negative predictive value (NPV).

5. What cut-off or diagnostic threshold should be used?
As previously mentioned, most depression screening tools have some predetermined cut-off point at which depressive symptoms are considered important. When deciding what threshold to use, it is crucial to consider the cost of false negatives or false positives to those suffering from the disorder. For example, in the case of depression in older adults there is an established evidence to support the fact that if untreated, patients have poor prognosis and increased mortality. There are also well established treatments available that are safe and inexpensive. In this circumstance, a highly sensitive test would be more important than a highly specific test.
Screening instruments specific for the elderly may be particularly useful in combating under-detection and under-treatment of depression in old age (Shulman et al., 1989; Katona & Shankar, 2005). In a recent systematic review by Watson and Pignone (2003), the accuracy of depression screening instruments for detecting depression in older adults in primary care was evaluated. The test characteristics of the seven instruments for detecting major depression are summarised in table 3.2. All studies used a criterion standard as comparison (ICD-10 or DSMIIIIR, DSM-IV) and the screening instruments reviewed had sensitivities of 67-100% and specificities of 53-98% for detecting major depression.

<table>
<thead>
<tr>
<th>Instrument (no of studies)</th>
<th>Description (cut points)</th>
<th>sensitivity</th>
<th>specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geriatric Depression Scale (9)</td>
<td>primarily 15 items, Y/N (3-5)</td>
<td>79-100%</td>
<td>67-80%</td>
</tr>
<tr>
<td>Centre for Epidemiologic Studies Depression Scale (5)</td>
<td>20 items ranking symptom frequency (9-21)</td>
<td>75-93%</td>
<td>73-87%</td>
</tr>
<tr>
<td>SelfCARE(D)(3)</td>
<td>12 item Likert scale(5)</td>
<td>77-90%</td>
<td>53-98%</td>
</tr>
<tr>
<td>Caribbean Culture-Specific Screen (2)</td>
<td>Number of items not reported(5-6)</td>
<td>82-92%</td>
<td>68-79%</td>
</tr>
<tr>
<td>Cornell Scale for Depression in Dementia (1)</td>
<td>19 item(7)</td>
<td>90%</td>
<td>75%</td>
</tr>
<tr>
<td>1-question screen (from Mental Health Inventory of the SF-36)(1)</td>
<td>1-item, 6 point scale(2)</td>
<td>67%</td>
<td>60%</td>
</tr>
<tr>
<td>Brief Assessment schedule Cards (1)</td>
<td>Number of items not reported(6)</td>
<td>92%</td>
<td>84%</td>
</tr>
</tbody>
</table>

The reality is that in everyday clinical practice, these screening tests are not used routinely in primary care or in acute hospital settings because of time constraints and there are even suggestions that very brief instruments such as the following two questions[ (1) "During the last month(2 weeks), have you often been bothered by feeling down, depressed, or hopeless?" and (2) "Have you been bothered by having little interest or pleasure in doing things?"], appear to perform as well as longer instruments(NICE, 2004; Whooley et al., 1994). According to the USPSTF and NICE, an affirmative response to the following two questions may be as effective as using longer screening measures or may indicate the need
for the use of more in-depth diagnostic tools. This “two-question screening” remains to be validated in older adults.

For the purpose of this thesis, the Geriatric Depression Scale (GDS), the Edinburgh Depression Scale (EDS) and the two NICE questions will be discussed in some detail.

**Geriatric Depression Scale**

The Geriatric Depression Scale (GDS) (Yesavage et al., 1982) was specifically developed to address the problems associated with using depression rating scales designed for adult populations in elderly patients. These rating scales often contain items that fail to discriminate the patterns of depressive symptoms from the general characteristics of the elderly population. The GDS screens for characteristics of depression in the elderly such as somatic concern, lowered affect, cognitive impairment, feelings of discrimination, impaired motivation, lack of future orientation, and lack of self-esteem. From an original list of 100 items, 30 questions addressing various symptoms of depression were extracted to create the 30-item version. The GDS is a self reported or interviewer administered inventory with a simple yes/no format that may be more acceptable in the elderly population. It does not contain somatic symptoms or other questions that might cause problems for older people. Because the GDS-30 is relatively time consuming, a shortened version containing 15 questions (GDS-15) was developed from the original 30-item version (Sheikh and Yesavage, 1986). There are also 10 item, 5 item and 4 item versions available. The GDS remains one of the most widely used instruments for screening depression in older adults.

The GDS (15 and 30 items) has been found to perform well among medically ill inpatients (Jackson and Baldwin, 1993) and has been validated among geriatric patients in various
settings and against various external case criteria/gold standards such as DSM III/IV, ICD 10, RDC and AGECAT. A recent systematic review of the screening accuracy of the GDS showed that for both GDS versions, similar validity indices were found (GDS-30: sensitivity 0.753, specificity 0.770 [cut-off value 10-11]; GDS-15: sensitivity 0.805, specificity 0.750 [cut-off value 5-6]). The internal consistency of the GDS-15 has been evaluated by a few studies and most have reported reliability values around 0.80 (Liu et al., 1998; D’Ath et al., 1994).

The GDS has been endorsed by the Royal College of General Practitioners for use in primary care (Williams and Wallace, 1993) as well as by the Royal College of Physicians and the British Geriatric Society (1992) for hospital use.

One of the disadvantages of using the GDS in hospitalized older adults is the fact that it has not been well validated in patients with significant cognitive impairment. The proportion of older adults with cognitive impairment and other sensory impairments in general hospital settings continue to rise as people live longer. Previous studies support the view that the GDS is not accurate when applied to persons with marked cognitive impairment. Kafonek et al., found the GDS to have sensitivities as low as 34% among patients with Alzheimer’s disease and 47% in a mixed group that included cognitively impaired participants.

The Edinburgh Depression Scale

The Edinburgh Depression Scale (EDS) (Cox et al, 1990) is a self-report questionnaire consisting of 10 items corresponding to non-somatic symptoms of depression. It was initially developed and validated to screen for depression in women following childbirth. Since then, the EDS has been further evaluated using community samples in the UK and
in a North American population. It has also been validated in women outside the postnatal period, in cancer patients and found to have acceptable levels of sensitivity and specificity (cancer patients: sensitivity 0.81, specificity 0.79 [cut-off value 13]; community UK sample: sensitivity 0.67, specificity 0.95 [cut-off value 12.5]).

The test can usually be completed in less than 5 minutes. Responses are scored 0-3 according to increased severity of the symptom. The total score is determined by adding together the scores for each of the 10 items. Validation studies have utilized various threshold scores in determining which patients were positive and in need of referral. Cut-off scores ranged from 9 to 13 points.

**Short screening instruments**

In a recent meta analysis by Mitchell and Coyne, screening tools were classified into 3 groups based on the length of the screening instrument and the time taken to complete the questionnaire. Ultra short screening instruments were defined as those with 1-4 items, like the GDS-4 and NICE-2, which take less than 2 minutes to complete.

Short screening instruments like the GDS-4 and the 2-item screen recommended by NICE (see Appendix 5) have been advocated as a means of ensuring that depressed people are identified and referred for appropriate treatment. The 2 questions suggested by NICE are: “During the past month, have you often been bothered by feeling down, depressed or hopeless?” and, “During the last month, have you been bothered by having little interest or pleasure in doing things?” These 2 questions were originally part of a screening questionnaire (the Primary Care Evaluation of Mental Disorders- PRIME-MD) designed to facilitate the diagnosis of common mental disorders in primary care. The theory behind using these 2-items as a method of screening for depression comes from the
DSM-IV criteria for a major depressive episode which recognises 2 core symptoms i.e. depressed mood and loss of interest or pleasure as been essential symptoms of depression. It is easy to see the attraction towards using shorter screening instruments in older adults with physical problems. In primary care and on general hospital wards where detection rates remain low, the use of these short screening instruments which take less than a minute to complete would appeal to practitioners and patients who may find longer screening measures cumbersome and time consuming for routine use. Although NICE suggested that these two questions were useful and probably sufficient enough to screen for depression in primary care and general hospital settings, the validity of this method of screening has not been fully established. The evidence used by NICE in recommending the 2-item screening instrument came from a paper by Whooley et al. (1997). In this study, the 2–item screening tool was found to have a sensitivity of 96% and specificity of 57% when compared to the Quick Diagnostic Interview Schedule (cut-off was “yes” to one of these two questions). In another study evaluating the 2 questions in verbal form, sensitivity of 97% and specificity of 67% was achieved compared with the computerized composite international diagnostic interview (Arroll et al., 2003). The accuracy of these short screening tools was further evaluated in a recent meta-analysis by Mitchell & Coyne in 2007. Overall sensitivity and specificity for two- and three- item tests on pooled analysis was 73.7% and s 74.7% respectively. Although the levels of sensitivity from these studies appear acceptable, there were a significant numbers of false-positive cases but very few false-negatives.

The conclusion to be drawn from these studies is that although the use of ultra-short screening tools may appeal to primary care physicians, they tend to be better at ruling out depression and are useful first steps in identifying cases that may be depressed. Screening patients with ultra-short screening tools must be accompanied by systems to ensure monitoring and follow up of positive cases.
4. SUMMARY OF THE INTRODUCTION

The literature from which this introduction was synthesized was sourced from the electronic databases MEDLINE, PsychINFO, Web of Science as well as Cochrane database of systematic reviews. Searches were limited to over 65s, English language citations and publication year 1996 – 2011. The search terms include various combinations of the following: depression, depressive disorder, major depression; prevalence, epidemiology; hospitals, hospitalization, hospitalized patients, inpatients; mass screening, sensitivity and specificity; aged, older or elderly. Expert reviews including UK, US and Canadian clinical guidelines on depression and screening were also included. Relevant articles were selected from the abstracts and the references of these articles were searched for additional references. It has to be highlighted that the purpose of the literature review was to look at what was known on the subject of the thesis in order to generate some rationale for the study. It was not intended to be a systematic review.

From the preceding review, it is clear that depression in older people who are physically ill is common yet undiagnosed and undertreated. Mental disorder in this population is an independent predictor of poor outcome. These poor outcomes include increased mortality, greater length of stay, loss of independent function and higher rates of institutionalisation. The cost of these disorders to sufferers and carers is substantial and the cost to services considerable.

There is some evidence to support screening high risk groups such as those with chronic physical illness (NICE 2004, NICE 2009), but only when used within a collaborative care system (Gilbody et al. 2006). In other words screening for depression can improve recognition of depressed patients and lead to a reduction in risk associated with depression
provided it is a first stage process that is linked to further assessments by mental health specialist (Pignone et al, 2002; NICE 2004; NICE 2009).

The use of a brief, simple screening tool which does not rely on symptoms of depression that could be attributable to their physical illness should improve the low detection rates of depression in older people who are physically ill.

Validating the EDS for use in physically ill patients would clearly be beneficial in hospital settings as it could mean using a single instrument in a variety of settings in general hospitals, thereby making health professionals more familiar and comfortable with one tool. The 2 questions advocated by NICE for use as a screening instrument in the 2004 and recently updated 2009 depression guidelines, has not been validated in older people in General Hospital settings.

4.1 Rationale for the study aims

In light of the above, this study seeks to address the following questions as set out in the aims and objectives

- How does the EDS perform as a screening tool for depression in older people who are physically ill when compared with a gold standard test?
- How does the performance of the EDS compare with an established screening measure of depression in older adults including the 2 questions suggested by NICE in their depression guidelines?
SECTION TWO: METHODOLOGY
5. METHODOLOGY

5.1 Aims

1. The primary aim of this study was to establish whether the Edinburgh Depression Scale is a valid tool for screening for depression in older people who are physically ill in hospital.

2. As a secondary aim, the study also sought to compare the performance of the EDS to: the GDS-15 and two brief screening tools – the GDS-4 (a component of the GDS-15) and the two NICE questions in terms of sensitivity, specificity, negative predictive value and positive predictive value. Analysis using the Receiver Operating Characteristics (ROC) curves will enable comparison of the various scales and indicate the most appropriate cut-off in this sample of patients. The clinical utility /performance of the screening tools in screening for milder forms of depression will also be examined.

5.2 Location and samples

The study took place at the Leicester General Hospital, UK; which is one of three NHS teaching hospitals that make up the University Hospitals of Leicester NHS Trust (UHL). It is located about 3 miles east of the Leicester city centre and provides a wide range of medical services including geriatric medicine. It hosts the headquarters of UHL and has around 680 beds.

Subjects were recruited randomly from 4 acute medical teams whose patients were admitted into one of four wards at the Leicester General Hospital between October 2006 and June 2010. During the course of the study, the wards were reconfigured to offer single
sex accommodation for patients. Majority of patients admitted to these wards were over 65 and had an acute medical condition. The eligibility criteria for inclusion into the study were

- All participants over 65,
- English speaking,
- fit enough to be interviewed and able to give informed consent,
- and non-cognitively impaired (MMSE scores above 24)

All patients meeting the inclusion criteria were approached to take part in the study by the medical team. Patients were only interviewed when the medical team deemed them medically fit and capable of giving informed consent.

Patients who were not able to complete the assessments and screening tools due to acute confusion or severe cognitive impairment were excluded. Patients who had to be transferred out of the acute medical unit as soon as they were well or to intermediate care beds as a result of bed pressures were also excluded.

5.3 Assessments and Screening Tools

Four assessment tools were employed during the study. These were

1. The Edinburgh Depression Scale (EDS)
2. The 15-item Geriatric Depression Scale (GDS-15) incorporating the GDS-4.
3. 2 questions proposed by NICE in their clinical guidelines on depression (NICE, 2004 & NICE, 2009).
4. The Present State Examination – Schedules for Clinical Assessment in Neuropsychiatry (PSE-SCAN; WHO, 1982).
The Edinburgh Depression Scale (EDS)

The Edinburgh Depression Scale (EDS) (Cox et al., 1990) is a self-report questionnaire comprising 10 items corresponding to cognitive-affective symptoms of depression such as guilt feeling, sleep disturbance, subjective sadness, anhedonia, and suicidal ideation. Each item has 4 statements and the patient is requested to underline one statement from each item corresponding to how they have felt in the last 7 days. Responses are scored from 0-3, with the most negative statement getting the highest score. Overall assessment is done by total score, which is determined by adding together the scores for each of the 10 items.

The 15-item Geriatric Depression Scale (GDS-15)

The GDS is a 30-item self reported or interviewer administered inventory with a simple yes/no format. It was devised specifically for use in older people and consists of items representing non somatic symptoms of depression. The 15-item version of the GDS (GDS-15) was developed from the original 30-item version (Sheikh and Yesavage, 1986) to reduce fatigue or poor concentration arising from using lengthy scales. In scoring the GDS, each item is awarded a score of 0 or 1 depending on whether it is worded positively or negatively. The total score on the GDS ranges from 0 to 15 or 30 depending on the version used.

The 2 questions proposed by NICE (NICE-2) in their depression guidelines.

The NICE-2 was recommended by National Institute of Health and Clinical Excellence (NICE) as a simple, effective and quick measure of screening for depression in primary care and general hospital settings. The initial recommendation was made in 2004 and reiterated in the updated 2009 guideline. The 2 questions suggested by NICE are: “During the past month, have you often been bothered by feeling down, depressed or hopeless?” and, “During the last month, have you been bothered by having little interest or pleasure in
doing things?” These 2 questions were originally part of a screening questionnaire (the Primary Care Evaluation of Mental Disorders- PRIME-MD) designed to facilitate the diagnosis of common mental disorders in primary care. The evidence for the use of the 2 questions was based on a paper by Whooley et al. (1997) where an affirmative response to either of the 2 questions performed as well as longer case finding instruments.

The “Gold Standard”

The Schedules for Clinical Assessment in Neuropsychiatry (SCAN) is a set of instruments and manuals developed by the WHO with the aim of improving the accuracy and reliability of classifying psychiatric disorders. The core component of SCAN is the 10th version of the Present State Examination (PSE-10), a semi-structured interview, and successor to PSE-9 (Wing et al. 1990; Wing 1996). The PSE-10 consists of two parts: the first part deals with anxiety, mood, and other neurotic features and the second with psychotic and cognitive features, and observation items on speech, behaviour, and affect.

The interviews take the form of a detailed clinical cross-examination of the subject, resembling a clinical interview, with the aim of discovering whether each of a comprehensive list of symptoms is present and if so, in what degree of severity. For most symptoms, a suggested wording is provided for eliciting each of the symptoms, although it is ultimately the interviewer who must decide if the symptom definitions are fulfilled. The order in which the sections are done depends on the most important symptoms of the respondent, making the SCAN flexible and easy to use.

Assessing correctly the presence or absence of a particular symptom or psychopathological phenomena requires some experience and knowledge of psychiatry (Brugha et al. 1999a). Interviewers would have to undergo training to enable them distinguish normal from abnormal phenomena and also acquire the specific interviewing skills. Data obtained from
the interviews are entered into a computer algorithm to generate either ICD-10 or DSM-IV psychiatric diagnosis.

Using the SCAN as ‘reference or gold standard’ in clinical research provides clinicians with an independent, standardized and valid method of checking clinical diagnosis that is more valid than wholly structured interviews (Brugha et al. 1999a). It has also been widely used in research with good reliability (Rijnders et al. 2000) and is understood by most scientists whether or not their preferred nosology differs from ICD-10 or DSM-IV.

It is for these reasons the PSE-SCAN was chosen as the ‘gold standard’ and used to identify patients suffering from depression according to ICD-10 criteria.

The clinical interviews were carried out by 2 psychiatrists (CE & RA) (specialist registrars) who had undergone a 5–day training course at a recognized WHO SCAN centre in the UK prior to recruitment of participants.

The first component of the PSE-SCAN to be administered was the Mini-Mental State Examination (Folstein et al., 1975), which is a brief 30-item test of the patients’ cognitive functioning. Patients scoring less than 24 on this test were excluded as it indicated a level of cognitive impairment that may potentially affect responses to the depression screening instruments, rendering them unreliable. The other components of the PSE-SCAN administered were those necessary to identify patients suffering from depression, namely sections 6 (Depressed mood and ideation), 7 (thinking, concentration, energy, interest) and 8 (bodily functions). The interviewers were blind to the results of the screening instruments, and the interview data was entered into a password protected laptop computer.
5.4 Procedure

Ethical and Clinical governance approval

Ethical approval for this study was obtained from the Leicestershire, Northamptonshire and Rutland Research Ethics Committee 1, on the 25th of May 2006 while Clinical governance approval was obtained towards the end of July 2006, from the University Hospitals of Leicester Research and Development Directorate. As a condition of the approval, an annual progress report for the research was submitted to both organisations throughout the course of the study. The research team also had regular meetings to discuss the progress of the research.

The research

The formal start date for the study was the 10th of October 2006, although recruitment of participants only started in November 2006. Copies of the research protocol, inclusion criteria and patient information sheets were made available for the medical teams in designated offices in all the wards involved in the research. The medical teams identified and invited potential participants when they were deemed medically fit for interview and capable of giving informed consent. All in-patients over the age of 65 under the care of 4 acute medical teams at the Leicester General Hospital during the study period were considered for inclusion into the study. Potential participants were usually reviewed during ward reviews to see if they met all 4 inclusion criteria. The Abbreviated Mental Test Score (AMTS) (Hodkinson, 1972) was used by the medical team as a measure of cognitive impairment. Patients scoring less than 6 out of a maximum score of 10 were usually judged by the medical team as too cognitively impaired to be invited into the study. All eligible participants were given an explanation of the nature of the research and then given
an information sheet to read by the medical team. Those unable to read due to any form of sensory impairment had the information read to them. Participants expressing an interest in the study were usually approached again before the end of the day to obtain verbal consent to pass their details on to the researchers. Upon receiving details of those expressing interest, a member of the research team (trained psychiatrist) (ICE, RA) made an appointment to see them on the medical ward after a period of at least one working day.

The first visit by the research team involved a psychiatric interview by a trained psychiatrist using the PSE-SCAN. On this visit, the nature and purpose of the research was explained to the patient again, and any other questions clarified. Participants were then required to give valid written informed consent by signing or initialling a consent form (see appendix 6) which they kept with them. A copy of the signed form was also kept by the research team. After giving valid informed consent, participants were interviewed using sections of the PSE-SCAN comprising the MMSE and sections relevant to the diagnosis of depression (sections 6, 7 and 8). Baseline demographic details were also collected and data was entered into a password protected computer. The first component of the PSE-SCAN to be administered was the MMSE. If a participant scored less than 24 on the MMSE, the interview would be terminated and that participant excluded from the study. Only those with scores above 24 on the MMSE proceeded to answer sections 6, 7 and 8 of the PSE-SCAN. Participants were told to expect a second interview by another researcher within 5 working days.

The second part of the research involved another visit by a research fellow (SB), to administer the screening tools [GDS-15 (incorporating the GDS-4), the EDS, and the NICE two questions]. The research fellow visited the patient within a period of
5 working days of the PSE-SCAN interview, and was blind to the patient’s performance on the PSE-SCAN. The screening tools were presented to participants as self-completion questionnaires requiring the respondent to either tick a box or circle the most appropriate response. Participants unable to read any of the screening instruments had the information read out to them. Participants who needed help in ticking off or circling of their response to the questionnaires were assisted in doing so by the research fellow. If a patient had been discharged home within this period, an appointment was made to see them at home. Patients who completed only one part (either assessment) were not included in the final analysis of the results.

**Sample size/power calculation**

A sample size calculation was not performed before recruiting patients because the emphasis of this study (a diagnostic study) is to obtain accurate estimates of sensitivity, specificity and predictive values. Calculation of the sample size required for such a study requires knowledge of the proportion of patients who will test positive on both/all scales to be compared (Alonzo et al., 2002). As this was unknown at the onset of the study, sample size was initially determined by assuming that 100 acute admissions would yield approximately 20-30 patients with gold standard/ICD 10 depression based on the estimated prevalence of depression in older adults in hospital settings. Once the proportion of patients testing positive or negative on the 4 scales are known (i.e. at the end of the study), it is possible to use the AUCs to assess how accurate a particular test is compared to the gold standard.
Inclusion and Exclusion criteria

All participants had to be medically fit for interview and capable of giving informed consent as judged by the medical team responsible for their care. Non-English speakers were excluded because the screening instruments had not been validated in other languages. Participants scoring less than 24 on the MMSE were also excluded from the study because a score of less than 24 indicates a degree of cognitive impairment that may influence responses to the depression screening instruments rendering them unreliable.

Diagnosis of depression

Cases of depression were identified according to the ICD-10 criteria for depressive disorder using results obtained from the PSE-SCAN. A diagnosis of depression was made if within the past 4 weeks, the patients admitted to at least 2 of: depressed mood, loss of interest and enjoyment, and loss of energy plus 2 or more other criterion symptoms for 2 weeks or more. The ICD-10 criteria for depression are summarised in Appendix 2.

Data collection and handling

The interview data from the relevant sections of the PSE-SCAN were entered into a laptop computer using research ID code numbers for patients to ensure security and to comply with data protection regulation. The EDS, GDS and the two NICE screening questions were presented to participants as self-completing questionnaires that required them to tick, underline or circle the most appropriate response. Patients unable to read the questionnaire for any reason had the questionnaire read to them by the researcher. All information concerning subjects were kept in research files stored in a locked cabinet located in the research office of the University Department of Psychiatry for the Elderly, Leicester General Hospital, to which only the research team had access.
Data Analysis

Data were analysed using the statistical package PASW (previously known as SPSS), version 18. 2 x 2 tables were used to calculate sensitivity and specificity, as well as and positive and negative predictive values for each screening tool at various cut-offs. ICD-10 diagnosis of depression was used as ‘gold standard’ in calculating these values. Receiver Operating Characteristics (ROC) curve was calculated for each scale to determine the overall accuracy. Based on the area under the curve, a maximum value of 1.0 indicates perfect performance of the screening instrument. The cut-off value that optimizes sensitivity without compromising specificity was chosen for each screening tool and then compared.

A description of the various measures employed in the analysis of the performance of the screening instruments is described below:

1. **Sensitivity**

The sensitivity of a screening instrument describes the proportion (or %) of patients with a disorder who are correctly identified by the instrument (or have a positive result). Sensitivity is normally a constant property of the test and is therefore not affected by the prevalence of the condition. Screening instruments that are very sensitive tend to identify most people with the disorder and rule out the disorder with certainty when the results are negative.

Sensitivity = \( \frac{a}{a+c} \)

<table>
<thead>
<tr>
<th>Target disorder (Gold standard)</th>
<th>Present</th>
<th>Absent</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnostic test result</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>a</td>
<td>b</td>
<td>a+b</td>
</tr>
<tr>
<td>Negative</td>
<td>c</td>
<td>d</td>
<td>c+d</td>
</tr>
<tr>
<td>Total</td>
<td>a+c</td>
<td>b+d</td>
<td>a+b+c+d</td>
</tr>
</tbody>
</table>
2. **Specificity**

The proportion (or %) of patients without the disorder \((b+d)\) who have a negative test result \((d)\) is called the specificity of a screening instrument. It is also a constant property of the test that is not affected by the prevalence of a disorder in a study sample. A specific test therefore identifies most people without the disorder and hence when the results are positive, it tends to rule in the disorder (Lawrie et al, 2000).

\[
\text{Specificity} = \frac{d}{b+d}
\]

3. **Positive predictive value**

The positive predictive value of a screening instrument describes the proportion of people with a positive test result \((a+b)\) who have the disorder \((a)\). The PPV is a more clinically useful measure as clinicians are more interested in knowing the chances of a patient having the condition if they have a positive test result. The PPV is dependent on the prevalence of a condition in a study sample and tends to rise as the prevalence of the disorder increases.

\[
\text{PPV} = \frac{a}{a+b}
\]

4. **Negative predictive value**

The negative predictive value of a screening instrument describes the proportion of people with a negative test result \((c+d)\) who do not have the disorder \((d)\). The NPV is also dependent on the prevalence of the condition in a study sample and tends to rise as the prevalence of the disorder decreases.

\[
\text{NPV} = \frac{d}{c+d}
\]

5. **ROC curves and cut-off thresholds**

The optimum diagnostic threshold for any screening instrument depends on the nature of the condition and the costs of a false positive or a false negative result. Higher cut-off values tend to generate more specific but less sensitive tests, while lower cut-offs are generally more sensitive and less specific. In the case of depression in older adults in general hospital settings where detection rates are very low and the costs of false negatives
very high (i.e. poor prognosis of physical conditions if left undetected and untreated), a highly sensitive test becomes more important.

The ROC curve is a graphical plot of sensitivity on the Y axis and 1-specificity on the X axis, for each cut-off threshold of the screening instrument. The optimal cut-off is that point on the curve that lies closest to the top left corner and is also the point which maximises the area under the curve (AUC)

The ROC curve is also used to compare competing diagnostic or screening tests. A test with the greatest area under the curve will generally be the more superior test (Lawrie et al., 2000).

6. Comparing performance of screening instruments using area under ROC curves

In the case of this study where comparisons of the performance of diagnostic tests was applied to the same set of patients, it is appropriate to calculate the standard error of the difference between the 2 areas using a formula that takes into account the correlation(r) induced between the 2 areas by the study of the same set of cases. The equation is given below (Hanley and McNeil, 1983):

\[
\text{SE} (\text{Area1} – \text{Area2}) = \sqrt{\text{SE}^2(\text{Area1}) + \text{SE}^2(\text{Area2}) – 2r\text{SE}(\text{Area1})\text{SE}(\text{Area 2})}
\]

Once the SE is known, the \( z \) statistic can be calculated using another formula:

\[
Z = (\text{Area1} – \text{Area2}) / \text{SE} (\text{Area1} – \text{Area2})
\]

If this figure (\( z \)) is above the critical level of 1.96, then we accept the areas are different. If \( z \) is below this critical level, then no difference exists between the curves (i.e. a non significant result).
5.5 Ethical Considerations

The study did not commence until full ethical approval had been obtained. The information sheet and consent forms were produced in large print for patients who with eyesight difficulties. Participants unable to read the information sheet or any of the screening instruments had the information read out to them. Participants were given the opportunity to ask questions and had the telephone number of the chief investigator (the author) if they had any other unanswered queries. Valid informed (written and signed) consent was obtained from all participants and they were free to withdraw their consent at any stage. Medical teams or general practitioners were informed of any patient found to be suffering from depression following the psychiatric assessment, to enable appropriate intervention or further referral to specialist mental health services.
SECTION THREE: RESULTS
6. RESULTS

6.1 Participants

A total of 125 participants met the eligibility criteria and agreed to take part in the study. Out of these, 7 completed the PSE-SCAN but declined the follow-up interview with the research fellow, hence were excluded. That left 118 participants who completed both the PSE-SCAN (Gold-standard) and the three depression screening instruments (although one participant did not complete the GDS but completed all other assessments). Among the 7 that did not complete the follow-up interview, 5(71.4%) were male, mean age was 80.5 years and the mean MMSE score was 27.7. The ages of participants ranged from 65 to 95 years with a mean age of 81.4 years (standard deviation 6.7 years). Two-thirds were female (65.3%), more than half were widowed and all were cognitively intact (mean MMSE score was 27.1). This sample reflects the eligibility criteria with all participants being English speaking, MMSE 24 or more, and medically fit to be interviewed. Data on the individual medical condition and reason for admission were not collected.

Table 6.1 summarises the characteristics of the study sample.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Subjects (N=118)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age (years)</td>
<td>81.4</td>
</tr>
<tr>
<td>Median age (years, range)</td>
<td>82(65 – 95)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>41(34.7%)</td>
</tr>
<tr>
<td>Female</td>
<td>77(65.3%)</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>36(30.5%)</td>
</tr>
<tr>
<td>Cohabiting</td>
<td>1(0.8%)</td>
</tr>
<tr>
<td>Separated</td>
<td>1(0.8%)</td>
</tr>
<tr>
<td>Divorced</td>
<td>8(6.8%)</td>
</tr>
<tr>
<td>Widowed</td>
<td>62(52.5%)</td>
</tr>
<tr>
<td>Never married and not cohabiting</td>
<td>10(8.5%)</td>
</tr>
<tr>
<td>MMSE scores</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>27.1</td>
</tr>
<tr>
<td>Median</td>
<td>27</td>
</tr>
<tr>
<td>Mode</td>
<td>26</td>
</tr>
</tbody>
</table>
ICD 10 diagnosis
- Depressed: 26 (22%)
- Non-depressed: 92 (78%)

ICD 10 Depressive Episode
- Mild: 16 (13.5%)
- Moderate: 10 (8.5%)
- Severe: 0 (0%)

Mean GDS score: 4.5
Mean EDS score: 6.7
Mean NICE-2 score: 0.69
Mean GDS-4 score: 0.88

6.2 Prevalence of ICD 10 depressive episode

Twenty-six of the 118 patients who completed the PSE-SCAN interviews were identified as suffering from a depressive episode as defined by ICD-10 criteria. The prevalence of depression in this study sample was therefore 22%. When classified according to severity, 16 (13.5%) of the patients had mild depressive episodes, 10 (8.5%) had moderate depressive episodes and none had severe depression.

The age- and sex-specific prevalence of ICD 10 depressive episode is summarised in table 6.2. Depression was more prevalent among women than men (23.4% versus 19.5%) and nearly a quarter of older adults aged 85 or more were depressed. These findings were not statistically significant at the associated significance level of p<0.05.

<table>
<thead>
<tr>
<th>Variable</th>
<th>(N=118)</th>
<th>ICD 10 depressive episode</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Depressed</td>
</tr>
<tr>
<td>Males</td>
<td>41</td>
<td>8 (19.5%)</td>
</tr>
<tr>
<td>Females</td>
<td>77</td>
<td>18 (23.4%)</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>65 - 74</td>
<td>21</td>
<td>6 (5.1%)</td>
</tr>
<tr>
<td>75 – 84</td>
<td>59</td>
<td>11 (18.6%)</td>
</tr>
<tr>
<td>85&gt;</td>
<td>38</td>
<td>9 (23.6%)</td>
</tr>
</tbody>
</table>
6.3 Detection of depression using the screening instruments

The ability of the 4 screening instruments to detect depression at various cut-off thresholds was investigated and the findings are summarised below.

1. Edinburgh Depression Scale

The performance of the EDS as a screening tool for depression in older adults in acute general hospital settings was evaluated by calculating the sensitivity, specificity, positive predictive values and negative predictive values across a range of thresholds.

The mean score of the EDS in this sample was 6.7. At low cut-off thresholds, the EDS had high sensitivity but low positive predictive values. Increasing the threshold appears to decrease the sensitivity while increasing the specificity, positive predictive values and negative predictive values. For example, at a cut-off threshold of 7, sensitivity was 92% and specificity 67% while a higher threshold of say, 14 gave sensitivity of 42%; specificity of 96%; PPV of 78%; and NPV of 85%. Table 6.3 summarises the performance of the EDS across a range of thresholds.

<table>
<thead>
<tr>
<th>Cut off</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>7+</td>
<td>0.92(0.73-0.98)</td>
<td>0.67(0.56-0.76)</td>
<td>0.44(0.31-0.85)</td>
<td>0.96(0.88-0.99)</td>
</tr>
<tr>
<td>8+</td>
<td>0.88(0.68-0.96)</td>
<td>0.77(0.67-0.85)</td>
<td>0.52(0.36-0.67)</td>
<td>0.95(0.87-0.98)</td>
</tr>
<tr>
<td>9+</td>
<td>0.80(0.60-0.92)</td>
<td>0.81(0.71-0.88)</td>
<td>0.55(0.38-0.71)</td>
<td>0.93(0.85-0.97)</td>
</tr>
<tr>
<td>10+</td>
<td>0.73(0.51-0.87)</td>
<td>0.88(0.79-0.93)</td>
<td>0.63(0.43-0.79)</td>
<td>0.92(0.83-0.96)</td>
</tr>
<tr>
<td>11+</td>
<td>0.65(0.44-0.82)</td>
<td>0.93(0.85-0.97)</td>
<td>0.73(0.51-0.88)</td>
<td>0.90(0.82-0.95)</td>
</tr>
<tr>
<td>12+</td>
<td>0.61(0.40-0.79)</td>
<td>0.93(0.85-0.97)</td>
<td>0.72(0.49-0.88)</td>
<td>0.89(0.81-0.94)</td>
</tr>
<tr>
<td>13+</td>
<td>0.53(0.33-0.72)</td>
<td>0.95(0.88-0.98)</td>
<td>0.77(0.51-0.92)</td>
<td>0.88(0.79-0.93)</td>
</tr>
<tr>
<td>14+</td>
<td>0.42(0.23-0.62)</td>
<td>0.96(0.90-0.99)</td>
<td>0.78(0.48-0.94)</td>
<td>0.85(0.77-0.91)</td>
</tr>
</tbody>
</table>

The optimum threshold of the EDS at identifying cases of depression in this study sample was 9, which had a sensitivity of 80% and specificity of 81%. This cut-off was determined by plotting the figures in the table above to create a receiver operating characteristic (ROC).
The optimal cut-off is that point on the curve that lies closest to the top left corner and is also the point which maximises the area under the curve. Using this cut-off, thirty-eight patients (32.2%) scored above this threshold but only 21 of them (55%) were actually depressed according to ICD-10 criteria (true positives). Table 6.4 shows the proportion of participants scoring above a range of cut-off thresholds on the EDS.

**Table 6.4**

<table>
<thead>
<tr>
<th>EDS cut-off point</th>
<th>Proportion of participants scoring above cut-off</th>
<th>Proportion of participants actually depressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>7+</td>
<td>45.8%</td>
<td>20.3%</td>
</tr>
<tr>
<td>8+</td>
<td>37.3%</td>
<td>19.5%</td>
</tr>
<tr>
<td>9+</td>
<td>32.2%</td>
<td>17.8%</td>
</tr>
<tr>
<td>10+</td>
<td>25.4%</td>
<td>16.1%</td>
</tr>
<tr>
<td>11+</td>
<td>19.5%</td>
<td>14.4%</td>
</tr>
<tr>
<td>12+</td>
<td>18.6%</td>
<td>13.5%</td>
</tr>
<tr>
<td>13+</td>
<td>15.3%</td>
<td>11.9%</td>
</tr>
<tr>
<td>14+</td>
<td>11.9%</td>
<td>9.35%</td>
</tr>
</tbody>
</table>

The Receiver operating characteristic curve for the EDS for all possible thresholds is shown in Figure 6.1. The area under the curve was 0.908 [p<0.0005, 95% confidence interval (CI) =0.852 to 0.965], indicating very good performance. From the ROC curve shown below, it can be seen that the optimum cut-off threshold lies somewhere between 8 and 9.
2. The 15-item Geriatric Depression Scale (GDS-15)

The performance of the GDS-15 was evaluated in 117 participants as one patient did not complete the scale. Data from this patient was therefore excluded in the calculations of sensitivity, specificity, positive predictive value and negative predictive values for the GDS only.

The mean score for the GDS-15 in this sample was 4.5. At a cut-off of 2, the GDS achieved 100% sensitivity but only 17% specificity. Applying the predetermined cut-off of 5 as suggested by the original validation study of the GDS-15 (Sheik and Yesavage, 1986) to this sample, produced 84% sensitivity, 75% specificity but a low positive predictive value of 47%. This cut-off may therefore not be suitable for older adults in acute medical settings. Table 6.5 summarises the performance of the GDS across a range of thresholds.
Table 6.5
Performance of the GDS at various cut-offs.

<table>
<thead>
<tr>
<th>Cut off</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>2+</td>
<td>1.0(0.83-1.00)</td>
<td>0.17(0.10-0.27)</td>
<td>0.24(0.16-0.34)</td>
<td>1.00(0.75-1.00)</td>
</tr>
<tr>
<td>3+</td>
<td>0.96(0.77-0.99)</td>
<td>0.48(0.38-0.59)</td>
<td>0.33(0.23-0.46)</td>
<td>0.97(0.87-0.99)</td>
</tr>
<tr>
<td>4+</td>
<td>0.88(0.67-0.96)</td>
<td>0.59(0.49-0.69)</td>
<td>0.37(0.25-0.50)</td>
<td>0.94(0.84-0.98)</td>
</tr>
<tr>
<td>5+</td>
<td>0.84(0.63-0.94)</td>
<td>0.75(0.64-0.83)</td>
<td>0.47(0.32-0.63)</td>
<td>0.94(0.85-0.98)</td>
</tr>
<tr>
<td>6+</td>
<td>0.80(0.58-0.92)</td>
<td>0.85(0.76-0.91)</td>
<td>0.60(0.42-0.76)</td>
<td>0.94(0.86-0.97)</td>
</tr>
<tr>
<td>7+</td>
<td>0.80(0.58-0.92)</td>
<td>0.86(0.77-0.92)</td>
<td>0.62(0.43-0.78)</td>
<td>0.94(0.86-0.97)</td>
</tr>
<tr>
<td>8+</td>
<td>0.64(0.42-0.81)</td>
<td>0.91(0.83-0.95)</td>
<td>0.66(0.44-0.83)</td>
<td>0.90(0.81-0.95)</td>
</tr>
<tr>
<td>9+</td>
<td>0.60(0.38-0.78)</td>
<td>0.93(0.85-0.97)</td>
<td>0.71(0.47-0.87)</td>
<td>0.89(0.81-0.94)</td>
</tr>
<tr>
<td>10+</td>
<td>0.48(0.28-0.68)</td>
<td>0.94(0.87-0.97)</td>
<td>0.70(0.44-0.86)</td>
<td>0.87(0.78-0.92)</td>
</tr>
</tbody>
</table>

The proportion of participants scoring above various thresholds on the GDS is also shown in table 6.6. If the recommended threshold of 5 is applied to this sample, 44 patients (37.6%) would score above this cut-off but only 21 (47.7%) would have been correctly identified as depressed according to ICD-10 criteria.

Table 6.6
Proportion of participants scoring at various cut-off points on the GDS

<table>
<thead>
<tr>
<th>GDS cut-off point</th>
<th>Proportion of participants scoring above cut-off</th>
<th>Proportion of participants actually depressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>2+</td>
<td>86.3%</td>
<td>21.4%</td>
</tr>
<tr>
<td>3+</td>
<td>60.7%</td>
<td>20.5%</td>
</tr>
<tr>
<td>4+</td>
<td>50.4%</td>
<td>18.8%</td>
</tr>
<tr>
<td>5+</td>
<td>37.6%</td>
<td>17.9%</td>
</tr>
<tr>
<td>6+</td>
<td>28.2%</td>
<td>17.1%</td>
</tr>
<tr>
<td>7+</td>
<td>27.4%</td>
<td>17.1%</td>
</tr>
<tr>
<td>8+</td>
<td>20.5%</td>
<td>13.6%</td>
</tr>
<tr>
<td>9+</td>
<td>17.9%</td>
<td>12.8%</td>
</tr>
<tr>
<td>10+</td>
<td>14.5%</td>
<td>10.3%</td>
</tr>
</tbody>
</table>

The ability of the GDS to correctly identify patients suffering from depression was also examined using a receiver operating characteristic (ROC) curve. The ROC curve for the
GDS is shown in figure 6.2. The area under the curve (AUC) was 0.883\( p< 0.0005, 95\% \) confidence interval (CI) = 0.808 to 0.959\] indicating very good accuracy in identifying cases of depression.

**Figure 6.2**
Receiver operating characteristic curve for the GDS

The optimum cut-off threshold for the GDS in this population was 7. At this threshold, the GDS achieved sensitivity of 80\%; specificity of 86\%; PPV of 62\% and NPV of 94\%.

3. **The 4-item Geriatric Depression Scale (GDS-4)**

The test performance of the GDS-4 was also evaluated using the same methods described above for the EDS and the GDS-15. Using a cut-off threshold of 2, the GDS-4 had a sensitivity of 80\%, specificity of 84\%, positive predictive value of 58\% and negative predictive value of 93\%. However if the recommended cut-off threshold of 1(D’Ath et
al., 1994) is applied to this sample, the sensitivity increases to 88% but the positive predictive value (i.e., the probability of depression in a person scoring greater than or equal to 1 on the GDS-4) drops to 41%. A cut-off threshold of 2 may therefore be preferable in this sample as will be shown on the ROC curve analysis.

**Table 6.7**
**Performance of GDS-4 questions at various cut-offs**

<table>
<thead>
<tr>
<th>Cut off</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1+</td>
<td>0.88(0.67-0.96)</td>
<td>0.66(0.55-0.75)</td>
<td>0.41(0.28-0.55)</td>
<td>0.95(0.86-0.98)</td>
</tr>
<tr>
<td>2+</td>
<td>0.80(0.58-0.92)</td>
<td>0.84(0.75-0.91)</td>
<td>0.58(0.40-0.74)</td>
<td>0.93(0.85-0.97)</td>
</tr>
<tr>
<td>3+</td>
<td>0.32(0.15-0.53)</td>
<td>0.94(0.87-0.97)</td>
<td>0.61(0.32-0.84)</td>
<td>0.83(0.74-0.89)</td>
</tr>
</tbody>
</table>

Another way of assessing the performance of a screening tool is to look at the proportion of participants scoring at various thresholds and then calculating those who are correctly identified using the gold standard, in this case the ICD-10.

The proportion of participants scoring at various thresholds is shown in table 6.8. If the threshold of 2 is applied to this sample, 34(29%) will score above this threshold and 20(58%) will be correctly identified as depressed using ICD-10 criteria.

**Table 6.8**
**Proportion of participants scoring at various cut-off points on the GDS-4**

<table>
<thead>
<tr>
<th>GDS-4 cut-off point</th>
<th>Proportion of participants scoring above cut-off</th>
<th>Proportion of participants actually depressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1+</td>
<td>45.3%</td>
<td>18.8%</td>
</tr>
<tr>
<td>2+</td>
<td>29.1%</td>
<td>17.1%</td>
</tr>
<tr>
<td>3+</td>
<td>11.1%</td>
<td>6.83%</td>
</tr>
</tbody>
</table>

The test performance of the GDS-4 was also evaluated using the receiver operating characteristics (ROC) curve. Figure 6.3 shows the ROC curve for the GDS-4. The area under the curve is 0.845 indicating that it also performs well as a screening instrument in this sample [p<0.0005, 95% confidence interval (CI) =0.753 to 0.938]. A cut-off of 2 gave
optimal performance in this study with 80% sensitivity, 84% specificity and high NPV of 93%.

Figure 6.3
Receiver operating characteristic curve for the GDS-4

4. The 2 questions suggested by NICE depression guidelines (NICE-2)

All eligible participants (n=118) completed the NICE-2 questions along with the EDS and the PSE-SCAN. Table 5.9 shows the data on sensitivity, specificity, positive and negative predictive values at the two possible cut-offs (1+ yes to 1 question; and 2+ yes to both questions). Answering yes to either question gave a sensitivity of 100%, specificity of 70%, positive predictive value of 49% and negative predictive value of 100%. A yes response to the two questions reduces the sensitivity to 76% but increases the specificity to 90%.
Table 6.9
Performance of NICE-2 questions at various cut-offs

<table>
<thead>
<tr>
<th>Cut off</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1+</td>
<td>1.00(0.83-1.00)</td>
<td>0.70(0.60-0.79)</td>
<td>0.49(0.35-0.62)</td>
<td>1.00(0.93-1.00)</td>
</tr>
<tr>
<td>2+</td>
<td>0.76(0.55-0.90)</td>
<td>0.90(0.81-0.95)</td>
<td>0.69(0.49-0.84)</td>
<td>0.93(0.85-0.97)</td>
</tr>
</tbody>
</table>

The proportion of participants scoring above the two thresholds is shown in table 6.10. Applying the cut-off threshold of 1+, 53 patients (44.9% of the study sample) scored above this threshold (i.e. answered yes to either question), but only 26/53 (49.1%) were correctly identified as depressed according to ICD-10 criteria. Misclassification rates are reduced if the threshold of 2+ is applied (i.e. 29 patients scoring above this threshold with 20/29 (68.9%) correctly identified as depressed).

Table 6.10
Proportion of participants scoring at various cut-off points on the NICE-2

<table>
<thead>
<tr>
<th>NICE cut-off point</th>
<th>Proportion of participants Scoring above cut-off</th>
<th>Proportion of participants actually depressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1+</td>
<td>44.9%</td>
<td>22.0%</td>
</tr>
<tr>
<td>2+</td>
<td>24.6%</td>
<td>16.9%</td>
</tr>
</tbody>
</table>

The receiver operating characteristics curve for the two-question screening instrument (NICE-2) is shown in Figure 6.4 below. The area under the curve is the greatest of the three screening instruments at 0.922 indicating that it performs very well as a screening instrument in older adults in acute general hospital settings[ p<0.0005, 95% confidence interval(CI) = 0.874 to 0.970].
5. Test characteristics of the 4 screening instruments compared

Comparisons of the test characteristics of the 4 screening instruments using optimum cut-off thresholds are displayed in table 6.11. Sensitivity for all the screening instruments ranged from 80% (GDS-15, GDS-4 and EDS) to 100% (NICE-2). Specificity ranged from 70% (NICE-2) to 86% (GDS-15). Positive predictive value was highest for the GDS-15 but lowest for the NICE-2 questions while the negative predictive value was 100% for the NICE-2 screening questions.
Table 6.11
Comparison of performance of the 4 screening instruments

<table>
<thead>
<tr>
<th></th>
<th>GDS-15 (cut-off:7+)</th>
<th>GDS-4 (Cut-off:2+)</th>
<th>NICE-2 (cut-off: 1+)</th>
<th>EDS (cut-off:9+)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity (%)</td>
<td>80 (0.58 - 0.92)</td>
<td>80 (0.58 - 0.92)</td>
<td>100 (0.83 -1.00)</td>
<td>80 (0.60 - 0.92)</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>86 (0.77 - 0.92)</td>
<td>84 (0.75 - 0.91)</td>
<td>70 (0.60 - 0.79)</td>
<td>81 (0.71 - 0.88)</td>
</tr>
<tr>
<td>Area under ROC curve</td>
<td>0.88 (0.81 - 0.96)</td>
<td>0.85 (0.75 - 0.94)</td>
<td>0.92 (0.87 - 0.97)</td>
<td>0.91 (0.85 - 0.96)</td>
</tr>
<tr>
<td>PPV (%)</td>
<td>62 (0.43 - 0.78)</td>
<td>58 (0.40 - 0.74)</td>
<td>49 (0.35 - 0.62)</td>
<td>55 (0.38 - 0.71)</td>
</tr>
<tr>
<td>NPV (%)</td>
<td>94 (0.86 - 0.97)</td>
<td>93 (0.85 - 0.97)</td>
<td>100 (0.93 - 1.00)</td>
<td>93 (0.85 - 0.97)</td>
</tr>
</tbody>
</table>

The receiver operating characteristic curves for all screening instruments is displayed in figure 6.5. As data was collected for 117 participants for the PSE-SCAN and all the screening instruments, the ROC curve analysis shown below excluded the one participant who did not complete the GDS. There was very little difference in the results by omitting that sample. The ROC curve shows that the 2 questions suggested by NICE has the largest area under the curve (0.922) followed by the Edinburgh Depression Scale (0.908). The 4 item GDS had the lowest area under the curve. Using the area under the curve as a summary measure of performance of these screening tools, the NICE-2 questions appears to have performed as well if not better than the other screening instruments in this study sample using the cut-off threshold of 1+( i.e. answering ‘yes’ to one of the two questions).

6. Statistical analysis comparing AUCs of ROC curves

A secondary aim of this study was to compare the performances of the screening instruments under evaluation. The appropriate statistical analysis for this is to compare the AUCs of ROC curves, using the methodology developed by Hanley and McNeill (1983). This compares one test with another, and takes account of the correlation induced between the two areas when the tests are applied to the same set of patients, as was the case in this study. This method derives a z-statistic from the standard error of the difference in the
two AUCs, and if \( z \) is above a particular level (conventionally \( 1.96 / P=0.05 \)), then it is accepted that the tests are different. When the AUCs of the EDS and GDS-15 were compared, the \( z \)-statistic was 0.604, i.e. non-significant. The \( z \) statistic was also lower than the critical value of 1.96 when NICE-2 was compared with GDS-4.

It is important to bear in mind that this study was not powered in advance to identify differences between the performance of the instruments, so the lack of statistical significance found here may represent a false-negative (Type II) error.

Figure 6.5
Receiver operating characteristic curves for GDS-15, GDS-4, NICE-2 and EDS (\( N=117 \)).
6.4 Detection of milder forms of depression using the screening instruments

Although this was not the primary objective of this study, the ability of the 4 instruments to correctly identify patients with milder forms of depression using the optimum cut-off threshold was examined in this study sample (table 6.12). As ICD-10 mild depression with 4 symptoms is subsumed under the category of minor or sub threshold depression according to the DSM-IV research criteria, assessing detection rates of ICD-10 mild depression would give an indication of how the tools may perform in identifying other forms of sub threshold depression including those with between 2 and 4 symptoms. The results show that if the EDS is used as a screening tool for patients suffering from ICD-10 mild depression, it would fail to identify or miss 25% (1 in 4) of these cases. The GDS-15 misses 30%; GDS-4 misses 25% and NICE-2 just 6.3% (table 6.12). In other words, the NICE-2 appears to identify most patients with milder forms of depression. In the case of moderate depression, all 4 screening tools identified at least 90% of cases.
<table>
<thead>
<tr>
<th>Screening tool (Cut point)</th>
<th>Mild depression N=16</th>
<th>Moderate depression N=10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Identified</td>
<td>Missed</td>
</tr>
<tr>
<td>EDS (9+)</td>
<td>12(74.8%)</td>
<td>4(25.2%)</td>
</tr>
<tr>
<td>GDS-15 (7+)</td>
<td>11(68.6%)</td>
<td>5(31.4%)</td>
</tr>
<tr>
<td>GDS-4 (2+)</td>
<td>12(74.9%)</td>
<td>4(25.1%)</td>
</tr>
<tr>
<td>NICE-2 (1+)</td>
<td>15(93.7%)</td>
<td>1(6.3%)</td>
</tr>
</tbody>
</table>
SECTION FOUR: DISCUSSION
7. DISCUSSION

7.1 Overview

This study has evaluated the performance of the Edinburgh Depression Scale and other brief questionnaires as tools for screening and diagnosing depression in older people who are physically ill in hospital. The PSE-SCAN acted as gold standard and generated a diagnosis of depression according to ICD-10 criteria (see Appendix 2). The use of the PSE-SCAN ensured that the assessments were standardised and therefore more reliable than an unstructured psychiatric interview. All participants were cognitively intact, English speaking and well enough to be interviewed. The findings from this study can only be generalised to patients who meet the inclusion criteria of the study, but this sample reflects every-day NHS practice, as most assessments are usually carried out on patients when they are medically fit and not confused.

118 patients were recruited randomly from four acute medical wards during the working week, as long as they met the eligibility criteria and agreed to take part in the study. The ratio of females to males in this sample was 3:2 and the average age of participants was 81.4 years. This number, though small, is comparable to other similar studies in medical in-patients (Koenig et al., 1992 \(N=109\); Blank et al., 2004 \(N=150\); O’Riordan et al., 1990 \(N=111\); Pomeroy et al., \(N=87\); Cullum et al., 2006\(N=223\)), and yielded a prevalence of depression of 22% according to ICD-10 criteria. Several studies using standardized research interviews in hospitalized older people have reported prevalence rates of depression ranging from 10% to 50% (Koenig et al., 1992; Blank et al., 2004; O’Riordan et al., 1990; Jackson et al., 1993; Magni et al., 1986; Pomeroy et al., 2001; Cullum et al., 2006; and Ramsay et al., 1991). A recent study based in the UK reported similar prevalence rates (17.7%) for depression using ICD-10 criteria derived from the Geriatric Mental State.
The prevalence of depression in this study sample is therefore in keeping with those obtained from previous studies and confirms the widely held notion that depression is much more common amongst older people who are physically ill in hospital than in the general population.

The age and sex specific prevalence of depression in this sample was also examined. The prevalence of depression for women was 23.4% and for men 19.5%. Prevalence studies on community samples have consistently reported rates of depression to be at least twice as common in women as men, with rates for men of 8.6% and for women 14.1% (Copeland et al., 1999). It would appear that our sample had more depressed males than one would expect with rates of depression almost equalling those in women. Depression was also more prevalent among the over 85s with rates of approximately 20%. A possible explanation for this finding might be greater burden of medical co morbidity in men and those aged over 85 compared to their counterparts in the community.

### 7.2 Assessments and screening interviews

The PSE-SCAN interviews were conducted by two psychiatric registrars who had undergone training in the use of the research interview in a recognised WHO training centre in the UK (Leicester). The use of clinicians with experience in psychiatry makes the results more reliable and less prone to errors. Evaluating the utility of any screening instrument requires that the comparisons between the tool and the gold standard be ‘blind’ and that the gold standard test be applied to the whole sample regardless of the results obtained from the screening instrument. This blind comparison with the ‘gold standard’ provides the most robust approach used in the evaluation of performance of a screening tool and reduces the possibility of ‘interviewer bias’.

In this study, the clinicians conducting the PSE-SCAN interview were blind to the results of the screening questionnaires and the research fellow who conducted the screening
questionnaires was also blind to the results of the PSE-SCAN. All patients included in the analysis completed both the PSE-SCAN and the screening instruments regardless of the outcome of both assessments. Patients were excluded if they did not complete the screening questionnaires after completing the PSE-SCAN. The time interval between both interviews was short (5 working days), to ensure so far as possible that the diagnostic status of the patients remained stable between both assessments.

7.3 The setting of the study
This is probably the first study evaluating the performance of the EDS in older adults admitted to general hospitals with a physical illness. The EDS has been evaluated in cancer patients (Lloyd-Williams et al., 2000) as well as in patients who are terminally ill (Lloyd-Williams et al., 2004) and found to be useful as a screening instrument in these contexts. Extending the use of the EDS as a screening tool in our sample of older adults, with different characteristics as those described above, will further validate the performance of the EDS across a range of patients. Applying the screening tests to patients with diagnoses (i.e. physical illnesses) frequently confused with the target disorder (i.e. depression) is another way of assessing whether the screening tool is able to discriminate against those without depression who are physically ill. The data obtained from this study is therefore valuable to clinicians in general hospital settings because it was obtained from patients where the screening instruments are most likely to be used.

7.4 Patients who did not take part in the study
Although data was complete for 117 participants (one patient did not complete the GDS), there were a number of potential participants who did not wish to take part in the study even though they were judged by the medical team as meeting the eligibility criteria. Another group of potential participants who were not included in the study were those
who were discharged home or transferred to intermediate care wards for rehabilitation before discharge home. Our study was not designed to collect information on those who did not wish to take part or those who were discharged before we could assess them. This means that there was some selection bias in our sampling procedure which may have an impact in the interpretation of the results. The problem of selection bias which is common in studies conducted on in-patient samples is one potential weakness of the study. Patients who did not take part in the study may (or may not) be suffering from depression and are not included in our results. It may be possible that the prevalence of depression, including calculations of sensitivity and specificity in our sample may be overestimated (or underestimated) depending on the characteristics of those who declined.

7.5 Performance of the Screening instruments

1. The Geriatric Depression Scale

The Geriatric Depression Scale (GDS) remains one of the most widely used screening tools for depression in older adults. The 15-item GDS was developed from the 30-item version and has been validated against the two major diagnostic systems - i.e. the DSM-IV/IIIR (Rinaldi et al., 2003), and the ICD-10 (Pomeroy et al., 2001; Cullum et al., 2006) and found to have satisfactory sensitivity (0.805) and specificity (0.750) across a range of settings including the community and inpatient geriatric units (Wancata et al., 2006). The most often used cut-off threshold was between 5 and 6, although in a recent study carried out in medical inpatients (Cullum et al., 2006), the optimal cut-off threshold was found to be 7.

This study sought to examine the performance of the GDS along with three other depression screening instruments in older adults who are physically ill in hospital. The shorter version of the GDS was chosen because the 30-item version was thought to be too
cumbersome and time consuming both for the patient and for the busy physician who will want a tool that is quick to administer in this population. The performance of the much shorter 4-item version derived from the 15-item version was also examined simultaneously. The gold-standard diagnostic test was the ICD-10 diagnosis of depression derived from the PSE-SCAN research interview.

Validity indices of the GDS-15 compared with previous studies

In our general hospital setting, the performance of the GDS-15 was examined by obtaining data on sensitivity, specificity, positive predictive value, negative predictive value, and receiver operating characteristic (ROC) curve analysis at various cut-off thresholds. The ROC curve allowed us to calculate the optimum cut-off threshold and the point which maximises the area under the curve (AUC).

Sensitivity, Specificity and cut-off thresholds

The results of this study suggest that the GDS-15 is a useful screening instrument for depression in older adults with physical illness who are admitted into hospital. The optimum threshold for the GDS-15 in this sample was 7+ and this gave sensitivity and specificity of 80% and 86% respectively. This cut-off is higher than that recommended by studies validating the GDS-15 in community samples of older adults (D’Ath et al., 1994 [cut-off threshold: 5+]; Arthur et al., 1999 [cut-off threshold: 3+]; Nyunt et al., 2009 [cut-off threshold: 5+]) but consistent with studies carried out on inpatient samples. For instance, Cullum et al., 2006 reported that a cut-off of 7+ was optimum for the sample of medical inpatients they studied, and at this cut-off sensitivity and specificity was 74% and 81% respectively. In another study of hospitalised older adults, a cut-off threshold of 5+ yielded sensitivity of 90% and specificity of 81% for the GDS-15 (Rinaldi et al., 2003). This study further supports the evidence from a previous study that higher cut-off threshold of around one to two points above recommended thresholds, may be more suitable for older
adults with physical illness in general hospitals. This can be explained by the high prevalence of physical symptoms and chronic physical health problems in our sample of older adults which can in turn lead to higher scores on the GDS-15. Questions like ‘do you feel full of energy?’; ‘do you often get bored?'; ‘do you feel pretty worthless the way you are now?’; and ‘have you dropped many of your activities and interests?’ are easily answered ‘yes’ when in hospital and therefore contribute to higher scores on the GDS.

**Proportion of participants scoring above recommended cut-off on the GDS-15**

A significant proportion of study participants scored above the recommended cut-off threshold of 5+ (37.6%). Comparing these findings with a similar UK study on medical inpatients, fewer participants scored above the recommended cut-offs of 5+ (44% in the study by Cullum et al., 2006 vs. 37.6% in this study) but very similar proportions scored above the optimal cut-off threshold of 7+ (28% vs 27.4%) in this study. Put simply, approximately one third of older adults in acute general hospital settings present with significant depressive symptoms (i.e. scoring above 7+), but only 60% of these patients (17.1% of the study sample) will have clinical depression according to ICD-10 criteria. (See Table 6.6).

**ROC curves and AUC**

The area under the ROC curve (95% CI) for the GDS-15 in this study was 0.88(0.81 to 0.96), which compares well with studies in similar settings (Pomeroy et al., 2001[AUC=0.80]; Rinaldi et al., 2003[AUC=0.88]). Higher AUC values indicate more accurate and superior performance of a screening tool compared to the gold standard. A value of 0.5 (AUC value for the 45 degree diagonal), represents a test that does not discriminate between those with the condition (depression) and those without. An area
under the ROC curve of 0.80 or above, as in this study therefore indicates very good performance. These data are summarised in Table 6.11.

**Predictive ability of the GDS-15 in clinical practice**

For screening purposes, particularly in general hospital settings, a cut-off that identifies as many people with the disorder as possible (i.e. higher sensitivity) is preferable when the disorder, in this case depression, is associated with high morbidity if untreated but has safe, inexpensive and effective treatment (Lawrie et al., 2000). NICE recommends that screening should be the first step of an assessment process and that further assessment by practitioners competent in mental health should follow any depression screening programme (NICE CG91, 2009). If the objective of any screening tool is to be achieved, there has to be the right balance of sensitivity, specificity and predictive values at the chosen cut-off threshold. In clinical practice, a highly sensitive test is desirable provided it is also able to correctly identify true sufferers of the condition it was designed to screen for. A highly sensitive test that is only able to identify very few people with the disorder (i.e. low PPV) would lead to unnecessary referrals of patients without the disorder to psychiatric services, increasing the workload and the need to use the ‘gold standard’ test in almost all cases. The best balance therefore would be to choose a cut-off threshold that maximizes sensitivity and NPV as much as possible without compromising specificity and PPV. For example, in choosing the cut-off threshold of 7+ in this study, three in five (60%) of those screening positive on the GDS-15 would be clinically depressed according to ICD-10 criteria while in those screened negative (scoring 6 and below on the GDS-15), only a very small proportion of cases (6%) would be missed (i.e. incorrectly identified as not clinically depressed). Increasing or decreasing the threshold would decrease or increase the sensitivity while having the opposite effect on positive and negative predictive values respectively. This ability of the GDS-15 to predict case level depression will vary
depending on the prevalence of depression in the sample and is likely to be much lower in community samples where the prevalence of depression is in the region of 1-5%.
### Table 7.1
Comparison of the performance of the GDS-15 in a number of published validation studies in inpatient settings

<table>
<thead>
<tr>
<th>Author, year</th>
<th>This study</th>
<th>Cullum et al., 2006</th>
<th>Pomeroy et al., 2001</th>
<th>Jackson et al., 1993</th>
<th>Rinaldi et al., 2003</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td></td>
<td>九龙医疗</td>
<td>九龙医疗</td>
<td>九龙医疗</td>
<td>九龙医疗</td>
</tr>
<tr>
<td>_sample size</td>
<td></td>
<td>118</td>
<td>618</td>
<td>87</td>
<td>59</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td>65+</td>
<td>65+</td>
<td>60+</td>
<td>65+</td>
</tr>
<tr>
<td>Diagnostic System</td>
<td></td>
<td>ICD-10</td>
<td>ICD-10</td>
<td>ICD-10</td>
<td>AGECAT</td>
</tr>
<tr>
<td>Cut-off</td>
<td></td>
<td>7+</td>
<td>7+</td>
<td>5+</td>
<td>5+</td>
</tr>
<tr>
<td>Sn/Sp</td>
<td></td>
<td>0.80/0.86</td>
<td>0.74/0.81</td>
<td>0.82/0.60</td>
<td>0.86/0.66</td>
</tr>
<tr>
<td>AUC</td>
<td></td>
<td>0.88</td>
<td>not stated</td>
<td>0.82</td>
<td>Not stated</td>
</tr>
</tbody>
</table>

Sn = Sensitivity at optimum cut-off  
Sp = Specificity at optimum cut-off.  
AUC = area under ROC curve.
2. The Edinburgh Depression Scale

In the original validation study of the 10-item EDS, a threshold score of 12/13 was found to have satisfactory sensitivity (86%), specificity (78%) and positive predictive values (73%) for detecting women with Research Diagnostic Criteria (RDC) depression. Lowering the threshold to 9/10 improved detection of cases to 90% (Cox et al., 1987) and this cut-off was considered suitable for screening postnatal depression in primary care settings. Although the EDS was originally developed and validated to screen depression in women in the postnatal period, it has been evaluated in samples outside the postnatal period (Cox et al., 1996) in patients with advanced metastatic cancer (Lloyd-Williams et al., 2000) and in patients who were terminally ill (Lloyd-Williams et al., 2004). In all these samples the EDS was found to be a reliable and valid depression screening instrument.

The EDS was designed with an emphasis on excluding somatic symptoms of depression which may be caused by the normal physiological changes associated with childbearing. The authors of the scale also wanted to produce an instrument that could be used by health workers without any specialist knowledge of psychiatry. The EDS is therefore very similar in construct to the GDS, which was designed to exclude somatic symptoms of depression that may be a feature of chronic physical illness in older adults. Applying the EDS to patients with diagnoses (i.e. physical illnesses,) frequently confused with the target disorder (i.e. depression) is another way of assessing whether the EDS is able to distinguish those with/without depression who are physically ill.

To our knowledge, the EDS has not been validated in older adults with physical illness who are admitted to hospital. This study therefore sought to establish whether the Edinburgh Depression Scale and other brief questionnaires are valid tools for screening for, and diagnosing depression in older people who are physically ill in hospital. Cases of depression were identified according to ICD-10 criteria and acted as gold standard.
Validity indices of the EDS compared with previous studies

The performance of the EDS as valid screening instrument was examined by obtaining data on sensitivity, specificity, positive predictive value, negative predictive value, and receiver operating characteristic (ROC) curve analysis at various cut-off thresholds. The ROC curve allowed us to calculate the optimum cut-off threshold and the point which maximises the area under the curve (AUC).

Sensitivity, specificity and cut-off thresholds

In our sample of older adults with physical illness, a cut-off threshold of 9+ was chosen as it gave the optimum balance of sensitivity and specificity without compromising the predictive ability of the EDS. At this threshold, the sensitivity and specificity of the EDS was 80% and 81% respectively and the positive predictive value (i.e. the proportion of patients scoring above this cut-off who are truly depressed) was 55%. This compares with sensitivity and specificity of 70% and 80% respectively in the study on patients with terminal illness (Lloyd-Williams et al., 2004) and sensitivity and specificity of 81% and 79% in a similar study on cancer patients (Lloyd-Williams et al., 2000). The cut-off threshold is lower than those obtained from these studies (cancer patients: [cut-off threshold: 13+]; (terminally ill: [cut-off threshold: 13+]) and may simply reflect the greater burden of physical complaints, side effects from medication and psychological distress in these patients. Most of the patients recruited into the study on cancer patients had to have a prognosis of 6 months or less with majority of the patients surviving a mean period of only 32.3 days (Lloyd-Williams et al., 2000). There was also a greater burden of depressive symptoms in these patients with mean score for the 10-item EDS of 10.9 for the cancer patients and a mean score of 10.5 for palliative care patients. The mean EDS score for our sample of older adults was 6.7. Prevalence rates for depressive disorder in the three studies were similar and ranged from 22% (this study and the study on cancer patients) to 27% (cancer patients: [prevalence rate: 27%]; (terminally ill: [prevalence rate: 32%]).
study on terminally ill patients) and is consistent with rates among hospitalized older adults, palliative care and cancer patients.

This study thus suggests that lower thresholds on the EDS may be more appropriate in samples with less severe physical complaints or illnesses and in populations where the prevalence of depressive symptoms is less.

**Proportion of participants scoring above cut-off thresholds on the EDS**

Using the optimum threshold of 9+ in this study sample, the proportion of older adults scoring above this cut-off threshold was 32.2% but only 55% of these patients (17.8% of the study sample) had clinical depression using ICD-10 criteria. This compares with 27% scoring above the recommended threshold of 13+ in the study on cancer patients, with 53% (14.3% of the study sample) correctly identified as depressed.

Put simply, approximately one third of older adults in acute general hospital settings present with significant depressive symptoms (i.e. score at or above 9 on the EDS), but only 55% of these patients (17.8% of the study sample) will have clinical depression according to ICD-10 criteria.

**ROC curves and AUC**

The area under the ROC curve (95% CI) for the EDS in this study was 0.91 (0.85 to 0.96), and is higher than the figure obtained in an earlier study on palliative care patients (Lloyd-Williams et al., 2004 [AUC = 0.709]). Although the samples and clinical setting differ in both studies, higher AUC values indicate a more accurate and superior performance compared to the gold standard with values of 0.80 and above, indicating very good performance. The reasons for this difference in the AUC values may also reflect the choice of gold standard used in these studies. In our study, the ICD-10 acted as the gold standard whilst in the study on palliative care patients, the DSM-IV acted as the external criterion.
measure. As the DSM-IV is more stringent than the ICD-10 in its criteria for a depressive episode, the differences in the scores may simply reflect better performance of the EDS on patients with less severe depressive disorders.

**Predictive ability of the EDS in clinical practice**

The ability of the EDS to correctly identify depression in patients who scored above the optimum cut-off of 9+ has already been discussed and is in the region of 55%. This is similar to the predictive ability of the EDS when used in cancer and terminally ill patients. As the objective of this study was to test the utility of the EDS in older adults with physical illness who are admitted to hospital, a cut-off threshold that allows true sufferers to be identified (high sensitivity) while at the same time minimising the proportion of cases that are missed (high NPV), remains the reason for choosing the cut-off threshold of 9+. Using this threshold in clinical practice and in settings with similar prevalence of depression as our study only a very small proportion of cases of depression (7%) would be missed.

Table 7.2 summarizes some published validation studies comparing the EDS with various standardized research interviews across a range of settings.
Table 7.2
Comparison of the performance of the EDS in a number of published validation studies

<table>
<thead>
<tr>
<th>Author, year</th>
<th>This study</th>
<th>Lloyd-Williams et al., 2004</th>
<th>Cox et al., 1987</th>
<th>Murray &amp; Carothers, 1990</th>
<th>Lloyd-Williams et al., 2000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>UK medical inpatients</td>
<td>terminally ill patients’</td>
<td>3 months postnatal</td>
<td>6 weeks postnatal</td>
<td>cancer patients (hospice/oncology)</td>
</tr>
<tr>
<td>Sample size</td>
<td>118</td>
<td>618</td>
<td>84</td>
<td>646</td>
<td>100</td>
</tr>
<tr>
<td>Age</td>
<td>65+</td>
<td>28-89</td>
<td>20s</td>
<td>20-40</td>
<td>25-69</td>
</tr>
<tr>
<td>Diagnostic System</td>
<td>ICD-10</td>
<td>DSM-IV</td>
<td>RDC</td>
<td>RDC</td>
<td>ICD-10</td>
</tr>
<tr>
<td>Cut-off</td>
<td>9+</td>
<td>13+</td>
<td>13+</td>
<td>13+</td>
<td>13+</td>
</tr>
<tr>
<td>Sn/Sp</td>
<td>0.80/0.81</td>
<td>0.70/0.80</td>
<td>0.86/0.78</td>
<td>0.68/0.96</td>
<td>0.81/0.79</td>
</tr>
<tr>
<td>AUC</td>
<td>0.91</td>
<td>0.71</td>
<td>not stated</td>
<td>not stated</td>
<td>not stated</td>
</tr>
<tr>
<td>PPV</td>
<td>0.55</td>
<td>0.56</td>
<td>0.73</td>
<td>0.67</td>
<td>0.53</td>
</tr>
</tbody>
</table>

Sn = Sensitivity at optimum cut-off  
Sp = Specificity at optimum cut-off.  
AUC = area under ROC curve  
RDC = Research Diagnostic Criteria (for depressive illness)
3. **The two questions suggested by NICE depression guidelines (NICE-2)**

The National Institute for Health and Clinical Excellence (NICE) in its updated clinical guidelines on the management of depression recommends that everyone who may be depressed, including those with a chronic physical health problem, should be asked if in the last month they have been feeling down, depressed or hopeless; and whether they are having little pleasure in doing things. Answering ‘yes’ to either question should warrant further assessment by a mental health practitioner competent in mental health assessment (NICE, 2009). A number of studies have been published examining the validity of these two questions as well as other ultra-short screening instruments (Mitchell and Coyne, 2007; Arroll et al., 2003 and Whooley et al., 1997). None of these studies tested the characteristics of the NICE-2 questions in older adults in general hospital settings, or validated the instrument simultaneously with more established screening tools using the ICD 10 or the DSM IV as gold standard. This study examined the performance of a two-question screening instrument (NICE-2), along with more established screening instruments (EDS, GDS-15 and GDS-4) using the ICD-10 diagnosis of depression as the external case criterion.

**Validity indices of the NICE-2 compared with previous studies**

The test characteristics of the NICE-2 screening questions was evaluated by obtaining data on sensitivity, specificity, positive and negative predictive value and receiver operating characteristic curve analysis at various cut-off thresholds.

**Sensitivity, specificity and cut-off thresholds**

Answering ‘yes’ to either of the two questions (i.e. applying a cut-off threshold of 0/1) gave a sensitivity of 100%, specificity of 70%, positive predictive value of 49% and negative predictive value of 100%. This compares with sensitivity of 96%, specificity of 57% and positive predictive value of 33% in another study applying the same threshold of
0/1, but on a sample of mostly male patients visiting an urgent care clinic (Whooley et al., 1997). Specificity improved to 66% when participants with concurrent substance abuse were excluded from the sample. The characteristics of this sample were different from those in our study in many ways. In particular, they were younger (mainly middle aged men) and were generally not as ill as those in our sample. The screening instrument was also tested against a different external criterion standard (the quick diagnostic interview schedule). Despite these differences, the two-question screening tool had high sensitivity and acceptable specificity. In our study, it would mean that the NICE-2 questions would identify everyone with depression according to ICD-10 criteria but wrongly identify 3 in 10 non-depressed people as depressed. This may be acceptable for instance in primary care where time constraints do not allow for detailed assessments and where the priority is to detect as many patients as possible with the disorder.

In another heterogeneous community sample (median age: 46) based in New Zealand, the two questions performed just as well as it did in this study. Sensitivity and specificity at the cut-off threshold of 0/1 (yes to either question) was 97% and 67% respectively although the computerised composite diagnostic interview was used as the gold standard (Arroll et al., 2003). Mitchell and Coyne conducted a meta-analysis to determine the accuracy of ultra-short screening instruments like the NICE-2. The pooled analysis for two- and three-item tests revealed overall sensitivity of 73.7% and specificity of 74.7% with a positive predictive value of only 38.3%. However, the studies included in the analysis were heterogeneous and involved samples from different settings with prevalence rates varying from 5% to 37% (Mitchell and Coyne, 2007).

Increasing the threshold to 2+ (i.e. answering ‘yes’ to both questions) made the NICE-2 questions less sensitive (76% vs. 100%) and more specific (90% vs. 70%) but with better positive predictive value (69% vs. 49%). This finding was similar to that obtained in the
study by Arroll et al., (2003). In our sample, we found the optimum threshold for the NICE-2 questions was a ‘yes’ answer to either question as identified on the ROC curve analysis. At this threshold, the proportion of participants answering ‘yes’ to either question, who are actually depressed was 49%, implying a high false positive rate.

Proportion of participants scoring above cut-off threshold on the NICE-2

Using the optimum cut off threshold of 0/1 in this study sample, the proportion of older adults scoring above this threshold was 44.9%, but only 49.1% of these (22% of study sample) had clinical depression using ICD 10 criteria (see Table 9). So if we use the NICE-2 questions as a screening instrument for depression in older adults in general hospital settings, approximately five in ten patients (50%) would be identified as having significant depressive symptoms, needing further comprehensive assessment (i.e. screen positive), but only half of these will have clinical depression using ICD10 criteria.

ROC curves and AUC

The ROC curve for the NICE-2 showed good agreement between scores on the test and ICD 10 diagnosis of depression (see Figure 5.4). The area under the curve (95% CI) for the NICE-2 of 0.92 (0.87- 0.97) was the highest of the four instruments compared against the ICD 10 in this study (Table 5.11 & Figure 5.5). It was also higher than that obtained in the study by Whooley et al., (2007) [AUC=0.82(0.78- 0.86)]. The AUC of the NICE-2 is almost 1 indicating that the performance of the NICE-2 as a screening instrument in this study sample was not only very good but also superior to the EDS, GDS-15 and GDS-4.

Predictive ability of the NICE-2

Despite the high sensitivity of this instrument in this sample, clinicians would want to know what a positive or negative test means to the patient who receives the test result. The
ability of the NICE-2 to correctly identify people who are depressed in those testing positive (answering ‘yes’ to either question) is probably more relevant clinically than the sensitivity. This is because predictive abilities vary depending on the prevalence of the disorder in the sample being screened. For example, in our sample of older adults with a prevalence of 22%, the ability of the NICE-2 to correctly identify depression in those answering ‘yes’ to either question is only 49.1% despite it being 100% sensitive. The flip side is that all of those who tested negative (i.e. answer ‘no’ to both questions) would be correctly free from depression. Mitchell and Coyne (2007) arrived at very similar findings in their meta-analysis of 22 studies. They found that two-screening instruments were accurate in identifying eight out of 10 depressed cases compared with a full psychiatric interview. However, when it came to predictive abilities of these instruments, only four out of 10 participants answering ‘yes’ to either question (positive test) were actually depressed. Two-item screening instruments were much better at ruling out a diagnosis (NPV >90%) if patients answer ‘no’ to both question (negative test). Clinicians using the NICE-2 as a screening tool can be more or less confident in telling their patients who screen negative that they are not suffering from depression, provided that the prevalence of depression in the population from which they are drawn is similar to that obtained in this study.

4. The 4-item Geriatric Depression Scale

The original validation study of the GDS-4 suggested a cut-off threshold of 1+ (0/1) as the optimum threshold (D’ Ath et al., 1994). Brevity and ease of administration was thought to be an advantage, particularly for older adults with physical illness. The GDS-4 has since been validated across a range of settings, including acutely ill geriatric inpatients (Shah et al., 1997; Goring et al., 2004), rehabilitation in patient unit (Pomeroy et al., 2001) using different external case criteria. In all these settings, the GDS-4 was found to have comparable sensitivity and specificity to longer scales. This study seeks to examine if the
findings from previous studies could be replicated in an English speaking population of older adults who are physically ill in hospital. It also compares the performance of the 4-item GDS with the GDS-15, the EDS and another very short instrument (NICE-2). The ICD-10 simultaneously acted as gold standard.

**Validity indices of the GDS-4 compared with previous studies**

The performance of the 4-item GDS in our study sample was evaluated by obtaining data on specificity, sensitivity, positive and negative predictive values at various cut-off thresholds, and calculating the area under the curve (AUC).

**Sensitivity, specificity and cut-off thresholds**

In our sample, we found the optimum threshold for the GDS-4 was a score of 2 or more (2+). At this threshold, sensitivity and specificity results for the GDS-4 was 80% and 84% respectively with positive predictive value of 58%. In another very similar study validating the GDS-4 among acutely ill geriatric inpatients, a lower threshold of 1+ was found to give the optimum sensitivity (75%), specificity (90%) and positive predictive value (86%) for their study sample (Shah et al., 1997). Although the setting is similar to our study, the sample size was much smaller (53 versus 118) and there were differences in the characteristics of the sample (all were over 80) and the external case criterion used (the Brief Assessment Schedule[BAS]). The ICD-10 or DSM-IV remains the gold standards in evaluating the test characteristics of any depression screening instrument. When applied among rehabilitation inpatients, using the ICD-10 as gold standard, The GDS-4 also performed well (sensitivity [82.4%]; specificity [67.1%]; PPV [37.8%]) but at a lower threshold of 1+ (Pomeroy et al., 2001). Goring et al., (2004) validated the GDS-4 as a screening tool for depression in medical inpatients using the GDS-30 as a gold standard. Although not an ideal gold standard measure, the GDS-4 performed well using the same
threshold of 2+ as in this study, with sensitivity and specificity of 78% and 74.5% respectively.

Table 7.3 shows summary statistics for the performance of the GDS-4 in some published validation studies.

Decreasing the threshold of the GDS-4 to 1+ in our study improves the sensitivity to 88% and but at the expense of being less specific (66%). It is difficult therefore to explain why a higher threshold produced optimum test characteristics for the GDS-4 in our sample of older adults. Higher cut-off thresholds may indicate greater burden of the depressive symptoms or physical illness that can mimic depressive symptoms leading to higher scores on our sample as a whole. The GDS-4 may also be worse at predicting ICD-10 case level depression, as shown by comparatively low PPV in our study, but better at picking up depressive symptoms using the BAS.

Despite these methodological differences, the GDS-4 performed well (AUC=0.85) as a screening instrument when compared to the GDS-15, EDS, and the NICE-2, although it was the worst overall performer in identifying depression. In our study for example, it would mean that the GDS-4 would identify 8 in ten people with depression according to ICD-10 criteria but wrongly identify less than 2 in 10 non-depressed people as depressed using the optimum cut-off of 2+. Again, given the brevity of the GDS-4 and the expected intention of the instrument to act as a first stage screen for depression, clinicians particularly those in primary care and general hospital settings, would be happy to accept high false positive rates (i.e. lower predictive ability of the positive test), provided they are able to identify as many depressed patients as possible for further intervention.
Proportion of participants scoring above cut-off threshold on the GDS-4

Using the optimum cut off threshold of 2+ in this study sample, the proportion of older adults scoring above this threshold was 45.3% but only 58% of these (18.8% of the study sample) had clinical depression using ICD-10 criteria (see Table 6.8). So if we use the GDS-4 questions as a screening instrument for depression in older adults in general hospital settings, approximately five in ten patients will screen positive (i.e. score 2 or more on the GDS-4), but only two thirds of these will have clinical depression using ICD-10 criteria.

ROC curves and AUC

The area under the curve (95% CI) for the GDS-4 of 0.85 (0.75-0.94) indicates that it performed significantly better than chance [p<0.0005]. Against the other screening tools, it fared less well, being the screening tool with the lowest area under the curve. (see Figure 6.5 & Table 6.11). It was however in line with those obtained in earlier validation studies using the ICD-10 (Table 7.3).

Predictive ability of the GDS-4

For screening purposes, a tool that is sensitive, i.e. identifies as many people as possible with depression, is preferable so that positive cases can be treated or referred for intervention. It is acceptable to have false positives especially when a further assessment (gold standard) will be carried out to exclude them. The GDS-4 accurately identified eight out of ten depressed cases compared with the ICD-10. It was also very good at ruling out a diagnosis of depression in patients who screen negative (i.e. score 1 or less) on the GDS-4. This finding is similar to what was obtained for the NICE-2 in this study and in the meta-analysis by Mitchell and Coyne (2007).
Table 7.3
Comparison of the performance of the GDS-4 in a number of published validation studies

<table>
<thead>
<tr>
<th>Author, year</th>
<th>This study</th>
<th>Pomeroy et al., 2004</th>
<th>Goring et al., 2004</th>
<th>Shah et al., 1997</th>
<th>D’Ath et al., 1994</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>UK medical inpatients</td>
<td>rehabilitation inpatients’</td>
<td>UK medical inpatient</td>
<td>geriatric inpatient</td>
<td>community sample</td>
</tr>
<tr>
<td>Sample size</td>
<td>118</td>
<td>87</td>
<td>153</td>
<td>53</td>
<td>120</td>
</tr>
<tr>
<td>Age</td>
<td>65+</td>
<td>60+</td>
<td>65+</td>
<td>80+</td>
<td>65+</td>
</tr>
<tr>
<td>Diagnostic System</td>
<td>ICD-10</td>
<td>ICD-10</td>
<td>GDS-30</td>
<td>BAS</td>
<td>GMS</td>
</tr>
<tr>
<td>Cut-off</td>
<td>2+ (1/2)</td>
<td>1+ (0/1)</td>
<td>2+ (1/2)</td>
<td>1+ (0/1)</td>
<td>1+ (0/1)</td>
</tr>
<tr>
<td>Sn/Sp</td>
<td>0.80/0.84</td>
<td>0.82/0.67</td>
<td>0.78/0.75</td>
<td>0.72/0.90</td>
<td>0.89/0.65</td>
</tr>
<tr>
<td>AUC</td>
<td>0.85</td>
<td>0.80</td>
<td>0.80</td>
<td>not stated</td>
<td>not stated</td>
</tr>
<tr>
<td>PPV</td>
<td>0.58</td>
<td>0.38</td>
<td>0.85</td>
<td>0.86</td>
<td>NPV=0.94</td>
</tr>
</tbody>
</table>

Sn = Sensitivity at optimum cut-off
Sp = Specificity at optimum cut-off.
AUC = area under ROC curve
GMS = Geriatric Mental Status Schedule.
PPV/NPV = Positive/Negative predictive values
BAS = Brief Assessment Schedule
5. The four screening instruments compared

The test characteristics and ROC curves for the four instruments are summarised in Table 6.11 and Figure 6.5. All four screening instruments performed well with good sensitivity (80% - 100%), satisfactory specificity (70% - 86%) and high negative predictive values (93 – 100%). The ROC AUC values varied from 0.85 to 0.91; this is a narrow range, and comparisons between tests using the standard error of the difference in AUCs did not find any differences that were statistically significant at the P=0.05 level. The considerable overlap in confidence intervals for the ROC AUC values of the screening instruments in table 6.11 also suggests that the study was not powered to find differences between tests. Although all 4 screening instruments performed adequately in this clinical setting with both good sensitivity and specificity, the area under the ROC curve was greatest for the NICE-2 (with also the narrowest confidence intervals). The strong performance of the two NICE questions (cut-off 1+) was as a consequence of its excellent sensitivity, though to some extent this is offset by only moderate specificity and a PPV of 49%.

In clinical practice, the choice of screening instrument will be influenced by other factors; for example, very short screening tools like the GDS-4 and the NICE-2 may well be preferable in clinical settings where time is of the essence such as in primary care. The longer tools may be preferable in hospital in-patient or out-patient settings, although where there is collaborative work with liaison teams the NICE-2 could also be the preferred screening instrument for use by ward staff as it performs well. In this situation, the NICE-2 could be a first stage screening tool, followed by the longer EDS or GDS to improve accuracy and reduce the number of false positive cases referred to specialist mental health services. This possibility will be explored further in future analyses of this data set.
7.6   Subsyndromal or sub threshold depression

Milder forms of depression (subthreshold depression) are also recognised as being a risk factor for future major depression and therefore require some means of identification. Identifying these cases in practice and in research is fraught with some difficulty as the current diagnostic systems do not have this category of sub threshold depression as a formal diagnosis. The research criteria of the DSM-IV define sub threshold depression as minor depression with at least 2 symptoms but less than 5 symptoms from a list of 9 depressive symptoms. Sub threshold depression therefore includes ICD-10 mild depressive episode with only 4 symptoms.

As a secondary objective of this study, the ability of the screening tools to detect ICD-10 mild depressive episode was the closest estimation of how these tools would perform when used to screen for sub threshold depression. The results in table 6.12 reveal an interesting finding in that the shortest of the 4 screening tools (NICE-2), proved to be better at identifying milder forms of depression and may therefore be useful in primary care where these patients are likely to present. Although in identifying these cases there may also be correspondingly high numbers of false positives, it may then be possible to use a longer instrument like the EDS or GDS as a second stage screening process to eliminate the false positives.
7.7 Strengths and weaknesses of this study

Strengths

This study had several strengths which are summarised below

- **Sample composition**

Sample size was comparable to those of other studies conducted in inpatient settings and yielded a prevalence of depression of 22%. In this area of research, the opinion is that studies with less than 10 depression cases (i.e. persons fulfilling gold standard case criteria) are judged as insufficient for accurate estimates of sensitivity and specificity (Wancata et al., 2006). This study had a total of 26 cases of ICD-10 depressive episode with prevalence figures similar to those obtained from previous studies. It is therefore likely that the sample size of 118 was sufficient enough to enable accurate data to be obtained on sensitivity and specificity and predictive values for all of the screening instruments. Recruitment of participants was random and the inclusion criteria represented what could be described as real life everyday NHS practice.

- **The clinical assessments**

The use of the PSE-SCAN to generate diagnoses of depression according to ICD-10 criteria ensured that assessments were standardised and therefore more reliable than a clinical psychiatric interview. The research interviews were conducted by clinicians with experience in psychiatry hence making the results more reliable and less prone to errors. The psychiatrists performing the diagnostic interviews (PSE-SCAN) were blind to the results of the screening assessments and vice versa, eliminating any review bias that may have arisen. ‘Verification bias’ was avoided by applying both the ‘gold standard’ PSE-SCAN interviews and screening assessments to the whole sample regardless of the result of the PSE-SCAN or screening interview. The time interval between both interviews was
short (5 working days), to ensure so far as possible that the diagnostic status of the patients remained stable between both assessments.

- **Data analysis**

  Sensitivity and specificity was reported for various cut-offs for all screening tools and ROC curve analysis done to determine optimum cut-off for all four instruments. Screening tools were compared simultaneously on all patients using the ICD-10 as the external case criterion. The calculations of sensitivity and specificity were more accurate because the research interview was applied to all eligible participants regardless of the diagnosis or score on the screening instruments.

**Weaknesses/ Limitations**

- **Length of the study**

  One main weakness of this study was the length of time it took to complete. The expectation was that recruitment of participants would have been completed within a year of starting and data analysis and results of the study ready in another 6 months. Unfortunately there were several delays throughout the duration of the study. Wards were closed periodically as a result of hospital acquired infections and a significant proportion of over 65s on these wards had cognitive impairments that excluded them from the study. There was also the practice in most acute hospitals to reduce lengths of stay in hospital meaning that there was only a very limited window of opportunity to recruit potential participants. Patients deemed medically fit were very often transferred to intermediate care wards or community hospitals for rehabilitation, making it very difficult for the author who had one day a week dedicated to research to recruit potential participants. During the course of the study, the author became a consultant psychiatrist in Lincolnshire and had to
suspend recruitment for some time, until another registrar based in Leicester had undergone all the necessary training required for this research was available to assist.

- **Sample bias**

This sample may not be truly representative of older adults in acute medical settings as several patients who met the inclusion criteria had been discharged home or transferred to other wards making it difficult to collect data on them. There were also patients who did not wish to take part in a psychiatric research despite meeting our inclusion criteria. Unfortunately this study was not designed to assess the characteristics of those participants who did not take part or give consent. Some of these patients may have been suffering from depression but were not included in this study and it is therefore possible that those who refused to take part may have affected the results obtained from this study.

- **Sample size**

Although comparable to other studies, the sample size of 118 was relatively small. Sample size calculations are important in research because it determines whether a study is sufficiently powered to detect important effects. If the sample is too small, it may fail to demonstrate important effects which are truly present in a population (false negative result or type II errors) or obtain a false positive result by chance.

This study was not powered to find differences between tests and may therefore have failed to demonstrate important differences between the screening instruments. This lack of power is reflected in the considerable overlap in confidence intervals for the ROC AUC values of the screening instruments in table 6.11 and the comparisons between tests using the standard error of the difference in AUCs. The most likely explanation for this is the relatively small sample size hence future research would require a larger sample size to
improve statistical power and reduce the risk of type II errors or false negative results. Despite this limitation affecting the secondary aim of the study, so far as the principal aim is concerned the high prevalence of depression in this study sample permitted accurate statistical calculations on sensitivity, specificity and positive and negative predictive values of the screening tools.

• **Exclusion criteria**

This study excluded patients with MMSE less than 24 because the tools had not been validated in patients with significant cognitive impairment. Yet this particular group of patients are often depressed particularly in acute hospital settings. It is also possible that some patients with MMSE scores above 24 may also be cognitively impaired affecting the accuracy of the results. Validating these tools in these patients is therefore important and represents an area for future research.

• **Diagnostic criteria**

The ICD-10 was used as gold standard in this study and is a system based on symptom count. This method of diagnosis ignores the vast numbers of older adults suffering from subsyndromal depression who suffer as much or even more than those with syndromal or ICD-10 case level depression. In this study, the performance of the screening tools in evaluating sub threshold depression was not evaluated fully as it was not the primary objective of this research. However the results obtained from the preliminary analysis of mild ICD-10 depression (a subtype of DSM-IV minor/subsyndromal depression), indicates that most of the screening tools employed failed to detect these cases. The importance of identifying sub threshold depression was highlighted in the recently updated NICE depression guidelines and it was for this reason, they adopted the DSM-IV criteria. Further analysis using the Structured Clinical Interview for DSM-IV (SCID) to make
depression diagnosis would enable accurate estimates of patients suffering from minor or subsyndromal depression that were not captured by ICD-10.

7.8 Conclusions & Clinical implications

This study has added to the UK evidence base and further confirms the findings from other similar studies that depression is more prevalent in older adults who are physically ill in hospital. It has also evaluated the performance of the Edinburgh Depression Scale and the two NICE questions as screening tools for depression and found them to have performed as well as one of the most established depression screening tools in older people. This means that clinicians in acute medical settings now have a choice of at least three easy to administer screening instruments that can accurately screen for depression in a variety of settings. Answering positively to one or more of the two recommended NICE screening questions would appear an appropriate initial screening tool for depression in older people in the acute medical setting. People identified as possibly depressed by this instrument would, however, need to have a further more detailed assessment to help determine whether they have clinically relevant depression. Opinion from specialist liaison mental health professionals may be required to ensure adequate treatment and follow up of identified cases. The findings from this study may be implemented within the NHS to improve the low recognition rates of depression identified in community and acute medical facilities. This will not only help in reducing the length of stay in hospital by the depressed older adult, but also improve the quality of life of the patient leading to lower morbidity and mortality. A key consideration in using any of these instruments is choosing an optimum cut-off threshold as it likely to be different for different populations.
7.9 Suggestions for further research

Validating the screening tools in patients with cognitive impairment would be a useful next step as this group of patients are over represented on geriatric inpatient units and have prevalence rates for depression in the region of 30%

Secondly, the clinical significance of sub threshold depression amongst older people is well documented in literature including the recent UK depression guidelines. Sub-threshold depression requires identification and treatment as they are associated with significant morbidity and is a risk factor for syndromal depression. The performance of these screening instruments in identifying sub threshold depression should be a subject for further research. It may be possible to use the research criteria for DSM-IV or the SCID to accurately identify all cases of sub threshold depression and then carry out further analysis on these samples.

The 2 stage screening process advocated by UK clinical guidelines on the management of depression is another area for further research. The NICE-2 questions being 100% sensitive would act as the first stage screening instrument identifying all cases of depression, followed by either the GDS-15, EDS or GDS-4, depending on how they perform at the 2nd stage. The advantage of the two stage screening in primary care or general hospital settings is that by requiring patients to screen positive on the two tests, a number of false positive cases are eliminated hence reducing the workload for the psychiatric services.
SECTION FIVE: APPENDICES
Appendix 1
DSM-IV criteria for major depressive episode

A: Five or more of the following symptoms have been present nearly every day during the same 2-week period and represent a change from previous functioning; at least one is either (1) depressed mood or (2) loss of interest or pleasure. Symptoms that are clearly due to a general medical condition should not be counted.
   1. Depressed mood most of the day
   2. Markedly diminished interest or pleasure in all, or almost all, activities most of the day
   3. Significant weight loss when not dieting or weight gain, or decrease or increase in appetite
   4. Insomnia or hypersomnia
   5. Psychomotor agitation or retardation
   6. Fatigue or loss of energy
   7. Feelings of worthlessness or excessive or inappropriate guilt (which may be delusional)
   8. Diminished ability to think or concentrate, or indecisiveness
   9. Recurrent thoughts of death (not just fear of dying), recurrent suicidal ideation without a specific a plan, or suicide attempt or a specific plan for committing suicide.

The symptoms

B: Do not meet criteria for a mixed episode

C: Cause clinically significant stress or impairment in social, occupational, or important areas of functioning

D: Are not due to the direct physiological effects of a substance or a general medical condition.

E: Are not better accounted for by bereavement; the symptoms persist for longer than 2 months or are characterised by marked functional impairment, morbid preoccupation with worthlessness, suicidal ideation, psychotic symptoms, or psychomotor retardation

Specifiers can be coded for severity (mild, moderate, or severe); psychosis (delusions or hallucinations); and remission (partial or full)
Appendix 2

ICD-10 depressive episode

A: The depressive episode should last for at least 2 weeks; there is no history of mania; and depressive episode is not attributable to organic disease or psychoactive substance.

Mild depressive episode

B: At least two of the following three symptoms must be present:
   1. Depressed mood to a degree that is definitely abnormal for individual, most of the day and almost every day; and sustained for at least 2 weeks
   2. Loss of interest or pleasure in activities that are normally pleasurable
   3. Decreased energy and increased fatigability

C: An additional symptom or symptoms from the following list should be present, to give at least four:
   1. Loss of confidence or self esteem
   2. Unreasonable feelings of self-reproach or excessive and inappropriate guilt
   3. Recurrent thoughts of death or suicide, or any suicidal behaviour
   4. Complaints or evidence of diminished ability to think or concentrate
   5. Change in psychomotor activity, with agitation or retardation (either subjective or objective)
   6. Sleep disturbance of any type

Moderate depressive episode

At least two from B and additional symptoms from C, to give a total of at least six symptoms

Severe depressive episode

All three from B and at least five from C to give a total of eight symptoms
Appendix 3

EDINBURGH DEPRESSION SCALE

Please UNDERLINE the answer which comes closest to how you have felt IN THE PAST 7 DAYS, not just how you feel today.

I have been able to laugh and see the funny side of things
As much as I always could
Not quite so much now
Definitely not so much now
Not at all

Things have been getting on top of me
Most of the time and I haven’t been able to cope at all
Yes sometimes I haven’t been coping as well as usual
No most of the time I have coped quite well
No I have been coping as well as ever

I have looked forward with enjoyment to things
As much as I ever did
Rather less than I used to
Definitely less than I used to
Hardly at all

I have been so unhappy that I have had difficulty sleeping
Yes most of the time
Yes quite often
Not very often
No, not at all

I have blamed myself unnecessarily when things go wrong
Yes most of the time
Yes some of the time
Not very often
No, never

I have felt sad or miserable
Yes most of the time
Yes, quite often
Not very often
No, not at all
I have been anxious or worried for no good reason
Not at all
Hardly ever
Yes sometimes
Yes very often

I have been so unhappy I have been crying
Yes, most of the time
Yes, quite often
Only occasionally
No, never

I get a sort of frightened feeling as if something awful is about to happen
Very definitely and quite badly
Yes but not too badly
A little, but it doesn’t worry me
Not at all

The thought of harming myself has occurred to me
Yes, quite often
Sometimes
Hardly ever
Never
Appendix 4

Name/ID No:

**GERIATRIC DEPRESSION SCALE**

*Please choose the answer that best describes how you have felt over the last week.*

*Please answer all the following questions by ringing either “Yes” or “No”*

1.  *Are you basically satisfied with your life?*  
   Yes/No

2.  Have you dropped many of your activities and interests?  
   Yes/No

3.  *Do you feel that your life is empty?*  
   Yes/No

4.  Do you often get bored?  
   Yes/No

5.  Are you in good spirits most of the time?  
   Yes/No

6.  *Are you afraid that something bad is going to happen to you?*  
   Yes/No

7.  *Do you feel happy most of the time?*  
   Yes/No

8.  Do you often feel helpless?  
   Yes/No

9.  Do you prefer to stay at home, rather than going out and doing new things?  
   Yes/No

10. Do you feel you have more problems with memory than most?  
    Yes/No

11. Do you think it is wonderful to be alive now?  
    Yes/No

12. Do you feel pretty worthless the way you are now?  
    Yes/No

13. Do you feel full of energy?  
    Yes/No

14. Do you feel that your situation is hopeless?  
    Yes/No

15. Do you think that most people are better off than you are?  
    Yes/No

*Items included in the 4-item scale*
Appendix 5

Name/ID No:

NICE Depression Screening Questions

1. During the last month, have you often been bothered by feeling down, depressed or hopeless? Yes/No

2. During the last month, have you often been bothered by having little interest or pleasure in doing things? Yes/No
Screening for Depression in elderly physically ill using the Edinburgh Depression Rating Scale

CONSENT FORM

Please initial each box

I confirm that I have read and understand the information sheet dated 10.04.06 (version 3) for the above study and have had the opportunity to ask questions.

I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, and without my medical care or legal rights being affected.

I agree to my hospital doctor or GP being informed of the findings of the study, if it is felt appropriate by the research team.

I agree to take part in the above study.

Name of patient Date Signature

Name of researcher Date Signature

Principal Investigator: Dr Collins Esiwe, Specialist Registrar, Psychiatry for the Elderly, Department of Health Sciences, Leicester General Hospital. Telephone: 0116 258 4597

Consent Form – version 2.1 (10/05/06)
Appendix 7

National Research Ethics Service
Leicestershire, Northamptonshire & Rutland Research Ethics Committee 1

1 Standard Court
Park Row
Nottingham
NG1 5GN

Telephone: 01159123344
Facsimile: 01159123300

25 May 2006

Dr C I Esiwe
Specialist Registrar
Leicestershire Partnership NHS Trust
Clinical Division of Psychiatry, Dept., of Health Sciences,
Psychiatry for the Elderly
Leicester General Hospital
Gwendolen Road
Leicester, LE5 4PW

Dear Dr Esiwe,

Full title of study: The performance of the Edinburgh Depression Rating Scale (EDRS) as a screening tool for depression in older people in the acute general hospital setting

REC reference number: 06/Q2501/39

Thank you for your letter of 10 May 2006, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised.

Ethical review of research sites

The favourable opinion applies to the research sites listed on the attached form.

Conditions of approval

The favourable opinion is given provided that you comply with the conditions set out in the attached document. You are advised to study the conditions carefully.

This Research Ethics Committee is an advisory committee to East Midlands Strategic Health Authority
The National Research Ethics Service (NRES) represents the NRES Directorate within the National Patient Safety Agency and Research Ethics Committees in England

NP4 1972

125
Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

<table>
<thead>
<tr>
<th>Document</th>
<th>Version</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application</td>
<td></td>
<td>13 February 2006</td>
</tr>
<tr>
<td>Investigator CV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investigator CV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protocol</td>
<td>2.1</td>
<td>09 May 2006</td>
</tr>
<tr>
<td>Participant Information Sheet:</td>
<td>3</td>
<td>10 April 2006</td>
</tr>
<tr>
<td>Participant Consent Form:</td>
<td>2.1</td>
<td>10 May 2006</td>
</tr>
<tr>
<td>Response to Request for Further</td>
<td></td>
<td>10 May 2006</td>
</tr>
<tr>
<td>Information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Edinburgh Depression Scale</td>
<td>1.0</td>
<td>06 February 2006</td>
</tr>
<tr>
<td>Geriatric Depression Scale</td>
<td>1.0</td>
<td>06 February 2006</td>
</tr>
<tr>
<td>NICE Depression Screening Questions</td>
<td>1.0</td>
<td>06 February 2006</td>
</tr>
</tbody>
</table>

Research governance approval

The study should not commence at any NHS site until the local Principal Investigator has obtained final research governance approval from the R&D Department for the relevant NHS care organisation.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

06/Q2501/39 Please quote this number on all correspondence

With the Committee’s best wishes for the success of this project

Yours sincerely

Dr C Edwards/Ms L Ellis
Chair/Co-ordinator

Email: linda.ellis@rushcliffe-pct.nhs.uk

Enclosures: Standard approval conditions
Site approval form

Copy to: Mr D Clarke
Leicestershire Partnership NHS Trust
Lakeside House 4 Smith Way,
4 Smith Way,
Leicester, LE19 1SS

R&D Department for NHS care organisation at lead site - UHL
Leicestershire, Northamptonshire & Rutland Research Ethics Committee 1

LIST OF SITES WITH A FAVOURABLE ETHICAL OPINION

For all studies requiring site-specific assessment, this form is issued by the main REC to the Chief Investigator and sponsor with the favourable opinion letter and following subsequent notifications from site assessors. For issue 2 onwards, all sites with a favourable opinion are listed, adding the new sites approved.

<table>
<thead>
<tr>
<th>REC reference number:</th>
<th>Issue number:</th>
<th>Date of issue:</th>
</tr>
</thead>
<tbody>
<tr>
<td>06/Q2501/39</td>
<td>1</td>
<td>25 May 2006</td>
</tr>
</tbody>
</table>

Chief Investigator: Dr C I Esiwe

Full title of study: The performance of the Edinburgh Depression Rating Scale (EDRS) as a screening tool for depression in older people in the acute general hospital setting

This study was given a favourable ethical opinion by Leicestershire, Northamptonshire & Rutland Research Ethics Committee 1 on 24 May 2006. The favourable opinion is extended to each of the sites listed below. The research may commence at each NHS site when management approval from the relevant NHS care organisation has been confirmed.

<table>
<thead>
<tr>
<th>Principal Investigator</th>
<th>Post</th>
<th>Research site</th>
<th>Site assessor</th>
<th>Date of favourable opinion for this site</th>
<th>Notes (1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dr C I Esiwe</td>
<td>Specialist Registrar</td>
<td>Leicester General Hospital</td>
<td>Leicestershire, Northamptonshire &amp; Rutland Research Ethics Committee 1</td>
<td>25/05/2006</td>
<td></td>
</tr>
</tbody>
</table>

Approved by the Chair on behalf of the REC:

(Delete as applicable)

(Signature of Chair/Administrator)

(Name)

(1) The notes column may be used by the main REC to record the early closure or withdrawal of a site (where notified by the Chief Investigator or sponsor), the suspension of termination of the favourable opinion for an individual site, or any other relevant development. The date should be recorded.
Appendix 8

Leicestershire Partnership
NHS Trust

Screening for Depression in elderly physically ill using the Edinburgh Depression Rating Scale

INFORMATION SHEET FOR PARTICIPANTS

You are being invited to take part in a research study. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

What is the purpose of this study?
The study has been set up to look at a screening tool for depression in people over the age of 65 years in a general hospital. It is hoped that this will lead to the development of services that will be more sensitive to the needs of this population group.

Why have I been chosen?
We will be asking all patients over the age of 65 who are admitted to general medical wards at LGH to consider taking part in the research. We will be including approximately 100 patients in the study.

Do I have to take part?
It is up to you to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep and be asked to sign a consent form. You will also be given a copy of this to keep. If you decide to take part you are still free to withdraw at any time and without giving a reason. A decision to withdraw at any time, or a decision not to take part, will not affect the standard of care you receive.

What will happen to me if I take part?
An interview will be completed while you are a patient on a general medical ward, this may take up to 60 minutes and will involve you answering questions about how you are feeling, your mood, and testing your memory. The purpose of the interview is to find out if you are suffering from any mental health difficulties, in particular depression. The first part of the interview involves a memory test; if you have difficulties with the memory test then we will not be asking you any further questions. Within the following few days you will be asked by another researcher, to complete three brief questionnaires (also about your mood and how you are feeling), this should only take 10-15 minutes maximum.
If, following the interview and the questionnaires, we feel that you may be suffering from depression we will contact your medical team and advise them regarding appropriate treatment.

What are the possible benefits of taking part?
We do not anticipate any specific benefits for those who take part in the research. However, if we feel you are suffering from depression we will be able arrange suitable treatment for you. The information we get from this study may help us to be better at detecting when future patients over the age of 65 who are admitted to General Hospital wards are suffering from depression.
What if I have a complaint about the research?
If you wish to complain, or have any concerns about any aspect of the way you have been approached or treated during the course of this study, the normal National Health Service complaints mechanisms should be available to you.

Will my taking part in this study be kept confidential?
All information which is collected about you during the course of the research will be kept strictly confidential. Your GP will be informed of the findings of this study, if it is felt appropriate by the research team.

What will happen to the results of the study?
The results of this study will be reported to appropriate colleagues locally, and will be published in the relevant scientific journals. Findings from the study will be presented in a way that will not identify individual participants in any way.

Who is organising and funding the research?
The research has been organised by the University of Leicester, Leicestershire Partnership NHS Trust and University Hospitals of Leicester NHS Trust. Minimal running costs that may be incurred in carrying out this study will be borne by the Department of Health Sciences, Psychiatry for the Elderly, University of Leicester.

Thank you for taking the time to read this information sheet.

Contact for further information:
Please keep this sheet for your information. If you have any questions, please do not hesitate to contact the following:

Dr Collins Iheanyichukwu Esiwe, Specialist Registrar, Psychiatry for the Elderly, Department of Health Sciences, Leicester General Hospital. Telephone: 0116 258 4597.

For independent advice regarding participating in research you can contact the local Patient Advice and Liaison Service (PALS) at Leicester General Hospital – 0116 2588295.
SECTION SIX: REFERENCES


Almeida OP, Almeida SA. Short versions of the Geriatric Depression Scale: A study of their validity for the diagnosis of a major depressive episode according to ICD-10 and DSM-IV. Int. J. Geriat. Psychiatry. 1999; 14, 858- 865


Geerlings SW, Beekman ATF, Deeg DJH, Twisk JWR, and Van Tilburg W. Duration and severity of depression predict mortality in older adults in the community. Psychological Medicine 2002; 32: 609-618


Gilbody S, Sheldon T, Wessely S. Should we screen for depression? BMJ 2006; 332: 1027-1030


Lloyd-Williams M, Dennis M, Taylor F. A prospective study to compare three depression screening tools in patients who are terminally ill. *General Hospital Psychiatry* 2004; 26: 384-389.

MacHale S. Managing depression in physical illness. *Advances in Psychiatric Treatment* 2002; 8: 297-305.


Murphy E. Social origins of depression in old age. *British Journal of psychiatry* 1982; **141**: 135-142.


National Screening Committee UK; Criteria for appraising the viability, effectiveness and appropriateness of a screening programme 2009.


Reynolds CF, Frank E, Perel JM *et al.* Nortriptyline and interpersonal psychotherapy as maintenance therapies for recurrent major depression: a randomized controlled trial in patients older than 59 years. *JAMA* 1999; 281:39-45.


Wing JK, Babor T, Brugha T, Burke J, Cooper JE, Giel R, Jablenski A, Reiger D, Sartorius N. Schedules for Clinical Assessment in Neuropsychiatry(SCAN). *Arch Gen Psychiatry* 1990; 7: 589-593.
