Quantitative and Qualitative Evaluation of Drug-Induced Parkinsonism

Martin J. Dobson

PhD

The University of Edinburgh

2001

Acknowledgements

I would like to thank my supervisors, Dr. D.G.C. Owens and Professor E. Johnstone, for all their help and support during the completion of this thesis. Dr. Owens conducted the clinical assessments for the study and provided encouragement during the later stages of my work. The assistance of Dr. M. Glabus and Dr. I. Macleod who prepared accelerometry hardware and software was invaluable. Dr. E.G. Walsh must be thanked for the generous loan of his rigidity assessment device. I am also indebted to my partner, Philippa Matthews, without whose support I would not have finished this work. Finally, I am extremely grateful to all the participants in this study for their time and co-operation.

Declaration

I declare that all instrumental assessments for this study were undertaken by me, that all data analysis was conducted by me, and that this thesis was composed by me.

16/02/2007

Martin J. Dobson

Abstract

Antipsychotic medications are ubiquitous in the treatment of psychosis. However, relief from positive symptomatology comes at a price. Extrapyramidal side-effects such as drug-induced parkinsonism (DIP) are common and superficial similarities between features of parkinsonism and those of psychosis hinder efforts to calibrate dosages. The boundary between psychopathology and drug-induced disorder is a major conceptual issue in psychiatry. Instrumental assessment promises the opportunity to more accurately gauge this boundary.

Three hypotheses were developed: that instrumentation has a role in the assessment of DIP, that bradykinesia is the predominant feature of DIP, and that cognitive and subjective features of parkinsonism are present in DIP. Instrumentation procedures were selected to objectively assess the three major features of parkinsonism: bradykinesia, rigidity, and tremor. Subjective ratings of symptomatology associated with psychosis and antipsychotic medication were taken. All the measures used were evaluated empirically relative to standard observer rating criteria and the constructs underlying the assessments were examined.

The instrumental assessment techniques demonstrated moderate to high accuracy though most did not display significant advantages over clinical rating procedures. However, a role was proposed for performance measures in regular monitoring of bradykinesia. Stronger support was found for the latter two hypotheses. Results indicated that a greater degree and prevalence of abnormality relative to the control group was present in bradykinesia than the other features of parkinsonism. Empirical evidence demonstrated the presence of a cognitive deficit in behaviour associated with the presence of parkinsonism.

The evidence from the study also bears on issues of drug tolerability. Support was provided for suggestions that the atypical antipsychotic, clozapine, has a uniquely low liability to induce parkinsonism.

Table of contents

1 Introduction

1.1 Psychosis

1.1.1 Conceptions of psychosis

Psychosis has been recognised in human societies for millennia and has been accounted for in many different ways. In many societies psychosis was explained in terms of supernatural phenomena such as possession by malevolent spirits or demons (Zilboorg & Henry, 1941). However, there is evidence that since at least as early as the Middle Ages psychosis could be regarded in terms of sickness, and thus might be amenable to medicine rather than exorcism (Allderidge, 1979).

Depressive disorders, often referred to as melancholy, were known to be distinct from the florid insanity of schizophrenia and mania. Melancholy was thought to result from a disturbance of the humours, specifically a predominance of black bile. The first formal account of psychosis was produced by Kraepelin in the 19th century (see Kraepelin, 1986) who classified the previously uncategorised insane masses into those suffering schizophrenia (which he termed *dementia praecox*) and those suffering affective disorder.

The major psychoses have always been viewed in the context of prevailing theories of mental function and dysfunction. Different schools of thought have accounted for psychosis using, among others, biological models, psychodynamic models, and learning models. In recent years, biological accounts of psychosis have increasingly taken precedence over other accounts.

Despite the influence of Kraepelin's classification of psychosis into two forms, the concept of a single psychosis is an enduring one. Noted formulations of this hypothesis have been published by Greisinger, who used the term "enheitepsychose" to identify the concept he described, and Crow (1995).

Crow's account posits a continuum disorder with varying degrees of affective and cognitive dysfunction. At one end of the continuum lies a disorder characterised by purely affective disturbance, at the other a disorder characterised by purely cognitive disturbance. Terms such as schizophrenia and major affective disorder are thus labels for patients whose condition may be represented as being towards one other end of the continuum.

Accounts such as these emphasise the considerable overlap between schizophrenia and major affective disorder in a number of parameters, including clinical phenomenology, treatment response, outcome measures, and psychosocial competence. It is true that significant similarities in presentation exist between many cases of schizophrenia and major affective disorder, that similar if not identical medication regimes are frequently used to treat schizophrenia and major affective disorders, that the same rating scales may be used to assess severity of symptomatology in schizophrenia and major affective disorder, and that social competence may be similarly impaired in schizophrenia and major affective disorder. However, most workers now regard the distinction between schizophrenia and affective disorder as clinically valid.

1.1.2 Models of symptomatology

Recent research investigations into the aetiology of psychosis have been concerned primarily with biological correlates of observable features of psychosis, particularly the negative symptoms of schizophrenia. The distinction between positive and negative symptoms is a concept used originally in neurology (Berrios, 1985). Positive symptoms are those abnormal by their presence. In schizophrenia these include hallucinations, delusions, certain forms of thought disorder, and some bizarre behaviours (Carpenter et al., 1988). Negative symptoms are those characterised by an absence of normal function. These may include affective blunting, ideational constriction, poverty of speech, diminished sense of purpose, and reduce social drive (Carpenter et al., 1988).

The terms "negative symptoms" and "deficit symptoms" are sometimes used as if interchangeable though other authors distinguish them in terms of the permanency of the symptoms (Carpenter et al., 1988; Fenton & McGlashan, 1994). Within this conception, the term "negative symptoms" is used only as a descriptive term and does not imply causality. Deficit symptoms are thus characterised as negative symptoms which are intrinsic to the disease process and are enduring and permanent.

Crow (1980) postulated the existence of two forms of schizophrenia, termed type I and type II. Patients with type I schizophrenia exhibit normal brain morphology and display only positive symptoms of schizophrenia. In contrast, patients with type II schizophrenia exhibit abnormal brain morphology and may display prominent deficit

symptoms in addition to positive symptoms. Within this account, all patients will exhibit positive symptoms at some time, though only those with type II schizophrenia will exhibit deficit symptoms. Although the deficit symptoms of type II schizophrenia may be absent early in the course of the illness, once they develop they are irreversible.

Type I	**Type II**
acute onset	insidious onset
positive symptoms only	prominent deficit symptoms
normal intellectual performance	progressive dementia
good response to medication	poor response to medication
no abnormal brain morphology	abnormal brain morphology

However, other workers have suggested that there may be more than two types of schizophrenia. Andreasen et al. propose that three types of schizophrenia may be distinguished (Andreasen et al., 1990). Using the Scale for the Assessment of Negative Symptoms (SANS; Andreasen, 1983) and the Scale for the Assessment of Positive Symptoms (SAPS; Andreasen, 1984), patients may be divided into "positive", "negative", and "mixed" groups dependent upon the form of symptoms predominating. A similar pattern of results has been found (Kay, 1991) using the Positive and Negative Syndrome Scale (PANSS; Kay et al., 1987).

Further, cross-sectional analyses of symptom prevalence have suggested that the delineation of symptomatology into two (positive and negative) categories may be inadequate. On the basis of factor analysis Liddle (1987) segregated individual symptoms into three syndromes. The symptoms formed clusters which Liddle termed

psychomotor poverty, disorganisation, and reality distortion. Other authors have described four or even five domains of psychopathology (Carpenter et al., 1988).

1.1.3 Accounts of deficit symptomatology

Despite the above caveats, the concept of deficit symptomatology in psychosis remains a valid and useful one. Deficit symptomatology has been recognised as a central component of psychosis since at least the time of Kraepelin (who termed schizophrenia dementia praecox), and retains a pivotal role in modern conceptions of schizophrenia such as that of Crow (1980). Within Crow's formulation deficit symptomatology is regarded as being present only in a proportion of cases of schizophrenia but progressive and irreversible when present. The severity of deficit symptomatology is held to be independent of the severity of positive symptomatology.

However, other authors argue that the severity of positive and negative symptoms are negatively correlated, and that the positive-negative symptom distinction represents opposing ends of the symptom continuum (Andreasen and Olsen, 1982).

The independence of positive and deficit symptomatologies is supported by the work of Carpenter et al., (1988) who make the fundamental distinction previously noted between primary and secondary negative symptoms. Primary negative symptoms are persistent and pervasive deficits which are intrinsic features of the illness. In contrast, secondary negative symptoms are more effervescent state phenomena present due to factors such as drug effects or a lack of social stimulation.

Carpenter et al. (1988) state that well-informed physicians are able to reliably distinguish primary and secondary negative symptoms, enabling the categorisation of patients into deficit and non-deficit groups. In a later study, the diagnosis procedure was repeated with a previously categorised group, the authors (Amador et al., 1999) finding 83% agreement on designation of deficit status, and 88% agreement on the non-deficit categorisation.

The importance of distinguishing primary and secondary deficit symptoms lies in the implications for treatment outcome. The presence of primary deficits has been linked with greater periods of hospitalisation, a poorer employment record, impaired social functioning, greater severity of overall symptomatology, and lower scores on global outcome measures, the strongest association being with impaired social functioning (Fenton & McGlashan, 1994). Further, while primary deficits are enduring features of the disease process and relatively unresponsive to treatment, secondary deficits can be alleviated by modification of treatment regimes or provision of greater social stimulation.

In addition to observer-ratings of the permanency of negative symptoms, a wealth of evidence has been presented of impairment on common neuropsychological tests. It is now clear that a constellation of genuine deficits exists in psychosis. In schizophrenia, impairment has been found in long-term episodic memory (Stip & Lussier, 1996) and in semantic memory (Frith, 1992); a common theme in these studies is that access is impaired rather than the stores themselves. Contrary to these

assertions, evidence has been presented of impairment to memory stores (Pantelis et al., 1997), though this deficit was less debilitating than the impairment in processes controlling access to the contents of stores. Extensive evidence also exists of impairment in short-term memory processes.

Many recent investigations into neuropsychological impairment in schizophrenia have made use of the working memory model developed by Baddeley (1990). The proposed system is composed of three elements: the central executive, the phonological loop, and the visuo-spatial sketchpad. Shallice's Supervisory Attention System (SAS; Shallice, 1988) is cited as a model which performs the functions of the central executive and is consistent with existing evidence.

A deficit in executive function is well-established in schizophrenia, impairment being found on neuropsychological tests such as the Wisconsin Card Sort Test (WCST, Nelson, 1976; Goldberg et al., 1987), Stroop (Liddle & Morris, 1991), Tower of London (Pantelis et al., 1997), and Continuous Performance Task (Frith et al., 1991). Impairment has been demonstrated in both verbal (Fleming et al., 1995) and spatial (Fleming et al., 1997) short-term memory. Nathaniel-Jones et al. (1996) suggested that the apparent impairment in executive performance is simply the product of these sub-system impairments. Though evidence exists of impairment in both verbal and spatial skills, it has been suggested that relative impairment in verbal abilities may be characteristic of schizophrenic performance (Goldberg et al., 1993; Taylor et al., 1981).

Other authors have hypothesised that the deficit in schizophrenia is not specific to either the central executive or any of its subsystems (Pantelis et al., 1997; Salamé et al., 1998); they propose that impairment results from a generalised deficit in processing efficiency, termed bradyphrenia. This form of impairment has been frequently described in parkinsonism and represents a generalised slowing of information processing. Parallels are drawn with changes due to ageing, Brébion et al., (1998) suggesting that "processing speed may be the primary limit to cognitive performance in schizophrenia as it is in the elderly". However, other evidence indicates that not all cognitive processes are affected evenly, this being "counter to a hypothesis that the cognitive dysfunction in schizophrenia is due to a single, generalised deficit" (Schatz, 1998).

It should be noted that a pattern of differential deficits, performance being relatively more impaired in some domains than in others forms a strong argument that deficits are genuine and do not result from factors such as lack of motivation (Gruzelier et al., 1988). Further, it has also been confirmed that the neuropsychological deficits found in schizophrenia are independent of medication effects (Pantelis et al., 1997), and are present in first-episode non-medicated patients (Saykin et al., 1994). However, there is considerable variation among individual patients: on any given task, only about half perform in the subnormal range (Stip, 1996).

Evidence of similar impairment exists in the major affective disorders though the literature is far less comprehensive than that for schizophrenia. This is particularly true for bipolar disorder; probably due at least in part to the difficulties of testing

patients in a manic phase. However, there is evidence for impairment in visual-spatial tasks (Taylor et al., 1981) and in executive function (McGrath et al., 1997). Amongst depressed patients there is similar evidence for impairment in spatial tasks (Taylor et al., 1981) and executive function (Austin et al., 1999; Hart et al., 1998). As in schizophrenia, neuropsychological impairment within major affective disorder groups (bipolar and unipolar) is apparent in only a limited proportion of the group, other members of the group performing within the normal range (Goldberg et al., 1993).

The majority of papers published in this field indicate a pattern of relatively more severe impairment in verbal abilities in schizophrenia, and relatively more severe impairment in visual-spatial abilities in major affective disorder (Goldberg et al., 1993; Taylor et al., 1981), though some investigators did find evidence for the opposite patterns (Austin et al., 1999; Hart et al., 1998). However, it is worth noting that in almost all studies directly comparing schizophrenia patients with major affective disorder patients, the evidence suggests relatively greater overall impairment in the schizophrenia group (Goldberg et al., 1993; Taylor et al., 1981)

Though there are differences between the patterns of impairment found in schizophrenia and major affective disorder there are commonalities. Executive function is impaired in both schizophrenia (Frith et al., 1991; Goldberg et al., 1987; Liddle & Morris, 1991; Pantelis et al., 1997) and major affective disorder (McGrath et al., 1997), and significant psychomotor slowing may be present in both schizophrenia and depressive disorders (Purcell et al., 1997). However, even where

there are similarities in performance it should not be assumed that there are common mechanisms mediating impairment.

The difficulties inherent in studying these issues must not be underestimated. The existence of efficacious treatment for psychosis provides an ethical imperative to treat though the use of this treatment may (as discussed later) contaminate the assessment of the underlying disorder.

1.2 Antipsychotics

The introduction of effective antipsychotics in the 1950s revolutionised psychiatry. Other types of drugs have also had a great impact, for instance that of efficacious antidepressants. However, despite the fact that antipsychotics have promised so much their use remains plagued with difficulties. These difficulties, combined with the ubiquity of antipsychotics in psychiatric practice, make the continued study of antipsychotics both more relevant and more urgent than that of other drugs.

1.2.1 The development of antipsychotics

That there are continuing problems with antipsychotics stems in part from their empirical development. The first antipsychotic to become publicised was chlorpromazine which was brought to the world's attention by Henri Laborit, a naval surgeon. Laborit found that chlorpromazine, investigated as an agent to dampen autonomic activity during and after surgery, could induce affective and behavioural changes. Patients were described as being "calm and somnolent, with a relaxed and detached expression". This "twilight state" of complete equanimity was later termed "ataraxy", meaning "without anxiety". The state of ataraxy was contrasted with the effects of existing agents such as morphine or the barbiturates.

Recognition of this, at the time unique, action provided the impetus for the commercial development and production of chlorpromazine. Within two years, following the seminal paper of Delay and Deniker (1952), chlorpromazine had become widely used and was a commercial success. In the following years many

other antipsychotics were developed. At first these agents were highly derivative of chlorpromazine but later agents were less closely related. However, all shared the same goal of reducing the symptoms of psychosis without sedation.

Efficacy of treatment of psychosis can be evaluated in a number of different domains. The efficacy of antipsychotics is clearly established in terms of their action on positive features (Cole, 1964). Levels of arousal and anxiety are reduced and hallucinations and delusions are suppressed. This specific action is the basis for class membership, distinguishing antipsychotics from tranquillisers, benzodiazepines etc. However, it is widely stated that around 25-30% of patients do not respond well to antipsychotic treatment and the true figure may be higher than this (Kane, 1995). Despite this, the action is sufficiently specific to be used as a criterion for class membership. Treatment efficacy in other domains is suggestive rather than proven. In particular, evidence for efficacy in treatment of negative or deficit features is controversial.

Much recent research effort has been directed to discovering the mechanisms of antipsychotic efficacy. It is hoped that more efficacious and more tolerable drugs may arise from a greater understanding of the pathophysiological mechanisms of psychosis. Existing evidence from both laboratory (Creese et al., 1976) and clinical (Johnstone et al., 1978) settings suggests that the efficacy of antipsychotic agents such as chlorpromazine is related to central blockade of D_2 dopamine receptors. However, these drugs have actions at a huge variety of other receptor sites.

1.2.2 Extrapyramidal signs

Though alike pharmacologically, the different types of antipsychotics are very varied chemically. These variations affect factors such as the degree of specificity of action. This can in turn alter the tolerability of the agent. The most pervasive adverse effects associated with antipsychotics, extra-pyramidal signs (EPS), have, like the beneficial effects, been linked with central dopamine blockade. The propensity of antipsychotics to cause EPS was noted even at their introduction by Delay and Deniker (1952). However, this action was viewed as unimportant, and a formal report of the propensity of chlorpromazine to induce EPS was not produced until 1954 (Steck, 1954).

EPS comprise a number of different forms of movement disorder:

Parkinsonism	All major signs of parkinsonism (bradykinesia, rigidity, tremor) and possibly affective and cognitive changes in mental state.
Acute dystonia	Involuntary motor activity in which muscle action is sustained at a point of maximal contraction, frequently resulting in a twisting distortion of the affected part.
Akathisia	Visible signs of discomfort and unease, difficulty in sitting still. Subjective symptoms of inner restlessness, anxiety, and disquiet.
Tardive dyskinesia	Involuntary movements; often predominantly in orofacial regions but may be found in all body parts.

Over the period since EPS were first recognised, opinions on their importance have changed dramatically. At times the development of EPS in a patient has been used as a means of adjusting the dosage required, the presence of EPS indicating that the

dose was sufficient to ensure therapeutic efficacy (see Haase, 1985). More recently, and to the present day, EPS are viewed as undesirable side effects. However, the fact that these highly prevalent effects are an intrinsic component of the action of antipsychotics led to the adoption of the term "neuroleptic", or "that which grips the nerve", as a coverall term for typical antipsychotic agents. Had other suggestions been adopted these drugs may have been named after their beneficial qualities, as 'ataractics'. The relationship between the beneficial and adverse effects of these agents is only slightly clearer now.

These disorders are undoubtedly very common. Though a wide range of different figures have been presented for the prevalence of EPS in differently defined patient groups, most are in the range 40-85% of typical antipsychotic treated patients (Casey, 1989).

The different forms of EPS may also be classified by their relationship with antipsychotic drug treatment (after Owens, 1999).

Mode of onset / course	Acute		Chronic
Duration of exposure	Early	Intermediate	Late
Syndromal	Acute dystonias	Parkinsonism Akathisia (acute)	Tardive dyskinesia Tardive akathisia
Relationship to pharmacological intervention	Direct / intimate neurological responsivity		Indirect / delayed / paradoxical neurological responsivity

The EPS literature is to a large extent dominated by tardive dyskinesia (TD). A recent search of Medline found 478 papers concerned with tardive dyskinesia while only 94 concerned with drug-induced parkinsonism (DIP) were published within the same period of time. This may be due to medico-legal issues, particularly as much of the work originates from the US. Dystonias are much less studied than other forms of EPS. This is perhaps due to a widespread belief that they are less common than other forms of EPS though this may not be so when high potency antipsychotics are used. Akathisia is perhaps least often the subject of systematic study. This neglect is possibly due to the difficulties of distinguishing akathisia from psychomotor agitation occurring as a feature of psychosis. Despite being the first form of EPS to be identified, DIP has been less studied in recent years. Since its first recognition, DIP has at different times been viewed as an inevitable consequence of antipsychotic medication, being a marker of treatment efficacy, or an unfortunate but treatable side effect unrelated to efficacy.

The concept of a threshold dose for the development of EPS was noted above. Proponents of this hypothesis stated that a threshold, particular to each individual patient, existed. Below this threshold lay therapeutic efficacy, above it toxicity, presenting usually as parkinsonism. Haase proposed that an optimal dose of antipsychotic for the individual could be determined by increasing dosage slightly until the first indications of parkinsonism were apparent (the development of micrographia was to be used as an indicator) and then reducing it slightly.

Recent use of brain imaging techniques has revived threshold theories of antipsychotic efficacy. Studies have indicated the existence of a threshold of D2 receptor occupancy in the striatum, above which EPS is apparent (Farde et al., 1992; Scherer et al., 1994). This implies the presence of a narrow dosage band of therapeutic efficacy. Below this band dosage is insufficient for treatment to be effective, above it dosage is sufficient to cause EPS. However, these accounts are yet to trigger any wholesale revolution in psychiatry.

Though often dismissed as merely troublesome side effects, the boundaries between EPS and features of psychosis form a major issue in modern psychiatry. In particular, the relationship between parkinsonism and deficit features of psychosis bears on issues of diagnosis, medication, drug efficacy, and treatment outcome.

1.2.3 Atypical antipsychotics

Following the evermore widespread adoption of chlorpromazine, other antipsychotic agents were introduced, all of them derivatives of chlorpromazine, and with similar modes of action. Other types of antipsychotics were developed too, differing to greater or lesser extent in properties and actions.

Increasing awareness of the prevalence of EPS drew attention to the lack of tolerability displayed by all extant antipsychotics. The next wave of drug development was theory-driven, in contrast to the empirical development of chlorpromazine. The mesolimbic dopamine system had been identified as the site of schizophrenic pathophysiology, and the nigrostriatal dopamine system implicated in

the production of EPS. Dopamine receptors had been recognised as falling into D_1 and D_2 groups (since further sub-divided into various D_1-like and D_2-like types), D_2 receptors being more common in the mesolimbic dopamine system and D_1 receptors being more common in the nigrostriatal system. Thus, drugs were developed with the attention of focussing on the mesolimbic system, acting more selectively at D_2 receptor sites. Sulpiride, a substituted benzamide, and the first antipsychotic to be labelled "atypical," was one of these.

In this context, "atypical" refers to antipsychotic agents which differ from "typical" antipsychotics in their having lowered propensity to induce EPS in the presence of equivalent therapeutic efficacy. Preliminary examination of sulpiride led to the conclusion that it did indeed have a lower propensity to induce EPS. However, this optimism was short-lived. Further studies were conducted in which care was taken to ensure that the sulpiride group received doses of equivalent therapeutic strength to those received by the comparison group (receiving a typical antipsychotic). The results of these investigations indicated that the advantages of sulpiride were less striking than previously thought.

Single system pharmacology is not always viewed as the solution to the problem of neurological side effects. Less selective drugs, previously derided as "dirty," may provide more of the benefits promised by so-called "clean" agents. Clozapine in particular has radically changed perceptions of schizophrenic psychopharmacology. A dibenzodiazepine, clozapine was developed as an antidepressant and was first registered in 1960. However, concern grew over its adverse effects on granulocytes

and it was withdrawn in 1975 following a number of deaths. It has since been re-introduced for use in schizophrenia though regular blood testing remains a condition of its use.

The evidence suggests that clozapine is truly different from typical antipsychotics. In a large multi-centre trial clozapine was compared with chlorpromazine with prophylactic antiparkinsonian medication in cases of treatment-resistant schizophrenia. The results demonstrated a clear advantage for clozapine in terms of reduction of positive symptomatology and neurologic tolerability. Further studies have since confirmed that clozapine is not only superior to typical antipsychotics in selected groups but at least as effective as other agents in non-selected groups of schizophrenics.

Equally extensive evidence has accumulated of a strikingly low liability to cause EPS. Clozapine causes little or no dystonia and the respective incidences of akathisia and DIP are greatly reduced. The risk of tardive dyskinesia is probably also very low though the evidence for this is less clear.

Clozapine is a drug apart from typical antipsychotics in pharmacological terms too. It exhibits only low occupancy rates of D_2 receptors and its range of actions is broader even than typical antipsychotics. Affinity for serotonergic receptors is particularly high. More recent conceptions of dopamine neurophysiology (Jaber et al., 1996) sub-divide D_1 and D_2 receptor types into D_1-like and D_2-like sub-types. To the extent that clozapine does act at DA receptor sites, its actions may be at D_4 receptors (D_2-like)

which are thought to be localised to the cerebral cortex. However, it is the broad spectrum of action of clozapine which is assumed to be key to its therapeutic advantages though the mechanism by which it achieves these benefits is as yet unclear.

Since the successful re-introduction of clozapine, efforts have been directed at developing other atypical antipsychotics which may achieve the therapeutic efficacy and low level of neurological side effects of clozapine without the increased risk of other adverse effects. To this end a number of other atypical antipsychotics have been introduced, all of which to some extent achieve their aims. However, many authors believe that clozapine's propensity to cause neurological side effects is still uniquely low (Miller et al., 1998).

1.2.4 Subjective experience of antipsychotics

The promise of atypical antipsychotics is of efficacious treatment of positive symptomatology free from the adverse neurological effects associated with typical antipsychotics. This freedom may extend beyond the overt physical signs of EPS to include the negative subjective experiences often associated with typical antipsychotic medication. These experiences are far from uncommon and may play a major role in treatment success.

Non-compliance with medication regimes is very high in clinical practice, particularly amongst outpatients. Assessments of clinically significant non-compliance in inpatients range from 7%-57% (Weiden et al., 1991), and up to 73%

amongst outpatients over a 2-year follow-up (Serban & Thomas, 1974). Earlier accounts of non-compliance most commonly attributed it, if not to factors associated with the illness, then to physical EPS (Van Putten, 1974). More recently the same author considered that this relationship might be mediated by the subjective components of EPS (Van Putten & May,1978). Other workers have confirmed this relationship between negative subjective experiences and non-compliance (Hogan & Awad, 1992, Awad & Hogan, 1994).

Descriptions of these negative subjective experiences frequently centre around complaints of feeling "fuzzy, woolly, lacking energy, unable to think clearly, like a zombie, restless, etc." A number of different terms have been coined for complaints of restricted cognition and emotion resulting from antipsychotic medication: "akinetic depression" (Rifkin et al., 1975; Van Putten & May, 1978), "neuroleptic dysphoria" (Hogan & Awad, 1992, Awad & Hogan, 1994), "neuroleptic-induced anhedonia" (Wise, 1991). Though these experiences are often subsumed under the catch-all term "dysphoria", a closer examination of the descriptors used indicates that the sensations are not common to all patients and that there may be different facets to the experience.

Some authors argue that the negative experiences may actually constitute depression. "Pharmacogenic depression" was noted in the German literature during the 1960s (Bandelow et al., 1992). Van Putten and May (1978) termed their conception of antipsychotic-associated dysphoria "akinetic depression", making an explicit link with DIP. However, it has also been argued that depression seen in schizophrenic

patients may not be linked with medication. Symptoms of depression may be accounted for in terms of "reactive depression" to the negative life events of being diagnosed with a serious mental illness, hospitalisation etc. This account has been dismissed (Knights & Hirsch, 1981) as "intellectually weak". These authors also discounted the account of Van Putten and May (1978), having found that depression persisted after treatment of even very mild signs of EPS. Knights and Hirsch argued that depression was an intrinsic part of the disease process of schizophrenia. More recently however, Bandelow et al. (1992) found higher levels of depression in patients treated with antipsychotics than in non-treated patients, and evidence of an association between EPS and depression. In light of all these findings it seems possible that depression may occur both as a component of psychosis and as a consequence of antipsychotic medication.

Reports of the subjective experience of antipsychotics in normals (Belmaker and Wald, 1977) are consistent with those of patients. Belmaker and Wald reported sensations of inner restlessness, anxiety, inability to relax, poor concentration, and irritability, coincident with a "paralysis of volition" and a lack of physical and psychic energy. They described a feeling that they felt unable to initiate tasks though they could perform them if demanded to do so. A 1992 review of literature concerned with the effects of typical antipsychotics in normals (Hollister, 1992) found evidence that chlorpromazine and reserpine had been associated with complaints of restlessness, depression, and feelings of unreality and depersonalisation. Similar complaints of nerves and apprehension were reported after a double-blind trial of reserpine.

The different, and sometimes apparently paradoxical, facets of the negative subjective experience may reflect different features of physical EPS. Particular negative experiences have been found to be reliably associated with particular forms of physical EPS (see table below). In fact, negative subjective response to antipsychotics early in treatment is predictive of physical forms of EPS later in treatment (Hogan & Awad, 1992).

Form of EPS	**Characteristics of associated subjective experience**
Dystonia	Fear and anxiety (Casey, 1994).
Akathisia	Inner restlessness, anxiety. Poor concentration and irritability (Casey, 1994).
Parkinsonism (particularly bradykinesia)	Lack of physical and psychic energy. Feeling "mummified and dull" (Van Putten and May, 1978), or "like a zombie" (Awad, 1993).

The efficacy of atypical antipsychotics has also been assessed in terms of quality of life. Quality of life measures are widespread in other areas of medicine but are underused in psychiatry. Awad and Hogan (1994) argue that quality of life provides a framework in which many aspects of medication response may be considered. Relief from schizophrenic symptomatology, side effects, psychosocial factors.

Where quality of life measures have been used they have indicated clear benefits of atypical antipsychotics over typical drugs. Meltzer (Meltzer et al., 1990) reported the cases of 38 treatment-resistant schizophrenic patients who were switched to

clozapine treatment. It was found that significant improvements on quality of life measures (Quality of Life scale; Heinrichs et al., 1984) were apparent after six months of clozapine treatment. After 12 months of treatment, continued improvements were seen on this index (Meltzer, 1992). Meltzer stated that these improvements reflected "highly significant clinical changes that are rarely, if ever, seen after switching typical neuroleptic drugs in patients who are poor responders to three or more other typical neuroleptics and in the relatively older schizophrenics studied here."

Naber (1995) reported the use of a measure of Subjective Well-being under Neuroleptics (SWN; Naber et al., 1994). This measure is intended to be specific to the negative effects of typical antipsychotics on quality of life. Naber states that significant correlations are found between results of this scale and other measures of quality of life. Scores on this scale were found to be significantly higher in a clozapine treated group than in a group receiving typical antipsychotics (haloperidol and flupenthixol), despite the fact that the clozapine group had been negatively selected for this medication due to therapy resistance or major side effects with typical antipsychotics.

Despite the influence that negative subjective experiences of antipsychotics may have on treatment success, they are often ignored. The assessment and identification of these phenomena is a major obstacle, and misdiagnosis is common. The superficial similarities between the dysphoria induced by medication and features of the illness being treated are considerable. There may be significant

phenomenological overlap between a number of constructs: deficit features of schizophrenia, psychomotor slowing as a feature of depression, antipsychotic-induced deficits, parkinsonism. Efforts to disentangle these issues require a full understanding not only of the underlying psychoses, but of the disorders which may result from medication.

1.3 Parkinsonism

1.3.1 Features of parkinsonism

This section will cover the features of parkinsonism, noting their properties and the mechanisms proposed to underlie them. Much of the evidence that will be presented here derives from studies of Parkinson's disease. Though it would be preferable to rely only upon evidence derived from DIP, there is a scarcity of work in this category. Caution must be taken in inferring from one form of parkinsonism to another.

1.3.1.1 Bradykinesia

Bradykinesia is, in most cases, the most salient feature of parkinsonism, particularly upper body bradykinesia (Quinn, 1995). Taken literally, bradykinesia means simply "slowed movement", however it is much more than this. Manifestations of bradykinesia include diminution or poverty of background motor activity, slowed execution of movements, difficulties in initiation, increased fatigability, diminishing amplitude of repetitive movements, impairment in sequencing of movements. This symptom complex is difficult to describe and define, its expression varying not only from patient to patient but from day to day within the same patient.

Different authors have used varying terms to refer to bradykinesia. 'Akinesia' is common, particularly amongst neurologists. However, strictly used the prefix 'a-' must refer to a total lack of movement rather than the more moderate abnormality of most cases of parkinsonism. 'Hypokinesia' might appear more appropriate though

while it has been used by some authors (e.g. Bloxham et al., 1984) for whatever reason it is not favoured. It has been suggested that 'akinesia' be used to refer to a poverty of movements produced and 'bradykinesia' be used to refer solely to a slowing in the execution of movement (Delwaide and Gonce, 1988). Though there is some evidence to support a distinction of this nature (1.3.1.1.1) this proposal has not been widely adopted.

1.3.1.1.1 Manifestations of bradykinesia

The most common major manifestations of bradykinesia were noted above. In this section these manifestations will be more comprehensively described.

Diminution or poverty of background motor activity presents as a lack of normal non-purposive movements. Adjustments of posture, the continuous non-goal-directed background motor activity seen in normals, are absent. The patient with parkinsonism may sit almost immobile. When movements are made the execution is usually slowed to at least some degree. Though actions may be performed competently and even accurately, they occur at a slower speed than normal. This has been demonstrated in tests both of pure movement speed (Evarts et al., 1981) and in tasks of motor control (Meier & Martin, 1970).

Difficulties in initiating movements are common in parkinsonism. The patient may have an action in mind to perform but is unable to do so. Vaughan (1986) a Parkinson's disease patient describes both the extent of his disability in normal activities of daily living (due to an inability to initiate movements) and his ability to

run for miles once started. Frequently an external stimulus may be necessary to trigger movements. Sacks (1973) describes devices which have proven effective as aids to initiate movement: lines on the floor to step over, tiny balls of screwed-up paper which can be dropped with a minimum of movement to then trigger a much greater movement. This reliance upon external stimuli to trigger movements, and the associated lack of self-initiated movements has led to parkinsonian patients being described as 'environmentally-driven' (Sacks, 1973). However, this term implies not only an inability to initiate actions held in mind, but a failure to plan self-driven, 'willed' movements. The extent to which this is true will be addressed later (1.3.1.5).

One of the first indications of incipient parkinsonism is often a report of abnormal tiredness and lassitude. The patient may report that they are easily fatigued and become rapidly tired. When investigated experimentally, strength may be normal at first yet rapidly tail off (Onuaguluchi, 1964). Over the course of a time period in which normal performance remains constant, performance levels amongst parkinsonian patients decay rapidly. This inability to maintain strength over relatively short periods of time may play a role in the manner in which the amplitude of repetitive movements can be seen to diminish.

This particular manifestation of bradykinesia is commonly used as an assessment technique in clinical practice. The patient is asked to hold their hands outstretched horizontally in front of their body, and then to repeatedly pronate and supinate the hands, turning them from palm-down to palm-up. Normals asked to perform this task do so until instructed to stop. Parkinsonian patients may complete only a few cycles

before the amplitude of the movements has decreased to nothing and the hands are still.

Few patients display all of the manifestations of bradykinesia described. Bradykinesia is a symptom complex that resists precise description and definition. Not only do some patients not exhibit some of the manifestations, those they do exhibit may only be apparent in some situations. A comprehensive account of bradykinesia must describe the situations in which impairment is found and not found. It must note where impairments thought to be associated can be proven to be dissociable. And it must attempt to clarify apparently contradictory findings.

Impairments in movement initiation and speed have been long demonstrated, most often by the use of reaction time tests (Wilson, 1925). Though these two impairments do show an association, evidence exists that they can occur independently. Later workers have used a reaction time paradigm which allows the separation of the latency before the response is initiated from the time taken to complete the response movement (Evarts et al., 1981). Evarts found that in some trials delayed initiation was followed by normal speed of movement, in others a normal time reaction time to initiate the movement was followed by an abnormally slow movement time. The different patterns could be seen in different patients or in the same patient in opposite arms. This pattern of performance can only be found if the two impairments, a deficit in initiating movements and a decrease in speed of execution, are mediated by different mechanisms.

The demonstrated dissociation between impairments in movement initiation and speed of execution supports the differential use of the terms 'bradykinesia' and 'akinesia' noted earlier (Delwaide & Gonce, 1988). In fact the authors of the work (Evarts et al., 1981) use the term bradykinesia to refer solely to the impairment in speed of execution and not to the impairment in initiation (as indicated by slowed reaction time). However, as previously stated this practice has not been more widely adopted.

The above dissociation between performance on measures of movement initiation and speed of execution is far from the whole story. In a second experiment by the same authors, performances in simple reaction time (RT) and choice reaction time (CRT) conditions were compared. In a simple RT condition, the same stimulus is presented on every trial; the required response is constant too. In the CRT condition, the stimulus differs from trial to trial; for each stimulus a different response must be made. It was found that parkinsonian patients were almost unimpaired on CRT though they had shown a significant deficit on tests of RT (Evarts et al., 1981).

Evarts et al. note that this is opposite to the expected pattern of results, formed on the basis of similarities between ageing and parkinsonism. Parallels between the effects of parkinsonism and of ageing are not uncommon (Dobbs et al., 1992). However, this finding is directly opposite to that found in elderly subjects in whom CRT performance is relatively more impaired than simple RT performance. In elderly subjects this is due an increase in initiation time in CRT, attributed to slowed information processing. Unfortunately the published results of Evarts et al. do not

allow the contribution of initiation and movement times to reaction time performance to be investigated in this experiment.

If it were assumed that movement speed was similar in the two conditions (it is certainly unlikely that movement is quicker in the CRT condition), the results appear to imply a much quicker speed of information-processing in the CRT condition in the parkinsonian group than in the control group (Bloxham et al., 1984). Given the unlikelihood of this explanation, one must conclude that the parkinsonian group are unimpaired (or relatively so) on CRT but are impaired in the simple RT condition. Bloxham et al. argued that normals respond more quickly in the RT condition than CRT because they know what movement will be needed and can pre-program it. Patients with Parkinson's disease, in contrast, do not make use of this information and select their response only as the target is seen. Thus, the initiation time deficit in the simple RT condition is relatively much greater than that in the CRT condition.

An earlier work using a different paradigm (a target-tracking task) found similar evidence of an inability to make use of prior information in planning movements (Flowers, 1978). It was demonstrated that parkinsonian patients performed poorly on a manual tracking task in which the target followed a repetitive (and thus predictable) path in one dimension. In certain trials, the target changed direction while concealed; controls could nevertheless predict the target's movement and continue to track it accurately but the patient group was disadvantaged further. Flowers suggested they were, "tied more directly to current sensory information, responding to events rather than anticipating them."

These findings too have since been clarified further (Bloxham et al., 1984), elucidating more accurately the situations in which performance is impaired. In the study of Bloxham et al., patients performed well on a tracking task (even at speeds too quick for use of visual feedback). This was true for both predictable (normals known to pre-program movements) and unpredictable tracks (both groups rely on feedback so prediction and pre-programming are not factors). In order to accommodate the results of Flowers, Bloxham et al. discussed the concept of "segmenting" in control of action. In this account, a new unit of movement is initiated (consciously) at each segmentation point; once initiated, control is automatic and ballistic. This study differs from that of Flowers in that the movement required is circular rather than a one-dimensional sweep. Therefore there is no obvious segmentation. When a new unit of movement must be initiated at the end of each sweep an impairment is found but a circular task is unaffected by initiation problems. According to Bloxham et al., patients are able to use prior information to control the form of a movement but not to initiate or pre-select it (particularly without an external trigger). A delay in initiation may still occur, but the hypothesis has been refined to explain a low efficiency of response to prior warnings (i.e. a failure to pre-select motor programs).

Perhaps the most notable findings are not the situations in which performance is impaired but the number of situations in which performance is normal. These findings have a bearing on determining the mechanisms responsible for bradykinesia.

1.3.1.1.2 Pathophysiology of bradykinesia

An account of the pathophysiology underlying bradykinesia must explain not only the performance impairments exhibited by patients but also the situations in which performance is normal.

That performance can, in some circumstances be normal argues against peripheral causes, such as factors directly affecting the muscles. It has been demonstrated that muscle innervation is not dysfunctional in parkinsonism (Dietz et al., 1981). In a study of gait in parkinsonism, using EMG measures of muscular innervation, it was found that the muscles are provided with the stimulation to perform actions planned at a higher level of the motor programming hierarchy. Thus impairment in performance must result from impairment in central motor control and not peripheral factors.

It has been stated that evidence for the location of the pathophysiology lacks detail, though bradykinesia correlates well with striatal dopamine deficit (Delwaide and Gonce, 1988). However, Marsden (1982) states that dysfunction may be traced to the basal ganglia, and suggests that Parkinson's disease provides a model of basal ganglia dysfunction.

1.3.1.2 Rigidity

Hypertonia, or rigidity, is found not only in parkinsonism but as a component of other disorders too. Three major forms of rigidity have been described: lead-pipe, clasp-knife, and cogwheel. These are distinguished in clinical practice by manual

palpation of the limb. Lead-pipe rigidity is felt as a uniform stiffness as the joint is moved passively. Where rigidity is of the clasp-knife type the limb remains immobile as increasing force is applied until it relaxes abruptly and offers no resistance to flexion. Cogwheel rigidity is a form in which passive extension or flexion is felt by the assessor to be occurring in a series of steps, though a constant force is applied. The frequency of these steps varies between patients; at higher frequencies it may be felt as "rippling" (Findley et al., 1981).

1.3.1.2.1 Manifestations of rigidity

Two forms of rigidity are seen in parkinsonism, lead-pipe and cogwheel rigidity. In addition, the neutral angle at the elbow has been found to be significantly decreased (Watts et al., 1986). Data for the prevalence of rigidity in parkinsonism are hard to come by though it is present in most cases of Parkinson's disease and is slightly less common in DIP. Caligiuri et al. (1989) found clinically apparent rigidity in 42% of antipsychotic treated patients.

Lead-pipe rigidity is, as described above, a simple increase in resting muscle tone. Cogwheel rigidity is more complex. The frequency of the "cogs" has been found to be between 6-6.6Hz and 7.5-9Hz (Findley et al., 1981), at the same frequency as postural tremor (in the respective patients). Of their 40 patients, 15 (38%) exhibited cog-wheeling in the 6-6.5Hz range and 18 (45%) exhibited cog-wheeling in the 7.5-9Hz range.

1.3.1.2.2 Pathophysiology of rigidity

As with tremor, two major primary causes of extrapyramidal rigidity have been proposed (Delwaide and Gonce, 1988). Firstly, mediation via spinal mechanisms has been proposed, citing "discrete troubles found in a few spinal cord reflex pathways", and a possible role for the tonic stretch reflex which correlates well with rigidity. Secondly, Delwaide and Gonce describe a possible mechanism involving hyperactivity in long-loop reflex pathways. However, Watts et al. (1986) found that rigidity was still present when EMG measures demonstrated that there was no muscle activity. Rigidity even persists after preparation for surgery when total muscle relaxants have been administered (Walsh, 1992). It is suggested that this occurs as a consequence of inactivity due to akinesia (Watts et al., 1986) but Delwaide and Gonce state that it is unlikely that this mechanism alone can explain rigidity.

Even less evidence exists to explain the production of cog wheeling. Some authors have suggested that rigidity overlaid with tremor may be responsible (Lance et al., 1963; Findley et al., 1981). These assertions were made primarily on the basis of similarities in frequency of tremor and cog wheeling. However, this hypothesis may well be an oversimplification (Owens, 1999),

In general, the literature concerning the mechanisms underlying rigidity is far from extensive; rigidity is often considered solely as a diagnostic marker for parkinsonism. Delwaide and Gonce (1988) concluded that "a firm conclusion on the mechanisms responsible for rigidity seems premature".

1.3.1.3 Tremor

Tremor may be defined as the regular movement of a body part about a fixed point. Movement occurs in more than one dimension though this may not be apparent to the naked eye. A degree of tremor, physiologic tremor, is normal and is exhibited by all humans. Symptomatic tremor may be exhibited as essential tremor, as a feature of parkinsonism, or in a number of other disorders. Within Parkinson's disease the prevalence of tremor is high; prevalence of tremor at presentation has been put at 70% (Quinn, 1995). However, tremor is held to be less prevalent in DIP (Ayd, 1961).

Categorisation of tremor is far from simple. It is often described simply in terms of context: whether it is present at rest, in posture, when the intention to move is formed, or during the actual performance of the action. Otherwise, it may be described as being slow or fast, coarse or fine, or of large or small amplitude. Though it is often stated that large tremors are of low frequency, and small tremors are of high frequency (Owens, 1999) this is a relationship of association rather than causation.

The characteristic parkinsonian tremor, as seen in Parkinson's disease, is a slow resting tremor. The combination of slow tremor with finger flexion gives this tremor a "pill-rolling" appearance.

1.3.1.3.1 Characterisation of tremor

Though tremulous movement occurs along more than one dimension, the movement is predominantly along the vertical axis when the hand is held pronated. For the purposes of measurement or description the tremor is most commonly considered as if movement occurs along only one axis (Caligiuri et al., 1989b; Arblaster et al., 1993). Along this axis, the tremulous movement occurs as a waveform.

Waveforms may be described in terms of their frequency and amplitude. The frequency represents the rate at which the limb is oscillating. This is expressed in terms of cycles per second (Hz). The amplitude of the tremor represents the distance that the tremulous limb moves from the fixed point in space about which it is moving. An alternative indicator of amplitude is acceleration. This method relies upon the fact that the acceleration that the limb is undergoing at any point in the cycle is closely related to its displacement from the neutral fixed point. The advantage of this is principally in terms of instrumentation factors (many studies use a means of instrumentation that measures acceleration).

So far so simple. However, it is rare that a tremor forms a regular wave. In almost all cases, it is a very irregular waveform comprising activity at a number of different frequencies. Mathematical techniques (FFT; 1.3.2.3) may be used to break down a sample of an irregular wave into its component parts. In many studies, the peak frequency of tremor is used as a comparator. The mean peak frequency for one group may be compared with that for another group (Tyrer et al., 1981), or for the same group after changes in medication (Pullinger & Tyrer, 1983). Alternatively, comparisons may be made using the number of cases within a group whose peak

frequency falls below some threshold (Arblaster et al., 1993). Another method calculates the amplitude of tremor within frequency bands (Caligiuri et al., 1989b; Caligiuri et al., 1991; Caligiuri & Lohr, 1993). The amplitude of tremor present within the frequency bands 3-7 Hz (Caligiuri & Lohr, 1993) or 4-6 Hz (Caligiuri et al., 1991) has been suggested as an index of the severity of parkinsonian tremor.

1.3.1.3.2 Physiologic tremor

The characteristics of tremor need to be considered for both normal and abnormal tremors. Normal physiological tremor is of relatively small amplitude and of individually characteristic frequency though this can be affected by many factors (stress, food and drink, drugs etc.) Marsden (Marsden et al., 1969) stated that 95% of normal adults show a single dominant frequency and that the pattern from each person shows a 'signature' which they found to be constant over three years.

Within a larger time-scale, tremor characteristics are known to change with age. Specifically, the dominant tremor frequency is known to decrease (or the proportion of tremor in lower frequency bands increases). Marsden (Marsden et al., 1969) also found a decrease in the dominant frequency with age from 9Hz at ages 20-40 years old to 7.7Hz for a group over 60 years old. Similar figures (7Hz for a group under 70; 6Hz for a group over 70) were also found by Wade (Wade et al., 1982). The differences in the figures found by these two studies can most likely be attributed to differences in the manner in which tremor was assessed. However, the decrease in dominant frequency is of similar magnitude.

1.3.1.3.3 Parkinsonian tremor

The dominant frequencies to be found in Parkinson's disease tremor have also been investigated. Findley et al. (1981) found a slower (resting) tremor with a frequency range of 4Hz to 5.3Hz and a fine postural tremor with a range of 5.8Hz to 6.8Hz. However, the 'resting' tremor could "continue in posture" and the 'postural tremor' was "sometimes visible at rest".

Tremor in DIP has been comparatively well investigated relative to the other features of parkinsonism. Arblaster et al. (1993) compared tremor frequencies in antipsychotic-treated patients with normal controls. Very few controls (3.2%) showed a dominant frequency below 7Hz but 29% of the patients did in (at least) one arm. The figure for the patient group could be much higher if a less inclusive entry condition was used; the group was defined as those who had taken anti-psychotic drugs for at least one month in the previous year and while DIP is believed to continue for a period of time after the discontinuation of drug treatment this period may be less than a year. Despite this, the results demonstrated a lower frequency of tremor in the patient group that the authors attributed to DIP.

Caligiuri et al. (1991) found that the percentage of overall tremor activity occurring within the 4-6Hz frequency band is a valid indicator of parkinsonian tremor activity as assessed by observer ratings. Their results indicated that the tremor found in a patient group treated with typical antipsychotic medication was more parkinsonian than a control group but less parkinsonian than a group of Parkinson's disease patients.

1.3.1.3.4 Lithium-associated tremor

If the effects on tremor characteristics of treatment with antipsychotic agents are to be considered, the possible actions of other agents on tremor must be taken into consideration when selecting patients. In particular, lithium is known to be associated with tremor.

That a fine tremor may result from lithium therapy is long known (Schou, 1959). It is a postural tremor, which may be an exaggerated physiologic tremor occurring at around 8-12 Hz (Hallett, 1986). Lithium tremor can be distinguished from cerebellar tremors, and from parkinsonian resting tremor. It is not responsive to antiparkinsonian drug treatment (Schou et al., 1970; Tyrer et al., 1980). Rather, it responds to beta-blocking medication, consistent with an attribution to an adrenergic mechanism. The incidence of symptomatic tremor in patients receiving lithium therapy has been found to range from 4% to 65% (Gelenberg & Jefferson, 1995).

It is possible to distinguish between the characteristics of tremor in acute lithium therapy and those after chronic lithium treatment (duration of at least six months; Pullinger & Tyrer, 1983; Tyrer et al., 1981). Acute lithium tremor occurs at rest and in posture. Relative to tremor characteristics before treatment commenced there is some increase in amplitude though no change in peak frequency (Pullinger & Tyrer, 1983). After longer-term lithium therapy there is a greater increase in amplitude and also a (small) decrease in peak frequency towards the parkinsonian range (Tyrer et al., 1980).

1.3.1.3.5 Physiology and pathophysiology of tremor

The mechanism that mediates parkinsonian tremor is still unclear and there is little consensus. In fact, there is not full agreement on the mechanism underlying normal physiological tremor.

Authors	**Hypothesised mechanism**
Findley et al., 1981	Gamma efferents and alpha motor-neurones have both been implicated
Lakie, Walsh & Wright, 1986	In part due to the mechanical properties of the postural system
Marsden et al., 1969	No mechanism proposed
Wade et al., 1982	Components from neuromuscular activity, cardio-ballistic thrust and passive resonance in the tissues of the hand

Though no mechanism was proposed, the results of Marsden et al. (1969), taken from both arms simultaneously, showed that while the shape of the frequency spectra from the two arms was very similar there was little coherence and no phase relation. This indicates that the arms do not share a common source of activation, although the many factors which affect the arms equally (muscle changes, hormones and other blood-borne agents, temperature, fatigue) do ensure great similarity in the patterns. Though the systems causing tremor in the two arms are separate they share the same environment.

Similarly there is a paucity of evidence concerned with the production of parkinsonian tremor. Following from suggestions of similarities between the effects

of ageing and parkinsonism, the mechanisms hypothesised to account for age-related decreases in tremor frequency may be considered. Arblaster et al. (1993) suggest the decrease is due to age-related changes in the basal ganglia, specifically decreases in the dopamine levels, paralleling those seen in parkinsonism. These decreases are not, in most people, sufficient to lead to parkinsonism but may be sufficient to lower the dominant frequency of tremor (dopamine levels may decrease to 20% of normal before parkinsonism is apparent clinically). However, Wade et al. (1982) attributed the age-associated changes in tremor frequency which they found (7Hz for a group under 70; 6Hz for a group over 70) to changes in the natural resonant frequency of the tissues as there was no change in amplitude or spectral pattern. This is consistent with the attribution of tremor to mechanical properties of the postural system.

Two major causes of parkinsonian tremor have been proposed (Delwaide and Gonce, 1988). The first of these is a supraspinal mechanism in which rhythmic activity in the thalamus drives contralateral limb tremor. The second is a spinal mechanism with a role for oscillatory properties of the myotatic arc. Delwaide and Gonce describe a dual mechanism that uses the concept of long loop reflex pathways to integrate peripheral influences with a thalamic determination of tremor frequency.

Within the specific context of DIP, an attribution of tremor to extrapyramidal dopamine-blockade is invariably deemed sufficient (e.g. Caligiuri & Lohr, 1993). A similar situation is found in the case of lithium tremor. It has been found to be non-responsive to antiparkinsonian medication (Schou, 1970; Tyrer et al., 1980), though Tyrer (Tyrer et al., 1981) suggests that it is extra-pyramidal in nature.

1.3.1.4 Other physical features

Mild postural instability is a common feature of parkinsonism and is commonly used as a diagnostic technique. However, in some cases, postural instability can be severe. In these cases there can be an almost total lack of normal reactions to tilting. The patient fails to react appropriately if tilted and may make little or no effort to prevent themselves from over-balancing. This impairment is attributed to a failure to make use of information from vestibular function (Purdon Martin, 1967), though the vestibular system itself is unimpaired.

The characteristic gait of Parkinson's disease is a stooped, kyphotic, festinating shuffle. Step length is greatly shortened (Kirollos et al., 1993), stepping occurs more quickly, and double support time (a measurement of the length of the period during which both feet are on the ground) is increased. Pendular arm swing is usually absent. Though a stooped bent posture (a "triple flexion") is common in Parkinson's disease, an upright poker-back posture with a marching gait may also be seen, particularly in DIP.

In a study of 130 cases of post-encephalitic parkinsonism, Purdon Martin found evidence for gait abnormalities in all patients. Within DIP, the most common feature is the lack of pendular arm swing which may be one of the most sensitive indicators to incipient parkinsonism (Owens, 1999). It has been suggested that many of the characteristic abnormalities of parkinsonian gait are secondary to the previously noted postural instability (Purdon Martin, 1967) .In particular, a failure to control tilt

the body sufficiently to produce an adequate step length may lie at the root of festination. If the body leans forward at an angle appropriate for a normal pace of locomotion yet step length is insufficient, the rate of stepping must increase to prevent the patient falling forwards. Purdon Martin localised the source of the impairments in postural stability and gait to the basal ganglia.

Facial masking is common in all forms of parkinsonism. Normal facial expressions are absent not through a lack of emotional experience but a bradykinesia in the facial musculature. However, this feature can be easily confused with the affective flattening of psychosis in which the physical ability to express is intact though the emotional range itself is restricted.

Autonomic disturbances including seborrhoea and sialhorrea are found in advanced cases of parkinsonism though there have been no studies of their prevalence in DIP.

1.3.1.5 Cognitive deficits

Considerable evidence has accumulated for the existence of cognitive deficits in parkinsonism. These deficits are present in almost all patients and in Parkinson's disease they are to be distinguished from dementia. Dementia in Parkinson's disease is much less common than often suggested with a prevalence of 10-15% (Brown & Marsden, 1984). Although the deficits in non-demented cases are more subtle, there is evidence that the cognitive decline is progressive. Areas of impairment include short-term memory and executive function, long-term memory, visuospatial processing, and sensorimotor dysfunction (Brown & Marsden, 1990).

As with studies of cognitive impairment in psychosis, much use is made of Baddeley's model of working memory (Baddeley, 1990; 1.1.3). Impairment has been demonstrated in all components of working memory, but dysfunction in the central executive is particularly well established. Impairment has been demonstrated on a number of different measures of executive function: Wisconsin Card Sort Test (Gauntlett-Gilbert et al., 1999), Stroop Colour Word Test (Lund Johansen et al., 1996), Continuous Performance Task (Hart et al., 1998).

Dalrymple-Alford et al., (1994) presented evidence that Parkinson's disease patients were not impaired in the performance of two tasks separately but were less able than controls to perform the tasks simultaneously. The authors state that parkinsonian patients are impaired only when tasks are demanding and effortful, and when they must rely on internally generated cues to guide attention and behaviour; impairment is not found when a task requires only automatic responses. They argue that only the central executive component of Baddeley's model is dysfunctional.

However deficits have also been found in visuo-spatial working memory and verbal working memory. Owen et al. (1997), found evidence of a systematic decline in which, "working memory deficits emerge and subsequently progress, according to a defined sequence." In newly diagnosed Parkinson's disease patients, impairment was found only in executive function but in more advanced cases, impairment was found in spatial working memory tasks and finally in visual and verbal working memory too. The authors argued that the sequence of decline in cognitive abilities, "may be

linked to the likely spatiotemporal progression of dopamine depletion within the striatum, in relation to the terminal distribution of its cortical afferents, " i.e. the deterioration of cognitive performance proceeds from abilities mediated by frontal cortical regions (executive function) to abilities mediated by slightly more posterior cortical regions (visuo-spatial working memory and verbal working memory).

Cognitive deficits in Parkinson's disease have frequently been accounted for in terms of "bradyphrenia" (Naville, 1922). This term denotes slowed cognitive processing, "a lethargy of the mind distinguished by a lack of interest, initiative, attention, concentration..." (Wilson, 1947) distinct from the effects of ageing, co-existent depression, or dementia. Bradyphrenia is sometimes considered "the mental equivalent of bradykinesia in Parkinson's disease" (Mayeux et al., 1987), a notion consistent with findings that the presence of cognitive impairment is associated with greater severity of bradykinesia (Mortimer et al., 1982). Bradyphrenia has been operationally defined (Brown & Marsden, 1990) as, "a slowing with increasing cognitive complexity above and beyond that shown by a control group".

Brown & Marsden (1990) reviewed the literature and, arguing that though "apparent slowing may be found on some tasks... this may imply a deficit relating to the tasks themselves, rather than reflect a non-specific slowing in cognition," they stated that an effect of slowness of thought over and above the effects of motor dysfunction on manual responses had not been demonstrated by any of the studies reviewed.

Dobbs et al. (1993) stated that the existence of bradyphrenia had been demonstrated by the standards of Brown & Marsden (1990). In a choice reaction time task, patients with Parkinson's disease were less efficient in their use of a warning of the response needed, i.e. they processed information more slowly. Similarly, Cooper et al. (1994) used a Choice Reaction Time (CRT) paradigm in which task complexity could be manipulated without changing response requirements; patients with Parkinson's disease were increasingly impaired as choice complexity increased. However, both these findings are equally compatible with impairment in specialised sub-systems. The results of Cooper et al. (1994) are consistent with a hypothesis of impairment in central executive function and those of Dobbs et al. (1993) are consistent with a deficit in action planning.

Though accounts postulating specific impairments may reflect the ability of parkinsonian patients to perform at normal levels in some circumstances, further parallels with bradykinesia, suggested as a physical model for bradyphrenia, should be considered. Even severely bradykinetic parkinsonian patients may at times exhibit normal levels of performance (Sacks, 1973). Thus it may be premature to conclude that either one of these accounts is wholly correct.

1.3.1.6 Subjective and affective changes

Numerous accounts exist of personality traits said to be characteristic of Parkinson's disease patients. These traits are believed to predate physical symptoms of parkinsonism by years or even decades. A review by Todes and Lees (1985) found consensus in a number of depictions of parkinsonian patients as having in common

an emotional and attitudinal inflexibility, a lack of affect and a predisposition to depression. Patients were frequently described as trustworthy, moralistic and diligent though with considerable suppressed aggressive drive. It should be noted that much of this work originates from a psychodynamic perspective in which personality factors may play a role in the aetiology of parkinsonism.

More recent work, from a neurological perspective supports the pre-morbid personality hypothesis. Ward et al (1984) examined identical twin pairs discordant for the presence of Parkinson's disease. It was found that the affected twin tended from early childhood to be less "usually the leader" and "more self-controlled", and a few years before the onset of the disease had become "less aggressive, quieter and less confident and light-hearted", than their unaffected twin. It has been suggested that the parkinsonian personality is indicative of minor changes in the dopamine-mediated mechanisms which underlie cognitive deficits in later parkinsonism (Lees & Smith, 1983).

Many of the accounts of a parkinsonian personality note the presence of depression in a significant proportion of patients (et al., 1980). In fact, the presence of depression in Parkinson's disease is well established and has been reported in a number of studies (Mindham, 1970; Tandberg et al., 1996). Prevalence has been reported as being between 37 and 90% (Santamaria et al., 1986). Significant numbers of Parkinson's disease patients have been found to have treatment for major depressive illness before the appearance of motor dysfunction (Shaw et al., 1980). It has been stated by some authors that this depression is an intrinsic component of

parkinsonism (Horn, 1974; Hoehn et al., 1976), and by others that it is a reactive depression secondary to the progressive physical disability of Parkinson's disease (Mindham et al., 1976).

Taylor et al. (1986) compared a group of depressed Parkinson's disease patients with a group of endogenous depression patients on a test of short-term memory known to be sensitive to the presence of primary affective disorder. The Parkinson's disease group did not show the same deficits found in the affective disorder group though they did exhibit deficits characteristic of impaired executive function. Taylor et al. concluded that depression in Parkinson's disease is a reactive depression.

In contrast, Santamaria et al (1986) found that depression in Parkinson's disease was associated with a younger age of onset, lower severity of motor impairment (as assessed by observer rating scales), and a positive family history of Parkinson's disease. The severity of depression was not related to severity of parkinsonism. The authors argued that depression is an intrinsic component of Parkinson's disease in a sub-group of patients.

Given the evidence for accounts of both reactive and primary depression in parkinsonism it may be safest to side with Rabins (1982) who concluded that depression in parkinsonism may occur as a primary feature in some patients and secondary to physical disability in others.

1.3.2 Instrumentation

Instrumentation procedures have been applied to the assessment of all the physical features of parkinsonism. This section will provide an overview of the methods investigated, commenting on the properties of the measures and their relative advantages and disadvantages. As in previous sections, some of the evidence derives from studies of DIP, some from studies of Parkinson's disease.

1.3.2.1 Bradykinesia

Attempts have been made to instrument most, if not all of the manifestations of bradykinesia noted earlier (1.3.1.1.1). The most simple methods time the sorts of tasks usually assessed by simple observation (hand turning, tapping etc). Other methods use more complex, or more easily standardised procedures.

Tests of grip strength may be provided using a sphygmomanometer cuff (Onuaguluchi, 1964). This procedure may provide a measure of absolute momentary strength or of the ability to maintain strength at a particular level over a period of time, i.e. fatigability. The ability to maintain repetitive movements over a period of time has been assessed using morse-key tapping. Very simple methods of instrumentation such as these may be effective in instrumenting very specific manifestations of bradykinesia but the information they can provide about the overall severity of the bradykinesia symptom complex is necessarily limited.

Measures of initiation time may be provided by reaction time tests using many different types of stimuli and response. Stimuli may be visual or "kinaesthetic".

Responses may be by pressing switches (Ebmeier et al., 1992), twisting a lever (Evarts et al., 1981), or by touching a computer screen (Riekkinen et al., 1998). Older studies tended to use custom-designed equipment (Evarts et al., 1981) to take measures of reaction time but more recent studies (Ebmeier et al.,1992) have utilised software run on a personal computer. Particularly valuable are methods in which reaction time and movement time may be separated. These allow the investigator to identify groups who exhibit difficulties in movement initiation, a slowed speed of movement, or who exhibit relative increases in choice RT or simple RT (1.3.1.1.). The advantages of reaction time instrumentation concern primarily the high level of accuracy in terms of presentation standardisation and performance measurement; very high test-retest correlations have been found for both simple and choice reaction time tests (Lowe and Rabbitt, 1998). Also, they provide a measure of pure speed of movement, uncontaminated by other factors such as motor control.

Other studies have chosen to use fine motor control tasks which provide an assessment of movement accuracy as well as simple speed of movement. One of the most simple is the pegboard. The plainest form of pegboard is a board drilled with holes into which round wooden pegs are to be inserted (Verkerk et al., 1990); in more complex versions the pegs are keyed identically and must be inserted into the randomly-oriented holes in the correct orientation (Meier & Martin, 1970). More complex tasks simulating activities of daily living are also popular, particularly amongst neurologists. These tests provide an assessment of functional ability (Jebsen, 1969). As such they have direct relevance to self-care and may be used as an index of treatment success. The advantages of tasks such as these stem in part from

their obvious face validity. In addition they require little or no training for the patient which may help to minimise practice effects. However, the properties of these tests (reliability, validity) vary greatly and have only been systematically investigated for a relatively small proportion of them.

Another form of tests which tap motor control abilities comprises pursuit tracking tasks. In these tasks, the patient uses a joystick to control a trace on an oscilloscope (Flowers, 1978) or on a computer screen (Bloxham et al., 1984). The task is to maintain the position of the trace as close to that of a moving target as is possible. Performance is scored on the basis of time spent on/near the target during the test period.

The range of instrumentation techniques which have been used to quantify bradykinesia is wide. It includes procedures which instrument simple speed of movement, the purest "bradykinesia," procedures which instrument other manifestations, such as initiation difficulties, or the inability to maintain repetitive movements, and also more complex batteries of tests which aim to provide a comprehensive assessment of the consequences of bradykinesia for functional ability. However, despite the range of techniques used, there has been little investigation into the validity of the techniques which tends to be assumed rather than demonstrated. Use of instrumentation methods is sometimes justified on the basis that they are procedures which detect impairment in parkinsonism rather than procedures which measure a particular feature of parkinsonism. Few of these techniques have been used with cases of DIP.

1.3.2.2 Rigidity

The range of procedures used to instrument rigidity is perhaps less wide than of those used in the case of bradykinesia. This is probably due more to the much narrower construct definition of rigidity than to any lack of imagination amongst investigators.

Some of the earliest instrumented methods of assessing rigidity used gravity to drive the limb(s) in which tone was to be assessed. The Wartenburg test seats the subject on the edge of a high chair or table with their legs able to swing freely; the physician lifts the legs to almost horizontal and then lets them drop. The more rigid the subject's legs, the fewer times they will swing before they come to rest. A computer has been used to record the number of swings made and the velocity attained (Brown et al., 1988a & 1988b). The most notable difficulty with this method is for the patient who must relax sufficiently to allow the legs to swing freely without either exaggerating or damping the motion. Brown et al. state that either of these situations may be easily detected from the computer record of the trial.

Rigidity assessment often involves the use of quite large and cumbersome machines to which the subject must be fastened in order that the limb from which the measurement is to be made may be passively moved back and forth. Numerous examples of this kind of work have been published since the 1950s (Long et al., 1964; Webster 1966; Caligiuri et al., 1989; Caligiuri & Galasko, 1992; Walsh, 1992).

Most recent studies of rigidity have used the activation paradigm developed by Webster (1966) to take account of the involvement of higher-level central nervous system influences in the production of parkinsonian rigidity. Rigidity is assessed in both "resting" and "activated" conditions and the results compared. In the activated condition a motor task is performed by the contralateral (i.e. non-test) limb. In normal subjects, no difference is found between the two measurements but in parkinsonism stiffness increases in the activated condition. Motor tasks which have been used successfully include pursuit tracking (Webster, 1966), isometric contraction (Kirollos et al.,1996), drawing circles in the air (Caligiuri & Galasko, 1992).

In a number of studies conducted by the same group of workers (Caligiuri et al., 1989; Caligiuri & Galasko, 1992), "stiffness values are obtained by applying known displacements and measuring the resultant force". In the study of 1989, stiffness was tested in the hand. The finger rests in a cradle mounted on a beam. The beam may be moved, by the experimenter, through a range of 30° (Caligiuri et al., 1989) or 40° where possible (Caligiuri & Galasko, 1992). A potentiometer was used to measure the displacement (in degrees) through which the finger was moved and a strain gauge mounted on the beam allowed the measurement of the resultant force (in grams). Stiffness is defined as the ratio of force/displacement. On each trial (a raising and lowering of the finger) the peak stiffness was obtained using a computer sampling at a rate of 100 samples/sec. The peak stiffness measures for 20 trials were averaged for both the resting and activated conditions. The coefficient of stiffness was obtained

from the ratio of the coefficient in the resting condition to that in the activated condition.

However, the methodology has serious problems limiting the information which can be extracted (Walsh, 1992). Measurements of rigidity in which movement is not continuous will be confounded by the presence of thixotropy in any estimation of stiffness. Thixotropy is a property seen in materials such as paints and sauces, such that stiffening occurs when the substance is at rest; the process may be reversed by agitation. In order to avoid confounding by this phenomenon measurements of stiffness should not be made from rest.

A later study from the group (Caligiuri & Galasko, 1992), used a series of trials recorded as one continuous movement, a practice which eliminates the confounding effect of thixotropy. The influence of thixotropy may be seen in the force and displacement traces published (fig. 2, page 4). The first (few) cycles of each condition, performed after a period of rest, show a higher level of resultant force than in the later cycles. This occurred, in the active condition at least, in the presence of a lowered figure for displacement. On these first three or four trials, the patient exhibited greater apparent stiffness than on the later trials; it is likely that this reflected the influence of thixotropy. However, the authors do not address this effect and do not specify whether the results of all trials were included in the analysis.

Despite the use of an improved method, the continued use of displacement as the independent variable causes further methodological difficulties. If the joint is moved

through a large range of movement, measures of stiffness can be confounded by differential viscosity at the extremes of movement. Smaller movements avoid this source of confounding but can only inform about the properties of the joint in a very limited range.

A further method of rigidity assessment was developed by Kirollos et al. (1996), who used a motor-powered mechanical device similar to that used by Webster (1966). Rigidity was assessed at the elbow from the work required to move the forearm through a fixed angle of 40° at a constant rate of 0.5 Hz. The arm was lightly strapped to a cradle mounted on a lever attached to the motor. The forearm moved in a horizontal plane, about a pivotal axis aligned to the elbow joint. Torque was measured using a semiconductor strain gauge and the angle of displacement using a potentiometer. The measure of stiffness used was the work required to move the arm per unit displacement (measured in $Nm^{\circ -1}$). Stiffness was compared using the activated stiffness above baseline.

An alternative method (Walsh, 1992) uses torque applied to the joint as an independent variable. This evades the confounding influences of thixotropy and differential viscosity. Positive feedback is added to the motion of the limb, causing the limb-device combination to oscillate. The man/machine combination forms an under-damped torsion pendulum, the behaviour of which can be described by an equation:

$$f = \frac{1}{2\pi}\sqrt{\frac{K}{J}}$$

f = resonant frequency

K = stiffness

J = inertia

When the motion of the device is tracked, the stiffness in the limb may be calculated from the frequency of oscillation.

Common themes in almost all the modern methods of rigidity instrumentation is the controlled movement of the limb under assessment by an external force, the measurement of forces applied to the limb and movement of the limb, and the use of the activation paradigm. The methods differ in the means of data analysis, principally as to whether the independent variable is taken to be displacement (Caligiuri et al., 1989; Caligiuri & Galasko, 1992), or torque (Walsh, 1992). These differences appear to be superficial but are central to the validity of the instrumentation procedure.

1.3.2.3 Tremor

Of the different methods which have been used to instrument tremor, by far the simplest is an objectified form of spirography. The patient is asked to draw a spiral between printed guidelines and the results are rated for accuracy (Bain et al., 1993); this form of assessment is, however, rater-subjective rather than truly objective. A modified method was developed by Verkerk et al. (1990). The time taken to draw the spiral was used as a score with time added for errors; three seconds for each time a

guide line was touched, five seconds for each time a line was crossed. As the authors do not indicate how these penalty times were determined the validity of this method must be regarded as doubtful at best.

As a means of assessing tremor, spirography appears of little use. Performance will be greatly affected by bradykinesia, fine motor control ability, and strategy (i.e. whether the patient aims to complete the spiral as accurately as possible regardless of the time taken or as quickly as possible despite the increased errors). Spirography methods may be of use where an assessment of hand function is required; no information about the properties of the tremor is derived, simply an indication of the level of impairment caused by the tremor. Other, more complex, methods of instrumentation provide additional, quantitative data about the properties of the tremor, most commonly the overall amplitude or information about its frequency components, e.g. peak frequency.

In recent work, the use of electronic accelerometers is ubiquitous but the literature contains examples of other methods. Walsh (1996) summarises a method dating from the nineteenth century, developed by Schaefer, which held the wrist motionless and transcribed the motion of the hand onto a smoked drum. The results obtained show an irregular waveform as is obtained with modern methods. However, Walsh describes the attainment of since replicated results as, at least partially, the product of good fortune. Another device, designed by Walsh (Lakie, Walsh & Wright, 1986), was described as a "hanging hand tremorgraph". The wrist was held still by one strap and the hand attached to an adjustable crank by another strap. The crank was

connected to an induction generator used as an angular accelerometer. A polygraph transcribed the signal from the accelerometer and the other data collected, including EMG.

Electronic accelerometers are now far more common than electro-mechanical devices. As with all electronic equipment there has been continuous development, from simple devices consisting of a valve, bridge circuits and a polygraph (Marshall & Schnieden, 1966), to integrated circuits and the use of computers for data collection, storage and analysis (Marsden et al., 1969; Hömberg et al., 1987; Caligiuri et al., 1991). The accelerometer detects acceleration in a given plane and emits a voltage proportional to that acceleration which may be recorded for later analysis.

To measure tremor in the wrist only, the forearm must be supported, usually with the hand pronated. The hand may be hanging (for a measurement of resting tremor) or held out horizontally (for postural tremor). The accelerometer may be attached to the dorsal surface of the hand, fastened to the middle finger by a velcro strap or gripped in the fist. Marsden et al. (1969) mounted the accelerometer at the tip of the fingers but other authors have argued that this may add the effects of tremor in the interphalangeal joints, confusing the pattern of results, or ignoring important components of tremor in disease (Wade et al., 1982).

Different studies have also focused on alternative features in the processed data. Much of the work by Walsh (1996) concentrates on the total movement present in

the tremor in units of acceleration. However, Arblaster et al. (1993) and most other authors argue that it is the peak frequency which characterises tremor in parkinsonism and distinguishes it from other tremors. In addition, there is great variation in the amplitude of tremor found in healthy adults, with Walsh (1992, 1996) finding a three-fold variation between some groups and much more between individuals. Much less variation is found in peak frequency in normal groups (Marsden et al., 1969).

For the purposes of neurological assessment of disability, tremor frequency is of little consequence relative to amplitude. However, if the goal is to distinguish parkinsonian tremor from that due to other causes then an indication of the frequency components present is essential. This requires the use of a Fast Fourier Transform procedure to analyse the frequency components of the tremor.

The Fourier transform is a mathematical technique for expressing a(n irregular) waveform as a weighted sum of sines and cosines (regular waveforms). The Fast Fourier Transform is an algorithm (a detailed sequence of actions to perform to accomplish a particular task; named after an Iranian mathematician, Al-Khawarizmi) for computing the Fourier transform of a set of discrete data values. Given a finite set of data points, for example a periodic sampling taken from a real-world signal (e.g. an accelerometer), the FFT expresses the data in terms of its component frequencies.

As methods of instrumenting tremor vary, so do the methods of analysing the data produced. Though almost all modern studies use FFT procedures to derive the

amplitude of frequency components present in the tremor, further processing may proceed in a number of different ways. In many studies, the peak frequency of tremor is used. The mean peak frequency for one group may be compared with that for another group (Tyrer et al., 1981), or for the same group after changes in medication (Pullinger & Tyrer, 1983). Alternatively, comparisons may be made using the number of cases within a group whose peak frequency falls below some threshold (Arblaster et al., 1993). Another method calculates the amplitude of tremor within frequency bands (Caligiuri et al., 1989b; Caligiuri et al., 1991; Caligiuri & Lohr, 1993). The amplitude of tremor present within the frequency band 3-7 Hz (Caligiuri & Lohr, 1993) or 4-6 Hz (Caligiuri et al., 1991) has been suggested as an index of the severity of parkinsonian tremor.

As noted before, accelerometry techniques are ubiquitous amongst modern methods of tremor assessment, as is the use of computers in data recording and processing, and FFT algorithms in analysis. These techniques now have well established properties; test-retest reliability is high ($r = .75$, $p < 0.01$; Caligiuri et al., 1989b). However, there is little concurrence on the position in which tremor should be measured or on the means by which the results of FFT analysis should be compared. This latter issue is a problem even between different studies conducted by the same group (Caligiuri et al., 1991; Caligiuri & Lohr, 1993). There seems little evidence on which to base a choice between these methods of analysis; they are mostly justified solely on the basis that tremor in Parkinson's disease is known to have a dominant frequency of around 4-6 Hz.

1.3.2.4 Other features

Instrumentation procedures have also been used as measures of gait dysfunction. Dietz et al. (1981) used a treadmill to keep their subjects stationary during gait analysis. Potentiometers were placed at the heel and at the ankle to measure changes in joint angle. The timing of ground contact was measured by separate electrical switches placed inside the shoes at the ball and heel of the foot. EMG was also taken from the medial head of gastrocnemius, tibialis anterior and soleus to assess the timing of muscle activity during gait.

In a number of studies, Dobbs and others (e.g. Kirollos et al., 1993) used a "gait assessment trolley" which the patient towed behind him. A three metre length of strong cotton clipped to the heels of the patient's shoes passes around a pulley attached to a "shaft encoder", mounted on a lightweight trolley. When walking, a length of cotton is transferred from behind one foot to behind the other; this rotates the shaft encoder and tows the trolley. The length of cord transferred represents the distance moved and the direction of encoder rotation indicates which foot has moved. The trolley is designed to maintain tension in the cotton. A battery-powered infrared transmitter sends encoded information to a receiver and a chart recorder.

Though highly ingenious, both of these methods, and especially that of Dietz (Dietz et al., 1981), are time-consuming, expensive and intrusive for the patient.

1.4 Aims

These hypotheses were formulated from the literature previously reviewed and will be explained here within the context of the literature. Similar attention will be paid to methods by which the hypotheses may be evaluated.

1. Instrumentation has a role in the assessment of DIP.
2. Bradykinesia is the predominant feature of DIP.
3. Features of parkinsonism other than the physical ones will be demonstrable in DIP.

1.4.1 Instrumentation has a role in the assessment of DIP

This assertion is central to the work done, and is the most important of the hypotheses. It must take precedence over the other hypotheses, here in considering the means to prove the hypotheses, in the analysis of the data recorded, and in the discussion of the findings.

In order to demonstrate that instrumentation has a role in the assessment of DIP it is necessary to demonstrate that instrumentation (or at least the particular methods of instrumentation evaluated) is an accurate means of assessing parkinsonism. The methods investigated may be evaluated as measures of a particular feature of parkinsonism, whether it is bradykinesia (Evarts et al., 1981), rigidity (Caligiuri & Galasko, 1992), or tremor (Caligiuri & Lohr, 1993), or as markers of the overall severity of parkinsonism (Arblaster, 1993).

The properties of the instrumentation methods used must be formally evaluated against certain criteria. These criteria assess the degree to which measurement occurs in a systematic and accurate manner. It is the performance of a test against these criteria which will determine its value.

1.4.1.1 Reliability

Reliability concerns the extent to which a measuring procedure yields the same results on repeated trials, i.e. the consistency of the procedure (Carmines & Zeller, 1979). This consistency is achieved by the minimisation of chance random error. However, measurements always contain a certain amount of random error and so an element of unreliability is unavoidably present. Reliability is thus a matter of degree rather than all-or-nothing. However, though a high level of reliability is necessary to a good form of measurement it is not sufficient in itself to recommend a procedure: validity is essential.

1.4.1.2 Criterion Validity

Validity can be assessed in different forms, most commonly as criterion validity or construct validity. Criterion validity is concerned with the presence of non-random error in measures taken, i.e. systematic bias in results. It is assessed with reference to some other, external, criterion; usually this other criterion is an existing Gold Standard of assessment for the phenomenon or property under consideration. For example, an abbreviated assessment procedure may be evaluated relative to a more

comprehensive form. The degree of criterion validity exhibited by the measure being evaluated is indicated by the level of correlation with the Gold Standard criterion.

A distinction may be made between concurrent and predictive criterion validity. Concurrent validity is assessed by the degree correlation between the measure and the criterion at the same point in time. Predictive validity concerns the ability of the measure to predict the results of the criterion measure at some future point in time. Essentially, the logic and procedures are the same for both concurrent and predictive validity; only the point at which validity is assessed differentiates them.

Criterion validity is expressed using a correlation co-efficient to describe the relationship between the results of the measure to be assessed and the criterion. An ideal measure would exhibit a perfect correlation with the criterion (correlation coefficient $r=1$). Realistically, a correlation co-efficient of $r>0.75$ is regarded as a high degree of correlation and $r>0.5$ is regarded as moderate. Commonly, the correlation co-efficient has an associated significance value which indicates the confidence with which the co-efficient value may be relied upon.

Though not theoretically complex and relatively simple to determine, criterion validity is wholly dependent upon the validity and reliability of the criterion chosen. In some circumstances there may be no suitable criterion. Or, one may wish to validate a measure without reliance upon existing measures. In these circumstances, construct validity must be investigated.

1.4.1.3 Construct Validity

Construct validity depends upon the extent to which the performance of a measure relates to theoretically derived hypotheses concerning the constructs being measured. This form of validation depends upon the existence of a relatively extensive theoretical background to the concept being measured.

Establishing construct validity involves three distinct steps. First, the theoretical relationship between the concepts themselves must be specified. Second, the empirical relationship between the measures of the concepts must be examined. Finally, the empirical evidence must be interpreted in terms of how it clarifies the construct validity of the particular measure.

Unlike criterion validity, construct validity cannot be expressed in a numerical value. It is for the investigator to interpret the supporting evidence and decide if it is sufficient to establish the validity of the measure.

If the evidence does not support the validity of the measure, there may be four possible interpretations (Cronbach & Meehl, 1955). Firstly, and most commonly, it may be concluded that the measure lacks construct validity, i.e. it does not measure what it purports to measure (it may measure some other construct, but not the construct of interest). Secondly, one may question the theoretical framework used to derive the predictions. Thirdly, the method or procedure used to test the hypotheses is inappropriate (this may be as simple as the use of an unsuitable statistical

technique). Finally, there may be a lack of construct validity or reliability in some other variable(s) in the analysis.

Construct validity demonstration is particularly important when wishing to develop a more accurate measure than the existing procedures. Criterion validity may be established against the existing measure as a starting point but the techniques used to establish validity in the absence of a criterion must be used to demonstrate that differences between the results of the two measures are due to greater accuracy in the new measure rather than in the old.

1.4.1.4 Sensitivity and specificity

Sensitivity and specificity describe the ability of a measuring procedure to quantify a particular property and that property alone. They are assessed relative to an existing criterion. As such they may be viewed as components in criterion validity. Further, assessment of these two properties is dependent upon the existence of a valid criterion for the property in question.

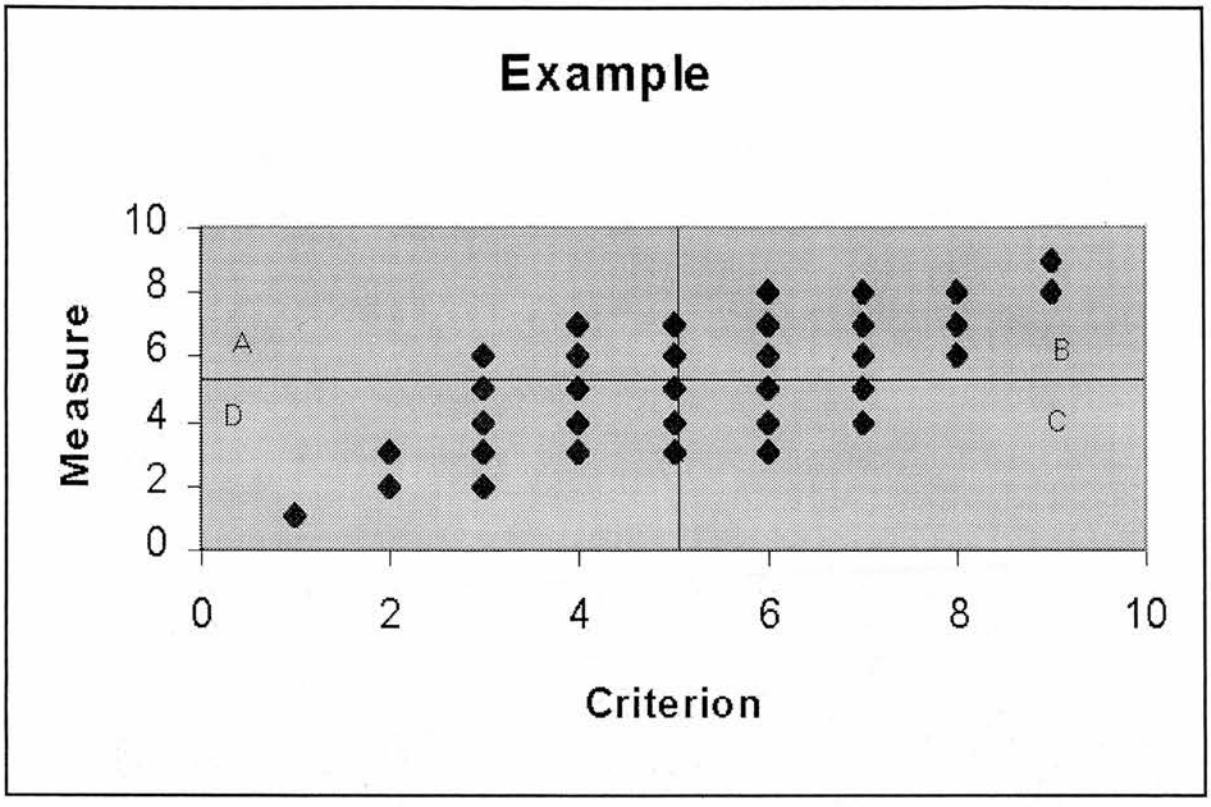
Sensitivity and specificity are often used to evaluate the performance of measures when the measure is used to divide cases into those which exhibit a particular property and those which do not, by means of threshold value. Cases which score above a particular threshold level are labelled positive, and those below the threshold are labelled negative. Within this context, sensitivity refers to the ability of the measure to detect genuine positive cases and specificity refers to its ability to avoid falsely identifying negative cases as positive.

These two properties may be quantified relative to the performance of a criterion (Greenhalgh, 1997). Sensitivity is defined as the percentage of genuine cases (as identified by the criterion) which are identified by the measure. Specificity is the percentage of cases identified by the measure which are genuine cases (again as identified by the criterion). These properties are not independent. If the threshold value is altered this will have opposing effects on the apparent sensitivity and specificity of the measure.

Using the graph below (Example 1) the relationship between sensitivity and specificity may be more clearly illustrated. The x-axis represents the results of assessment using the criterion. The y-axis represents the results of the measure under assessment. For both the measure and the criterion, cases exhibiting the property are scored more highly than those that are not. It can be seen that a good but not perfect degree of correlation exists between the results of the measure and the criterion.

Within the graph, cases which fall in sector A genuinely exhibit the property and are identified as doing so by the measure being assessed. Those which fall in sector B also exhibit the property but are not identified as doing so by the measure. Sector C contains those cases which are correctly identified by the measure as not exhibiting the property. Cases falling in sector D do not exhibit the property but are wrongly identified by the measure as doing so: they are "false positives".

Example 1



By shifting the threshold at which the measure identifies positive and negative cases, the sensitivity and specificity may be manipulated (though the criterion validity of the measure has not changed). If the threshold of the measure is increased the number of cases identified by the measure is decreased. There are fewer cases in sectors A and D and a greater number of cases in sectors B and C. Fewer cases are correctly identified but the number of false positives is reduced. Specificity has been increased but sensitivity is reduced.

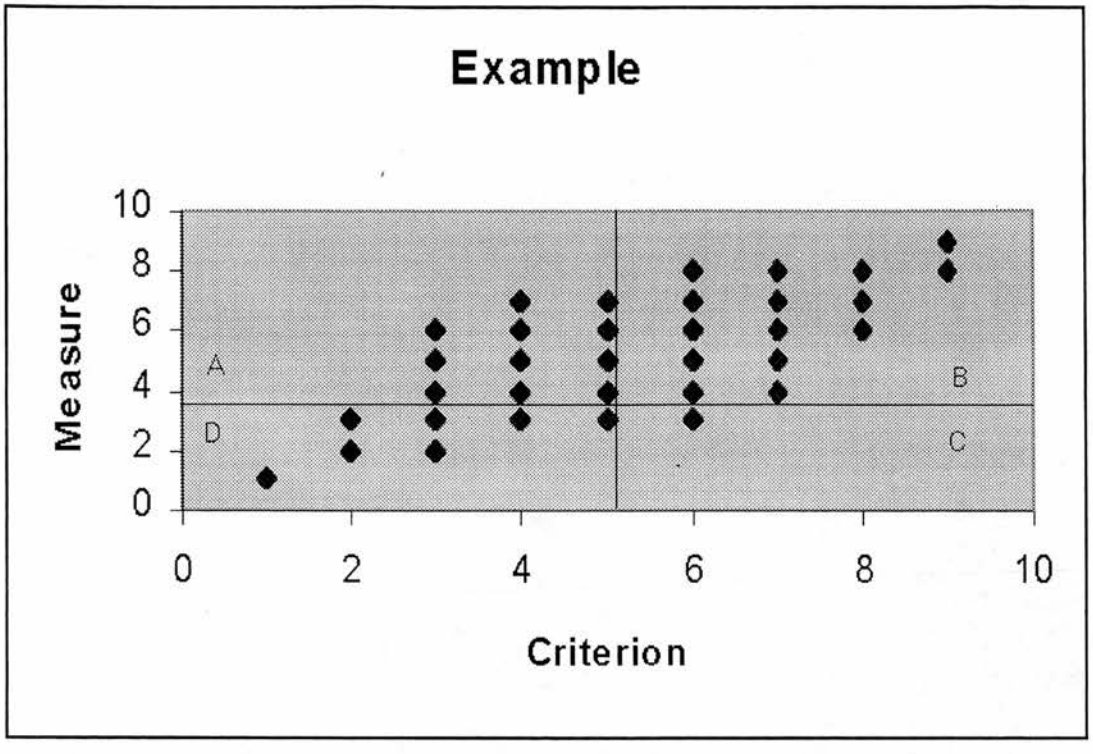
If the threshold of the measure is decreased the number of cases identified by the measure increases. There are more cases in sectors A and D and fewer cases in sectors B and C. A greater proportion of genuine cases are correctly identified but

there are more false positives. Sensitivity is increased at the expense of reduced specificity.

The optimum threshold level for a measure is dependent upon its intended purpose and the circumstances in which it is to be used. If it is desirable that all cases exhibiting the property are identified and the presence of some false positives is not important, high sensitivity is paramount and a low threshold value should be used. If, however, it is important to avoid false positives, specificity should be maximised by the use of a raised threshold.

Example 2 uses the same results as the earlier illustration but the positive/negative threshold on the new measure is lower. The small number of cases in sector C indicates how few of the cases identified as positive by the criterion were not identified by the new measure. In contrast, a great number of cases were detected by the new measure which were not identified by the criterion. This pattern of results may signify greater sensitivity in the new measure relative to the criterion, or less specificity. In order to determine which of these two explanations is more satisfactory the investigator must use the techniques detailed earlier for the demonstration of construct validity. This recourse to the theoretical framework will allow the investigator to decide which of the measures is the more accurate.

Example 2



1.4.1.5 Summary of test properties

The test properties described here are essential qualities for any measuring procedure. Use of a measure is only justified if it can be demonstrated that the measure exhibits high levels of reliability, validity, and sensitivity and specificity. The analysis techniques can enable the quantification of the extent to which a procedure fulfills these criteria. However, it must be noted that the criterion validity (and sensitivity and specificity) of a measure can only be established to the extent that the criterion has validity. Thus, the success of the enterprise is dependent upon the choice of a suitable criterion. This is dependent not only upon the property to be measured but the use to which the measure will be put.

1.4.1.6 Other considerations

The instrumentation methods will be investigated first as measures of particular features of parkinsonism, and then as markers of the overall severity of parkinsonism (where the literature suggests this is appropriate). This latter form of assessment relates most particularly to accelerometry assessment of tremor which has been described as providing an "early warning" of the development of sub-clinical parkinsonism (Arblaster et al., 1993).

However, demonstrating that instrumental measures provide a valid means of assessment of parkinsonism is not sufficient to demonstrate the existence of a role. The other characteristics of the measures must be assessed to determine whether they have some advantage over the observer ratings used in clinical practice. An advantage may be shown in terms of being a better, more accurate, measure of parkinsonism. Or the instrumentation may be easier to use for the investigator, less time-consuming for the patient and the investigator, less intrusive for the patient, less demanding of the investigator (in terms of experience, training and skill level required). It is hoped that instrumentation may allow the relatively unskilled investigator to quickly produce accurate, reliable and valid assessments of severity of parkinsonism without confounding by inter-rater differences.

1.4.2 Bradykinesia is the predominant feature of DIP

It is stated that bradykinesia (particularly upper-body bradykinesia) is the cardinal feature of parkinsonism (Quinn, 1995). Though there is an absence of comparative studies, clinical impression strongly suggests that this is the case. In other words, that

rigidity and tremor are of lower prevalence than bradykinesia and/or that rigidity and tremor are of lower relative severity than bradykinesia. Clinical impression also suggests that bradykinesia is the dominating symptomatology in DIP to an even greater extent than in Parkinson's disease.

A demonstration that one feature is of greater severity than another, or of greater prevalence is deceptively complex. It is necessary to determine on what grounds this increased severity or prevalence is to be assessed. Perhaps the simplest method is to consider which items on the observer rating scales make most contribution to overall score variance. If total scores are more influenced by bradykinesia item scores than say, tremor scores, one could conclude that bradykinesia has a more central role in symptomatology than tremor. However, to a large extent, this analysis would rely upon the construction of the scale to provide equal weight to each feature of parkinsonism. Use of a scale which gave undue to significance to ratings of a particular feature would confound efforts to assess the contribution of that feature and other features to overall severity of parkinsonism (this point will be revisited later 2.3.1).

In the previous section, the intention to assess instrumentation methods as markers of overall severity of parkinsonism (using whole-scale observer ratings of parkinsonism as criteria) was noted. The results of this form of evaluation may provide evidence to indicate the predominating form of symptomatology. An advantage of this method is that the observer ratings are used as whole scale assessments of parkinsonism, as they were designed to be used. If a particular instrumentation procedure can be

identified as being the best indicator of overall severity of parkinsonism, it could be concluded that the feature assessed predominated in the patient group used. However, it is also possible that a particular method may show greatest value as a measure of overall parkinsonism simply because the other instrumentation procedures do not measure the features of parkinsonism they are intended to measure. The greater accuracy of one form of measurement would not be valid evidence that the feature so measured was the predominant form of symptomatology.

Possibly the simplest method, and one which may avoid many of the difficulties noted above is to consider the numbers of patients who are outside the normal range on measures of each feature. This analysis may be completed principally using the instrumentation methods. Results from the control group can be used to determine a normal range of scores for each method of instrumentation, and then the numbers or proportions of patients who fall outside the normal range for each feature of parkinsonism assessed may be calculated. Alternatively, a similar method may be used using the observer ratings with which the rating criteria may be used to determine a normal range.

While it is hoped that the instrumentation methods may have a role in answering questions of this nature, perhaps by providing enhanced sensitivity of measurement, they cannot be used in this fashion unless their validity as measures of the features of parkinsonism has been satisfactorily demonstrated (see hypothesis 1).

1.4.3 Cognitive and subjective features of parkinsonism are present in DIP

Cognitive deficits and other subjective features such as changes in personality have been documented in Parkinson's disease. However, evidence of the existence of these phenomena in DIP is limited. It is hypothesised that these phenomena do exist in DIP and that their existence may be demonstrated. Again, this aim is deceptively complex. The principal difficulty is the extent of the superficial similarities between features of parkinsonism and deficit symptoms of psychosis. This phenomenal overlap between psychosis and parkinsonism was described earlier but includes cognitive deficits, psychomotor slowing, and can include motor abnormalities.

In circumstances such as this, longitudinal experimental designs are frequently used. It is often possible to assume that features which emerge after commencement of treatment are the result of treatment, in this case that they would be parkinsonian in nature. However, the fact that psychotic symptomatology is not stable over time makes it an intricate matter to determine that changes in deficit severity are due to treatment factors. An alternative method is to use a cross-sectional experimental design to consider the presence of cognitive deficits across a patient group.

In order to demonstrate that deficits found are features of DIP rather than the underlying psychosis it is necessary to demonstrate an association between the presence of the deficits and the presence of other features of parkinsonism. Within Parkinson's disease, associations have been demonstrated between cognitive deficits and bradykinesia (Mortimer et al., 1982). If the presence of a particular cognitive

deficit is associated with the severity of bradykinesia, it can be concluded that the deficit in question is a feature of parkinsonism rather than of the illness being treated.

2 Methodology

This section will consider the measures selected for evaluation, the observer-ratings to be used as criteria, and the statistical techniques to be used to analyse the data.

2.1 Measures

The measures selected will be described in the context of the features of which they are intended to provide an assessment.

2.1.1 Bradykinesia – Jebsen Hand Function Test

Two methods were chosen to instrument bradykinesia: the Jebsen hand function test (Jebsen et al., 1969) and the CANTAB reaction time test. The two procedures are dissimilar in structure and administration, being designed to measure different aspects of bradykinesia.

The Jebsen hand function test is a comprehensive battery of fine motor control tasks. The battery was intended for both clinical and research use as an assessment of functional capability and treatment efficacy. It is suitable for measuring a broad spectrum of manual function in different populations.

The authors originally used the test with seven tasks: (1) writing; (2) turning over cards; (3) picking up small common objects and placing in a container; (4) stacking checkers (draughts pieces); (5) simulated feeding; (6) moving large light objects (empty large tins); (7) moving large heavy objects (full tins). All tasks are assessed purely on the basis of time taken to complete the task.

The Jebsen test has many qualities to recommend its use. Primarily it is a test of functional ability, tapping the impaired performance in activities of daily living which is distressing to patients and has negative consequences for treatment success. In particular, the simulated feeding task instruments the patient's ability to perform a primary task of self-care. The variety of tasks in the test are sensitive to performance in many facets of motor function. Strength, speed of movement and accuracy are all required to perform the tasks swiftly.

The Jebsen test exhibits good psychometric properties of validity and reliability. The use of common household objects and the inclusion of tasks very obviously related to normal activities lends the Jebsen test a high degree of face validity. Further, the authors have demonstrated that it also exhibits a high degree of test-retest reliability ($r = 0.60$ to 0.99; Jebsen et al., 1969). Further, practice effects are negligible, perhaps a benefit of the use of everyday objects in simple tasks.

The study of the Jebsen group indicates that tasks 2-7 exhibit very similar properties, including a similar range of scores. However, task 1 shows a much higher rate of failure to complete, with a proportion of higher scores tending to be much more extreme than on the other items. While this may indicate a higher level of sensitivity to motor control dysfunction in the writing task, it may indicate that the task does not sit well with the other tasks.

Writing is known to be very sensitive to parkinsonism, and micrographia is often cited as an early marker of incipient parkinsonism (1.2.2). However, assessment of

writing performance is usually on the basis of observer-rated changes in size and quality of writing. It is the size and steadiness of the writing rather than the speed which it is produced which marks parkinsonism. There is little indication in the literature as to what extent writing speed is indicative of bradykinesia rather than non-medication factors such as cognitive impairment or educational level. Further, it is changes in the writing relative to previous assessments, rather than properties of the writing at any one time, which are used as indicators of dysfunction.

A pilot study here found very high levels of failure to complete task 1 (particularly with the non-dominant hand) and it was omitted from the battery for the study proper as it detracted from the integrity (and construct validity) of the battery as a whole.

Following the omission of task 1, six tasks from the Jebsen test were included in the assessment of bradykinesia. All tasks were performed with each hand separately, and timed.

1. Card turning using 3"x 5" cards (simulated page turning). Five cards are placed in a horizontal row, 2" apart, oriented vertically on the desk in front of the subject. Timing is from the word "Go" until the last card is turned over. No accuracy of placement after turning is required.
2. Picking up small common objects. A large empty tin is placed in front of the subject, 5" from the front of the desk. The objects (2 paper clips, 2 pennies, 2 bottle tops) are placed in a row alongside the can. Timing is from "Go" until the last object strikes the inside of the can.

3. Simulated feeding. Five kidney beans are placed on the board (a large wooden board, secured to the desk, with an upright central piece of plywood (2" high) glued to it), touching the upright and 2" apart. An empty tin is placed in front of the subject, and a teaspoon is provided. Timing is from "Go" until the last bean touches the inside of the can.
4. Checkers. Four draughts pieces were placed in front of and touching the board. The pieces are to be stacked one on top of another. Timing is from "Go" until the 4^{th} piece makes contact with the 3^{rd}.
5. Large light objects. Five empty 400g size tins were placed in front of the board, 2" apart with open end facing down. The tins are to be moved to stand on the board. Timing is from "Go" until the last tin is released.
6. Large heavy objects. The task is as the previous task but cans are full (400g weight).

Total time taken to complete all tasks in the battery with both hands was calculated. In the Jebsen study, it was necessary to calculate time taken for the dominant and non-dominant hands separately as one patient group participating in the study was a hemiparesis group. However, this consideration is not relevant to this study. The use of a combined time for the two hands may be justified on the grounds that the observer-rating criteria to be used for bradykinesia do not differentiate between the two hands (in contrast to those of rigidity and tremor), being ratings of overall speed of movement. The use of a combined total for all six tasks is intended to produce a comprehensive combined assessment of overall upper-body bradykinesia. Similar

properties were demonstrated for the tasks which all proved sensitive to impaired motor performance (Jebsen et al., 1969).

2.1.2 Bradykinesia – CANTAB

The CANTAB (Cambridge Neuropsychological Tests, Automated Battery) is a PC software package which contains a number of common neuropsychological tests, of which the reaction time test was selected. Stimuli are presented visually on the screen and responses are via a press-plate and a touch-sensitive screen. On all tests, feedback is provided to the subject immediately after each trial.

The CANTAB reaction time task contains both simple reaction time and choice reaction time conditions. In both conditions, the subject responds by touching the computer screen. In the simple reaction time condition a circle is presented in the middle of the screen (always in the same position); the subject is asked to touch the centre of the circle as quickly as they can after a dot appears in the circle. In the choice reaction time condition five circles are presented on the screen (again, always in the same position); the subject is asked to touch the centre of the circle in which the dot appears. A press-pad is used in some sections of the task, the hand resting upon the plate until the stimulus appears. The use of the press-plate enables the separation of reaction latency and movement latency. Reaction latency includes time to view and process the stimulus information, and plan and initiate the response. Movement latency is solely the time taken to move the hand from the press-plate to the screen. The separate recording of reaction and movement latencies means that a

differentiation may be made between patients who are slow to move and those who may be unimpaired in movement speed but are slow to initiate their movements.

The reaction time test is presented in five sections, of which three are forms of simple reaction time, and two are forms of choice reaction time:

1. subject must touch a circle on the screen when a yellow dot appears in the centre of the circle;
2. as previous stage, but dot may appear in one of five circles;
3. press-pad must be held down until dot appears, subject does not need to touch the screen;
4. press-pad is to be held until dot appears, and then screen touched (single circle);
5. as previous stage, but with five circles.

The first three of these function primarily as practice stages for the responses required in the final two sections.

Being administered by computer, the CANTAB has certain valuable properties. Presentation of stimuli is more consistent from patient to patient than can be achieved by a human investigator. Response latencies are measured with great accuracy. Finally, the administration of the test is simpler and quicker for the investigator.

The use of the CANTAB with the patient groups involved in this study is well established, its use having been validated in varied contexts and with different groups: schizophrenia (Pantelis et al., 1997), depression (Purcell et al., 1997),

Parkinson's disease (Riekkinen et al., 1998), tests of drug effects (Elliott et al., 1997).

The purpose of including the CANTAB reaction time task in addition to the Jebsen test is, in part, to provide a more pure assessment of motor speed. The Jebsen test demands the performance of complex sequences of controlled movements. In contrast, the CANTAB movement latency provides an uncontaminated measure of speed of movement after the response has been initiated. The capacity to analyse initiation latency distinct from movement speed is also valuable. Further, the CANTAB also allows the calculation of a motor planning latency which may be used as a marker of cognitive slowing (2.1.5).

2.1.3 Rigidity

Rigidity is to be assessed using a positive feedback device previously used by Walsh (1992). As with most other forms of automated rigidity assessment this device is used to examine the relationship between the force applied to the limb and the resultant motion of the limb.

The positive feedback device consists of a printed motor mounted with the axle positioned vertically. A metal beam is fastened perpendicular to the axle. Mounted on the beam is a padded cradle in which the subject's forearm rests. The arm is held in place with velcro straps. The axis of movement at the subject's elbow is positioned concentric with the axle of the motor.

The velocity through which the lever is moving is monitored by the device. When motion is detected, a force is applied to the lever in the same direction as the movement. The torque generated by the motor causes the limb to move until its momentum is restrained by stretching the tissues of the joint about which motion is occurring and a rebound occurs. As soon as the velocity reverses, the current in the motor (and hence the torque) reverses, pushing the limb backwards until it is again checked by elasticity. Once again the current reverses, accelerating the limb in the opposite direction. With sufficient gain in the loop the oscillations will become self-sustaining for as long as the system is energised.

The output from the device consists of three signals conveying displacement of the lever (and attached limb), velocity of the lever and torque applied to the lever. These signals are to be recorded on computer using custom software. Recording will begin once the motion of the lever has settled to a regular pattern of oscillation. Samples are to be of a minimum of ten cycles of oscillation.

The software used presents the data as graphs of displacement, velocity and applied torque against time. Accurate values for any point on the graphs may be produced. This information will be used to calculate the mean cycle length of oscillation (in seconds) for each sample. From this the frequency of oscillation may be derived.

It is possible, using the formula noted earlier (1.3.2.2), to calculate a raw stiffness value from the frequency of oscillation. However, it is argued (Caligiuri & Galasko, 1992; Webster, 1966) that an activation ratio is a more valid method of quantifying

parkinsonian rigidity. The ratio represents the effect of activation on stiffness and is calculated from the ratio of activated stiffness to resting stiffness. Because the ratio is derived by comparing data from one condition to another within the same patient, the effect of inertia on stiffness is cancelled out. This method is simpler for the investigator and less intrusive for the patient.

The positive feedback method has great advantages over other methods of rigidity quantification. Principally, the use of torque as the independent variable avoids confounding by thixotropy and differential viscosity. These phenomena affect, to some degree, most other instrumental procedures. Further, the relatively small influence of the investigator on the procedure (in other methods the force applied to the limb is produced by the investigator rather than by a motor) is also beneficial in ensuring validity.

2.1.4 Tremor

The instrumentation method selected to quantify tremor is an accelerometer procedure similar to that used in numerous other studies (Arblaster et al., 1993; Caligiuri et al., 1991; Caligiuri & Lohr, 1993; Pullinger & Tyrer, 1983; Tyrer et al., 1981). The accelerometer to be used (model no. ICS 3022-0022-N) is approximately 1cm square and 3mm deep; it is fastened to the finger by a velcro strap. The output signal is proportional to the acceleration undergone. The raw data is to be recorded and stored on computer. Variables representing the amplitude and frequency composition of the tremor may be calculated.

Samples of tremor will be recorded from both hands, of both resting and postural tremor. In both conditions, the patient will be seated with the hand pronated and the forearm supported on a cushioned rest. Cushioning is necessary to ensure against the presence of artefacts due to pulses in the forearm causing movement in the hand. Resting tremor is to be recorded with the hand hanging limp; postural tremor is to be recorded with the hand extended. All samples will be of at least 15 seconds in length, and will be recorded after the hand has settled into position.

Data will be analysed using three 5-second epochs taken from the sample. A Fast Fourier Transform will be performed on the data to determine the relative contribution of different frequency components in the sample to total amplitude. The software to be used presents this information on a frequency spectrum and as a measure of amplitude for each frequency component. Using the three analysed epochs, mean tremor amplitude in two frequency ranges will be calculated These ranges are 3.0-7.0 Hz (low frequency parkinsonian tremor) and 7.0-13.0 Hz (high frequency normal tremor). Comparisons may be made using total tremor amplitude, amplitude in the low and high frequency bands, and the ratio between high and low frequency amplitudes.

The use of accelerometry in tremor instrumentation is now long established and well validated. The same may be said for the use of FFT procedures to analyse the frequency composition of tremor. In this study, the values chosen both for the FFT procedure and the calculation of representative values are well supported by the literature.

The 5-second epoch to be used for FFT analysis was judged sufficient to ensure the validity of the analysis, given the frequency ranges to be assessed. An average of three epochs is to be taken as a precaution against the presence of artefacts in the record. The frequency ranges (3.0-7.0 Hz and 7.0-13.0 Hz) are chosen on the basis of existing work in this field. The literature provides evidence that tremor may be defined as activity occurring above a 3.0 Hz cut-off (Caligiuri et al., 1991). Many authors use a 4.0 Hz cut-off but a lower threshold may be more inclusive of extremely slow tremors. The 13.0 Hz upper limit is also common, it being stated that normal human tremor above this level is unusual (Marsden et al., 1969). The use of a cut-off value of 7.0 Hz between the two frequency ranges is also supported, a similar value being used by other authors (Arblaster et al., 1993; Caligiuri & Lohr, 1993).

2.1.5 Motor Planning Impairment

The indicator of cognitive impairment to be used is the motor planning variable calculated from the CANTAB reaction time test data. Though it is derived from results in the simple and choice reaction time conditions, the impairment is independent of motor response speed and slowed initiation.

The failure of some parkinsonian patients to make use of prior information concerning movement form has been noted earlier (Bloxham et al., 1984; 1.3.1.1.1). This can be illustrated in terms of the reaction time paradigm. In the simple reaction time condition the form of the response to be made is known before the stimulus is

presented, as the response is always the same. In the choice reaction time condition the form of the response cannot be known until the stimulus is presented.

Normal subjects make use of prior of their prior knowledge of response form in the simple condition and pre-plan their response. Their response time is thus shorter in the simple condition than the choice condition. In contrast, patients with parkinsonism often fail to make use of prior knowledge. The effect on their performance is that response times are lengthened to a greater extent in the simple reaction time condition than the choice condition. It is this relative impairment in the simple reaction time condition (and relative lack of impairment in the choice condition) which characterises the motor planning impairment. This characteristic pattern of performance may be exhibited as impairment in the choice condition and greater impairment in the simple condition or as impairment only in the simple condition.

Within the CANTAB reaction time test setting, a variable to indicate the extent of the motor planning latency (a component of the reaction latency) can be derived by subtracting the simple reaction latency from the choice reaction latency. The common components of the two latencies (time to view the stimulus, process the fact of its presence, and to initiate a response) are cancelled out, leaving only the extra time taken to prepare a motor plan in the choice condition. In subjects who do not make use of prior knowledge of the form of response needed this figure should be zero (though in practice it is unlikely to be).

An advantage of this procedure, over other methods of detecting cognitive impairment is that the motor planning variable may be calculated from data already collected. This removes the need for patients to complete any further tests.

2.1.6 Subjective Features – SWN

Two subject-rated scales have been selected. The SWN (Subjective Well-being under Neuroleptics; Naber et al., 1994) is a recently developed scale intended to be sensitive to the negative subjective effects of typical antipsychotics. It was originally developed in German but is now available in an English translation. No evidence has yet been published of the use of this version. Also to be used are three simple visual analogue scales. These form simple assessments of subjective sensations of slowing, sedation, and restlessness.

The SWN has been shown to distinguish the effects of typical antipsychotics from those of atypical agents, and to be predictive of future non-compliance with medication (Naber, 1995). The scale comprises 42 multiple-choice items, with five possible responses for each item. It is intended that the total score of the scale is used, in line with Naber (1995). However the authors also state that factor analysis demonstrates that the items form five sub-factors (emotional regulation, self-control, mental functioning, social integration, physical functioning) and one extra item which does not cluster with any others.

The SWN scale has been selected for the opportunity that it may provide to identify reliably those patients who have negative experiences of typical antipsychotics,

distinguishing them from those who may have a superficially similar appearance due to features of their illness.

2.1.7 Subjective Features – Visual Analogue Scales

The visual analogue scales have been devised specifically for this study. They comprise three 10cm lines, labelled as follows:

Scale	Maximum	Minimum
Slowing	I feel so slowed down it's like carrying a weight on my back	My movements don't feel at all slow
Sedation	I feel so sedated I can't keep my eyes open	I don't feel sedated at all
Restlessness	I feel so fidgety and restless I can't sit still at all	I don't feel fidgety or restless at all

The form of response is to mark a cross on the line at the point felt to most accurately represent the patient's subjective state. The response is scored by measuring the position of the cross from the maximum point on the scale. Thus for all three scales a minimum score (representing normality) rates 10.0 cm and a maximum score (representing extreme abnormality) rates 0.0 cm.

Use of visual analogue scales is well established in many situations in which they provide reliable and valid measures of subjective sensations (criterion validity has been demonstrated relative to objective measures of related features). The benefits of scales such as these derive principally from their ease of use. Visual analogue scales

are quick and simple to administer, and are easily understood by the patient. This rapidity of response may in itself help to increase the validity of the scales by encouraging instinctual responses (a very simple form of response may allow less leeway for artifice).

2.2 Patient Instructions

2.2.1 Jebsen Hand Function Test

In addition to the standard instructions below, patients were encouraged to treat the test as a game which challenged them to complete the tasks as quickly as possible.

2.2.1.1 Card turning

For the left hand: "Place your left hand on the table please. When I say 'Go', use your left hand to turn these cards over one at a time as quickly as you can, beginning with this one (indicate card to extreme right). You may turn them over in any way that you wish and they need not be in a neat pattern when you finish. Do you understand? Ready? Go."

For the right hand: "Now the same thing with the right hand beginning with this one (indicate extreme left card). Ready? Go."

2.2.1.2 Small common objects

For the left hand: "Place your left hand on the table please. When I say 'Go', use your left hand to pick up these objects one at a time and place them in the can as fast as you can beginning with this one (indicate paper clip on the extreme left). Do you understand? Ready? Go."

For the right hand: "Now the same thing with the right hand beginning here (indicate paper clip now on the extreme right). Ready? Go."

2.2.1.3 Simulated feeding

For the left hand: "Take the teaspoon in your left hand please. When I say 'Go', use your left hand to pick up these beans one at a time with the teaspoon and place them in the can as fast as you can beginning with this one (indicate bean on the extreme left). Do you understand? Ready? Go."

For the right hand: "Now the same thing with the right hand beginning here (indicate bean on the extreme right). Ready? Go."

2.2.1.4 Checkers

For the left hand: "Place your left hand on the table please. When I say 'Go', use your left hand to stack these checkers on the board in front of you as fast as you can like this, one on top of the other (demonstrate). You may begin with any checker. Do you understand? Ready? Go."

For the right hand: "Now the same thing with the right hand. Ready? Go."

2.2.1.5 Large light objects

For the left hand: "Place your left hand on the table please. When I say 'Go', use your left hand to stand these cans on the board in front of you, like this (demonstrate). Begin with this one (indicate can on extreme left). Do you understand? Ready? Go."

For the right hand: "Now the same thing with the right hand beginning here (indicate extreme right can). Ready? Go."

2.2.1.6 Large heavy objects

For the left hand: "Now do the same thing with these heavier cans. Place your left hand on the table. When I say 'Go', use your left hand to stand the cans on the board in front of you, like this. Begin with this one (indicate can on extreme left). Do you understand? Ready? Go."

For the right hand: "Now the same thing with the right hand beginning here (indicate extreme right can). Ready? Go."

2.2.2 CANTAB

In addition to the verbal instructions below, diagrams of the stimuli were produced. These illustrations of the circle with a dot in the centre and the five circles, could be

shown and explained before the trial commenced, benefiting patients who found the verbal instructions too abstract if given before the start of the test and too quick if given after the test had started.

2.2.2.1 Pointing to the circle

"A yellow spot will appear inside the circle. Touch the circle as soon as you can after the spot appears."

If the subject points TOO SOON prompt "Try to wait until the spot appears"

If the subject points TOO LATE prompt "Try and point a little quicker".

2.2.2.2 Five choice pointing

"Now we are going on to pointing with five choices. The spot may appear in any of the five circles. Point to the circle where you saw the spot appear. Point as soon as you can after you see the spot."

2.2.2.3 Single choice release

"The yellow spot will appear inside the circle soon after you press the pad. Let go of the pad as soon as you can after you see the spot. Don't let go of the pad until after you see the spot appear."

Prompts: "Try no to let go of the pad until after you see the spot."

2.2.2.4 Single choice release and point

"From now on, when you let go of the pad, touch the circle as soon as you can. Remember not to let go of the pad before you see the spot, but this time remember you have to touch the circle."

2.2.2.5 Five choice release and point

"Now we are going to give you five circles again. Remember, touch the circle where you saw the spot as soon as you can, but don't let go of the pad until you see the spot."

A slight modification was made to the instructions to read "...touch the 'centre' of the circle..." following the case of one patient who interpreted the instruction to "touch the circle" in an over-literal fashion, touching the line forming the perimeter of the circle rather than its centre.

2.2.3 Rigidity

The rigidity instrumentation required little to no conscious effort on the part of the patient. Instructions for this procedure were not formalised. The patient was simply asked to relax as fully as possible and let the machine do the work of moving their arm rather than trying to 'help' it.

2.2.4 Tremor

Like the rigidity assessment procedure, this form of instrumentation demanded little effort from the patient. For each section the hand position needed was explained and if necessary demonstrated. The patient was asked to hold their hand in the position demonstrated, either extended (in the resting tremor assessment) or hanging limp (in the postural tremor assessment), with in both cases the forearm supported and the hand pronated.

2.3 Criteria

The criteria to be used in the evaluation of the instrumental measures are all commonly used and well-established rating scales. Most are observer-ratings, one is a self-report scale.

2.3.1 Extrapyramidal Symptom Rating Scale

Designed using extensive input from neurology, the Extrapyramidal Symptom Rating Scale (ESRS; Chouinard et al., 1980) is a highly comprehensive rating of extrapyramidal symptomatology with very good psychometric properties. The parkinsonism section of this scale is a comprehensive assessment of parkinsonian features. Separate items for the severity of all major signs are included, including bradykinesia, and rigidity and tremor by different body areas. All items are rated on a scale of 0-6 (generally "normal" to "extremely severe" though extensive rating guidelines are provided).

The ESRS has advantages over other commonly used scales (e.g. Simpson & Angus, 1970). The tremor items use a novel feature, the score being influenced on a dual-axis basis by the extent to which tremor is present as well as the amplitude of the tremor. Bradykinesia is made distinct from gait problems (and facial masking), which may in part result from bradykinesia but should not be used as the sole indicator of bradykinesia. Further, similar weight is given to all the major features of parkinsonism (bradykinesia, rigidity, tremor), and correspondingly less weight is placed upon minor features such as facial masking.

Analysis may use a total of all items, a group of bradykinesia-related items (nos. 1, 2, and 4), and some individual items (tremor and rigidity in the arms, bradykinesia, and akathisia).

2.3.2 Abnormal Involuntary Movement Scale

The Abnormal Involuntary Movement Scale (AIMS; Guy, 1976) is an observer-rated assessment of dyskinesias. Items are included for the severity of movements in a number of different body areas (grouped in "facial and oral", "extremities", and "truncal" categories). Further items provide global scores for the overall severity of abnormal movements, incapacitation due to abnormal movements, and patient's awareness of abnormal movements. All items are rated on a 0-4 basis, these scores representing "normal", "minimal", "mild", "moderate", and "severe" degrees of dysfunction. The three global impression items and a total of all movement items may be used in analysis.

2.3.3 Targeting Abnormal Kinetic Effects scale

An observer-rated assessment of parkinsonian side effects (TAKE; Wojcik et al., 1980), designed as a companion to the AIMS scale The format of the TAKE parallels that of the AIMS, consisting of five items for individual features (bradykinesia, rigidity, tremor, autonomic nervous system (ANS) effects, akathisia) and three global scores for overall severity of side effects, incapacitation due to side effects, and patient's awareness of side effects. Like the AIMS, all items are rated on a 0-4 basis,

these scores representing "normal", "minimal", "mild", "moderate", and "severe" degrees of dysfunction. Analysis may be conducted using all individual item scores and/or a total of items 1-5.

2.3.4 Positive And Negative Syndrome Scale

This scale is an observer-rating assessment of psychiatric symptomatology (PANSS; Kay et al., 1987), most commonly used for schizophrenic patients. Items fall into three sub-scales (positive and negative symptomatology, and general psychopathology). Analysis may use the total overall score of all three sub-scales, the totals for the three sub-scales, a composite score (negative symptomatology – positive symptomatology), and some individual items.

2.3.5 Montgomery-Asberg Depression Rating Scale

The Montgomery-Asberg Depression Rating Scale (MADRS; Montgomery & Asberg, 1979) is an observer-scored rating of depression. The MADRS was derived from a comprehensive observer-rating of psychopathology, the Comprehensive Psychopathological Rating Scale (CPRS; Asberg et al., 1978), by selecting the items most sensitive to changes in depressive symptomatology.

When analysing the results of the MADRS, total score is most commonly used. Two groups of items may also be used in this study. The first comprises items which may overlap descriptively with features of extrapyramidal side effects. This "EPS-like" group of items allows the investigation of possible conceptual contamination; it

comprises items 3, 6, 7, and 8. The second group is comprised of all other items (nos. 1, 2, 4, 5, 9, and 10).

2.3.6 Beck Depression Inventory

The Beck Depression Inventory (BDI; Beck et al., 1961) is a very commonly used self-report scale for depression which has been well-validated in psychiatric populations. The BDI comprises 21 items which are rated on a 0-3 basis, the patient being asked to choose which of four statements best describes the way they have been feeling during the previous week. A simple total score will be used for analysis.

2.4 Statistical analysis procedures

2.4.1 ANOVA

Analysis of variance and associated post-hoc tests (Bonferroni and LSD) are used for the comparison of means of different groups where the number of groups is three or more. ANOVA tests to determine if the different experimental groups are likely to represent samples from the same population. Post-hoc tests to perform multiple comparisons allow the location of significant differences to be determined, i.e. which groups differ significantly from each other. The Bonferroni post-hoc test is the more conservative of the two. A significance level of $p < 0.05$ was considered significant and a level of $p < 0.01$ very significant.

2.4.2 T-test

Student's t-test likewise calculates the likelihood that two sample means tested represent the same population. Levene's test for equality of variances was performed with all t-tests, a probability of $p < 0.05$ indicated that the assumption of equality of variance had been violated. As with ANOVA, a significance level of $p < 0.05$ was considered significant and a level of $p < 0.01$ very significant.

2.4.3 Non-parametric ANOVA

The Kruskal-Wallis test is a distribution-free form of analysis of variance, i.e. it may be used for data which cannot conform to a normal population curve (when data values must be integers). As with the other tests for group differences, a significance

level of $p < 0.05$ was considered significant and a level of $p < 0.01$ very significant. Lieberman's index analysis was used as a non-parametric post-hoc test to determine the location of significant between-group differences.

2.4.4 Parametric correlation co-efficient

The Pearson's correlation co-efficient is calculated as an indicator of the strength and significance of a relationship between two variables. Relationships where $r > +0.8$ or $r < -0.8$ can be considered strong, those where r is between -0.5 and -0.8 or between 0.5 and 0.8 are classed as being of moderate strength, and those where r is between -0.5 and 0.5 are classed as weak. As with the tests of differences, a significance level of $p < 0.05$ was considered significant and a level of $p < 0.01$ very significant.

2.4.5 Nonparametric correlation co-efficient

The Spearman's ranked correlation co-efficient is a non-parametric equivalent of the Pearson's measure. This coefficient will identify non-linear relationships which may be missed by the Pearson coefficient, and may also be used with non-continuous variables. As with the Pearson coefficient, relationships where $r > +0.8$ or $r < -0.8$ can be considered strong, those where r is between -0.5 and -0.8 or between 0.5 and 0.8 are classed as being of moderate strength, and those where r is between -0.5 and 0.5 are classed as weak. Once again, a significance level of $p < 0.05$ was considered significant and a level of $p < 0.01$ very significant.

2.5 Demographic characteristics

The following tables contain the demographic data for the patient group and the control group. The information covers age, gender, diagnosis, and medication status.

2.5.1 Patient group case characteristics

Case	Age	Gender	Diagnosis	Medication	Code
AD	28	M	Schizophrenia	Depixol 80mg weekly	2
AG	50	F	Schizophrenia	Risperidone 4mg nocte Changed to Quetiapine 100 + 150	4
AH	30	M	Depression	Chlorpromazine100 OD Procyclidine 5mg TD Lithium 1000mg	3
BC	52	M	Bipolar Disorder	Lithium 1000 nocte Sertraline 200 OD Thyroxine 150 ug Temazepam 20	7
BW	47	M	Bipolar Disorder	Lithium 1000mg Thioridazine 25 mané + 100 nocte	3
CM	48	F	Depression	Nefazodone 200mg BD Mirtazapine 15-30mg	7
CMC	29	M	Depression	Thioridazine 75 nocte Temazepam 20 nocte Lithium 1000mg nocte	3
CR	36	F	Schizophrenia	Clozapine 500mg (200 OD + 300 OD) Chlorpromazine 50mg OD	5

DC	52	F	Depression	Imipramine 175mg nocte	6
DCL	23	M	Schizophrenia	Clozapine 750 OD Hyoscine	5
ED	40	F	Depression	Sulpiride 800 BD Procyclidine 5 mané Chlorpromazine 50 nocte	3
ET	20	F	Schizophrenia	Trimipramine 150 nocte Stelazine 30 nocte Droperidol 10 BD Procyclidine 5 TD Pindolol	2
GC	26	M	Schizophrenia	Clozapine 300 mg OD Paroxetine 20 OD	5
GG	20	M	Schizophrenia	Amitriptyline 200 mg nocte	6
GN	42	F	Bipolar Disorder	Olanzapine 10 mg Lithium 300 mg	4
HH	34	F	Bipolar Disorder	Amitriptyline 125mg nocte Carbamazepine SR 400mg nocte Lithium 800mg nocte Pericyazine 80mg nocte	2
HM	41	F	Schizophrenia	Clopixol 400 2/52 Fluoxetine 40 mg OD Temazepam 20 nocte Orphenadrine 50 TD	3
JA	38	M	Bipolar Disorder	Lithium 450mg BD Haloperidol 1mg BD Tegmetol SR 600mg	2
JG	38	F	Depression	Chlorpromazine 75-100mg Paroxetine 30mg	3
JH	36	F	Bipolar Disorder	Lithium 100-mg Haloperidol 5mg Procyclidine 2.5mg	2

				Thyroxine 125 ug	
JK	45	F	Depression	Amitriptyline 125mg nocte	6
JM	33	M	Schizophrenia	Clozapine 200 nocte	5
JMK	48	M	Depression	Seroxat Trazadone 100 nocte	7
JR	58	M	Bipolar Disorder	Haloperidol 5 BD	1
JS	28	M	OCD	Olanzapine 20mg	4
LB	46	F	Bipolar Disorder	Lithium 1200 Droperidol PRN Procyclidine 5 TD	1
LK	30	M	Schizophrenia	Chlorpromazine 100 BD Procyclidine 5 BD Depixol 60mg 2/52	2
LS	32	F	Schizophrenia	Olanzapine 15mg	4
MB	40	M	Schizophrenia	Depixol 30 mg 3/52	2
MD	22	M	Schizophrenia	Trifluoperazine 25mg	2
MH	59	F	Depression	Lithium 600mg nocte Imipramine 100mg nocte	6
MT	17	M	Schizophrenia	Trifluoperazine 20mg Procyclidine 5 BD	2
MWB	41	M	Schizophrenia	Clozapine 750mg OD	5
PC	38	M	Bipolar Disorder	Lithium 1200mg Carbamazepine 500mg	1
PMG	53	M	Schizophrenia	Clozapine 350 mg OD	5
PS	27	F	Schizophrenia	Clozapine 350 mg OD Imipramine 10mg BD Hyoscine 300ug BD	5
RF	31	F	Schizophrenia	Chlorpromazine 50mg (for 5 days only)	1
SM	24	M	Schizophrenia	Clopixol 200mg weekly	2

				Zopiclone 7.5	
SP	22	M	Schizophrenia	Risperidone 2mg OD	4
SPA	25	M	Schizophrenia	Clozapine 325mg OD Seroxat 30 OD	5
TIH	33	M	Schizophrenia	Clopixol 100 weekly Chlorpromazine 100 nocte	2
WM	36	M	Schizophrenia	Depixol 50 mg weekly Pimozide 15mg OD Clomipramine Procyclidine 5 PRN	2

The following codes were used to describe medication status for defining patient medication groups:

0 – Control

1 – Minimal exposure to psychotropic medication

2 – Typical antipsychotics (high potency)

3 – Typical antipsychotics (low potency)

4 – Atypical antipsychotics (other than clozapine)

5 – Atypical antipsychotics (clozapine)

6 – Antidepressants (tri-cyclic)

7 – Antidepressants (SSRI)

2.5.2 Control group case characteristics

Name	Age	Gender
AC	39	F
AHC	30	M
CH	36	F

DW	28	M
EA	26	F
EM	34	M
FC	24	F
GD	39	F
GT	45	M
HB	33	F
HP	24	F
J	N/A	M
JV	33	M
KD	29	M
MS	33	F
SC	28	M
VB	25	F

2.5.3 Group characteristics

This section will present and compare the demographic characteristics of the control group and the different patient groups identified. The analysis will first consider the control group and the overall patient group, then consider the patient medication groups which will be used.

2.5.3.1 Group numbers in the patient and control groups

Group	Number
Patient	42
Control	17

2.5.3.2 Gender composition of the patient and control groups

Group	Number		
	Male	Female	Total
Patient	25	17	42
Control	8	9	17

2.5.3.3 Age composition of the patient and control groups

Group	Age		
	Minimum	Maximum	Mean
Patient	17	59	36.14
Control	24	45	31.625

2.5.3.4 Patient medication group numbers

The first sub-division of the patient group is into those receiving antipsychotic medication and those not.

Group	Number
Antipsychotics	31
Typical	*18*
Atypical	*6*

Non-Antipsychotics	11
Antidepressants	*7*
Minimal exposure	*4*

2.5.3.5 Gender composition of patient medication groups

Group	Number		
	Male	Female	Total
Antipsychotic	20	11	31
Typical	*12*	*6*	*18*
Atypical	*8*	*5*	*13*
Non-antipsychotic		6	11
Antidepressants	5	*4*	*7*
Minimal exposure	*3*	*2*	*4*
	2		

2.5.3.6 Age composition of patient medication groups

Group	Age		
	Minimum	Maximum	Mean
Antipsychotic	17	53	32.94
Typical	*17*	*47*	*32.39*
Atypical	*22*	*53*	*33.69*
Non-antipsychotic	20	59	45.18
Antidepressants	*20*	*59*	*46.29*
Minimal exposure	*31*	*58*	*43.25*

2.5.3.7 Antipsychotic medication group numbers

The antipsychotic group may be further sub-divided by separating the typical antipsychotics group into those who are receiving high- or low-potency antipsychotic agents and by separating the atypical antipsychotics group into those who are receiving clozapine and those who are receiving other agents. The resulting sub-groups are small in number; other demographic data is not presented for these groups.

Group	Number
Typical antipsychotics	18
High potency	*12*
Low potency	*6*
Atypical antipsychotics	13
Clozapine	*8*
Other agents	*5*

2.5.4 Group comparisons

2.5.4.1 Patient and control group

No difference in gender composition (Chi sq. = 0.764, df = 1). Similarly no differences in age (Kruskal Wallis test, Chi sq. = 1.8, df =1). The control group is very similar to the overall patient group in terms of gender composition and average age.

2.5.4.2 Patient medication groups

Differences are found in gender composition (Chi sq. = 15.434, df = 3, $p < 0.01$) and in age (Kruskal Wallis test, Chi sq. = 13.612, df = 6, $p < 0.05$). Though the overall patient group is similar to the control group, the age and gender distribution in the patient group is not consistent within the different medication groups.

2.5.5 Multiple drug therapy

Data were also collected describing the prevalence of multiple drug therapy in this patient cohort. The following table illustrates the prevalence of multiple drug therapy in some of the patient medication groups participating in the study. The distinction between high and low potency typical antipsychotic agents is not used in most comparisons but is useful here.

No. of medications	Typical antipsychotic (high potency)	Typical antipsychotic (low potency)	Clozapine	Antidepressant
1	17%	-	50%	43%
2	25%	33%	50%	43%
3	25%	50%	-	14%
>3	33%	17%	-	-

Supplementary medications in the typical anti-psychotic groups include additional typical antipsychotics, lithium carbonate, tricyclic antidepressants, SSRIs, carbamazepine, and anti-parkinsonian agents (procyclidine). Even in the clozapine

group, half of the group were receiving supplementary medication including SSRIs, other antidepressants, and even a small regular dose of a typical antipsychotic.

2.5.6 Summary

The control group is a suitable comparison group for the patient group in this study, being non-significantly different in gender and age composition. Though there are differences between the patient medication groups these likely reflect differences in the diagnostic composition of the medication groups. Multiple drug therapy is common in all the patient medication groups. Though this prescribing practice is undesirable in terms of research methodology it is a common reality of clinical practice.

3 Results

This section contains the evaluation of the measures used in this study and, following this, a consideration of any other significant results.

3.1 Bradykinesia – Jebsen hand function test

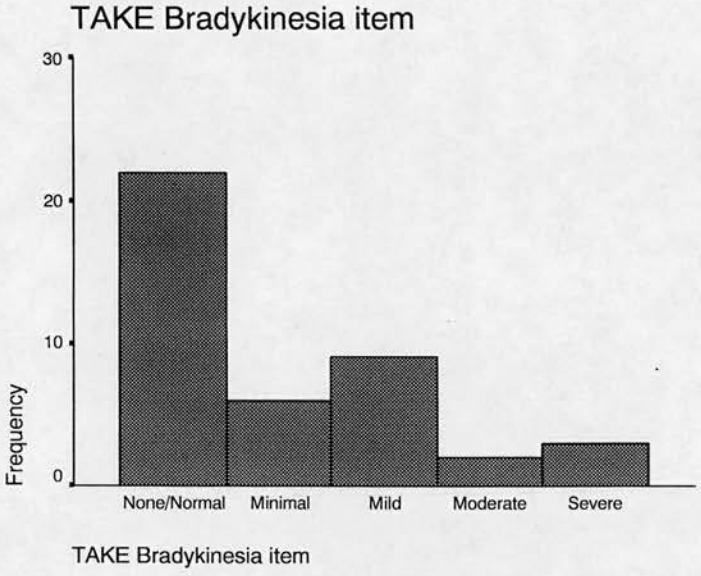
3.1.1 Observer ratings of bradykinesia

Observer ratings of bradykinesia are taken from the TAKE and ESRS scales. The TAKE item is a global assessment of bradykinesia, including manifestations of bradykinesia such as gait difficulties, lack of pendular arm swing and facial masking. The ESRS bradykinesia item is an assessment of "pure" bradykinesia, comprising speed of movement and initiation difficulties. Also used is the group of bradykinesia-related items from the ESRS scale.

3.1.1.1 Descriptive statistics for the TAKE bradykinesia item scores in the patient group only

	N	Minimum	Maximum	Median
TAKE bradykinesia item	42	0	4	0

3.1.1.2 Bar-graph of TAKE bradykinesia item score frequencies for the patient group



3.1.1.3 TAKE bradykinesia item score frequencies for the patient group only

	Frequency	Percent
None/Normal	22	52.4
Minimal	6	14.3
Mild	9	21.4
Moderate	2	4.8
Severe	3	7.1
Total	42	100.0

3.1.1.4 Summary of TAKE bradykinesia item scores

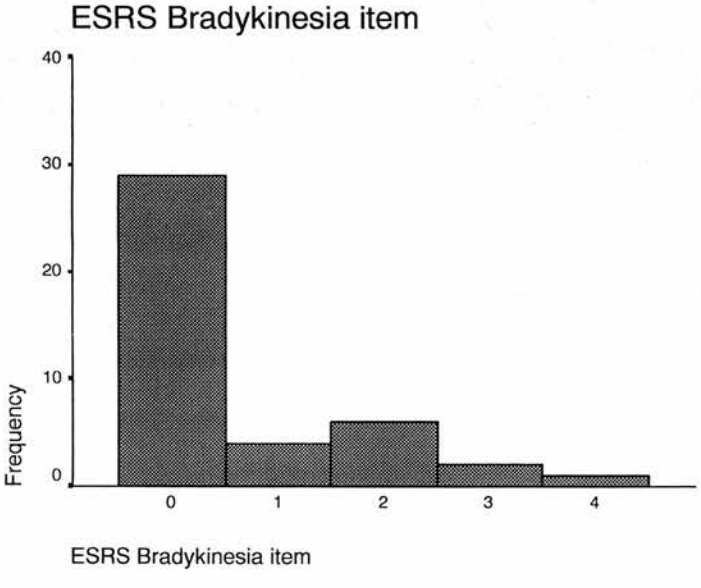
Though a small number of patients were rated as exhibiting severe levels of bradykinesia, the vast majority exhibited minimal or no bradykinesia. Over 50% of patients were rated normal, and two thirds were rated as normal or minimal

bradykinesia. The median score was zero. The scores do not exhibit a normal distribution.

3.1.1.5 Descriptive statistics for the ESRS bradykinesia item scores in the patient group

	N	Minimum	Maximum	Median
ESRS bradykinesia item	42	0	4	0

3.1.1.6 Bar-graph of ESRS bradykinesia item score frequencies for the patient group



3.1.1.7 ESRS bradykinesia item score frequencies for patient group

	Frequency	Percent

0	29	69.0
1	4	9.5
2	6	14.3
3	2	4.8
4	1	2.4
Total	42	100.0

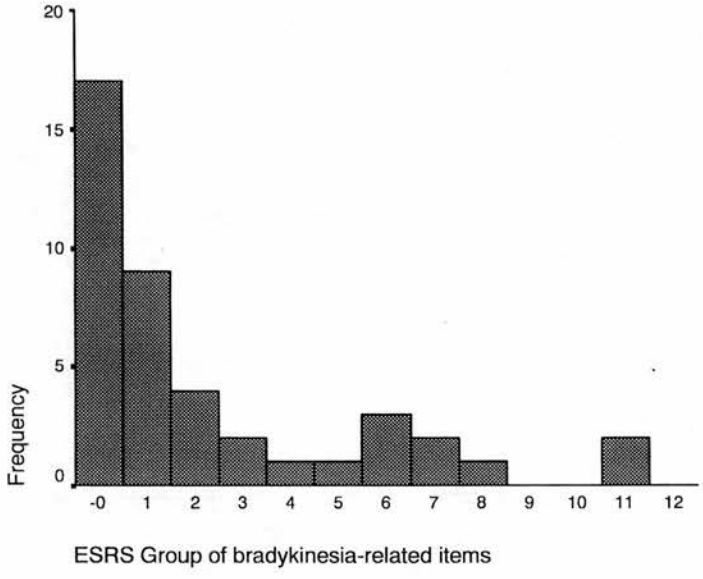
3.1.1.8 Summary of ESRS bradykinesia item scores for the patient group

The ratings made using the ESRS bradykinesia item exhibit an even greater preponderance of low scores than do those made using the TAKE bradykinesia item. Over two thirds (69%) of patients were rated as exhibiting no bradykinesia on this item, and the highest score used was four (of a maximum six). The median score was zero again. The scores on this item display a similar non-normal distribution pattern to those on the TAKE bradykinesia item.

3.1.1.9 Descriptive statistics for the group of ESRS bradykinesia-related items scores in the patient group only

	N	Minimum	Maximum	Median
Bradykinesia-related items group from ESRS	42	0	11	1

3.1.1.10 Bar-graph of score frequencies for the group of ESRS bradykinesia-related items



3.1.1.11 Score frequencies for group of ESRS bradykinesia-related items scores for patient group

Score	Frequency	Percent
0	17	40.5
1	9	21.4
2	4	9.5
3	2	4.8
4	1	2.4
5	1	2.4
6	3	7.1
7	2	4.8
8	1	2.4
11	2	4.8
Total	42	100.0

3.1.1.12 Summary of ESRS group of bradykinesia-related items scores

As with the other ratings of bradykinesia, scores on this criterion were mostly low. The maximum rating given was 11, and the median was only 1. The distribution of scores is also similar to the other bradykinesia criteria.

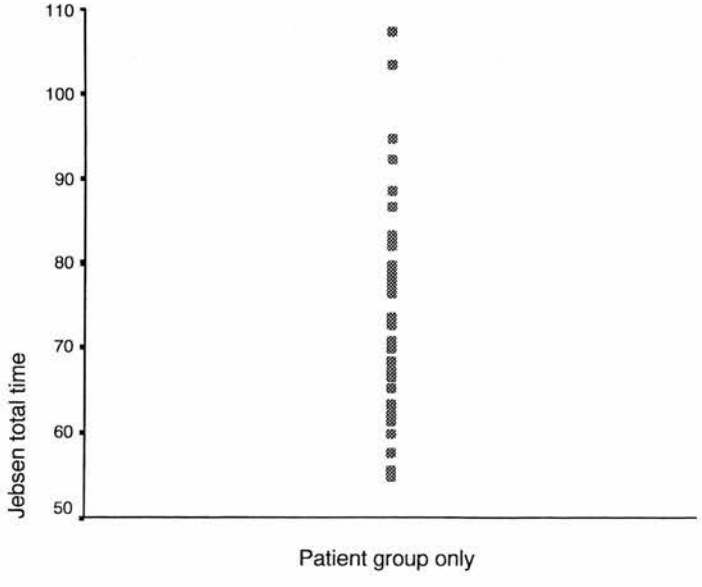
3.1.2 Instrumentation

The results of the Jebsen Hand Function Test are presented here for the patient group only.

3.1.2.1 Jebsen hand function test total time descriptive data from the patient group

	N	Minimum	Maximum	Mean	SD
Jebsen total time	40	54.76	107.37	74.9667	12.5430

3.1.2.2 Scatter plot of Jebsen test total time for patient group only



3.1.2.3 Summary of the Jebsen hand function test results

The results from the Jebsen test exhibit significant clustering towards the lower end of the total time range. The lack of a normal distribution parallels the pattern of scores found in the observer-rating criteria.

3.1.3 Evaluation of the Jebsen hand function test as a measure of bradykinesia relative to the TAKE bradykinesia item criterion

3.1.3.1 Correlations between TAKE bradykinesia item and Jebsen total time

N	Coefficient	Significance
40	0.325	0.020

Scatter plots of the relationship between Jebsen total time and TAKE bradykinesia item are displayed in section 3.1.3.6.

3.1.3.2 Identification of a bradykinetic group using the observer-rating criterion

The inclusion threshold was set at a score of 2 or greater.

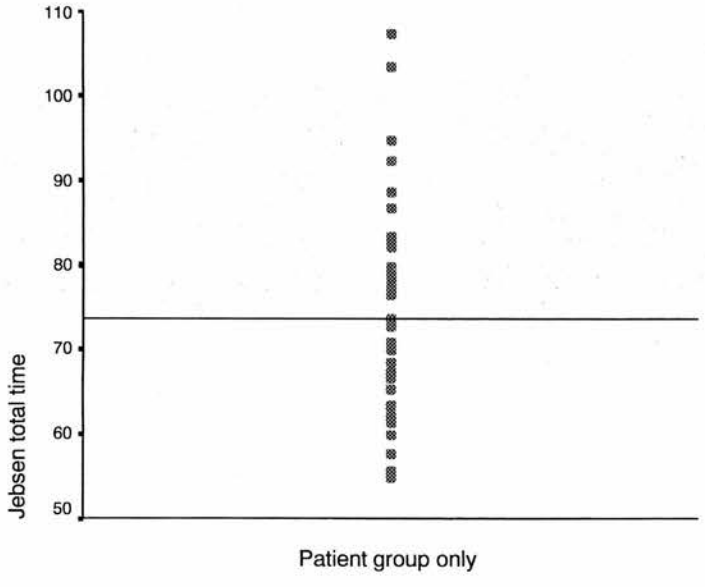
	Frequency	Percent
Normal/Minimal bradykinesia	28	66.7
Bradykinetic	14	33.3
Total	42	100.0

3.1.3.3 Identification of a bradykinetic group using Jebsen test

The inclusion/exclusion threshold for the Jebsen test groups was set at 2SDs above the mean of the control group.

	N	Minimum	Maximum	Mean	SD	Mean +2SD
Jebsen total time	17	45.15	75.53	60.3871	6.6229	73.6329

3.1.3.4 Scatterplot of Jebsen total time results in the patient group with group threshold (2SDs above mean of control group) marked

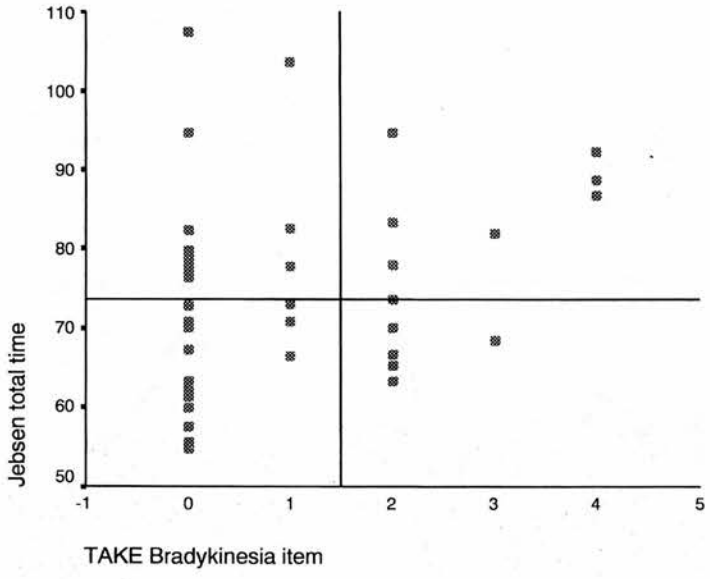


3.1.3.5 Frequency data for Jebsen test identified groups

Jebsen total time	Frequency	Percent
<2SDs above control group mean	21	50.0
>2SDs above control	19	47.5

group mean Total	40	100.0

3.1.3.6 Scatterplot of Jebsen total time against TAKE bradykinesia item score with group thresholds marked



3.1.3.7 Crosstabs procedure results

Measure: *Positive*	12	7
Negative	15	6
	Negative Criterion	*Positive*

Sensitivity = 53.8%

Specificity = 36.8%

3.1.3.8 Summary of the evaluation of the Jebsen hand function test using the TAKE bradykinesia item criterion

The correlation between the Jebsen test results and the TAKE bradykinesia item criterion is of only weak to moderate strength but is statistically significant ($p = 0.020$). The instrumentation group inclusion threshold at 2SDs above the mean of the control group provided comparison groups of roughly equal sizes (though the group identified as bradykinetic by the Jebsen test was slightly larger than that identified by the TAKE bradykinesia item criterion). The majority of patients were rated as non-bradykinetic by both the observer-rating and the Jebsen test. In this evaluation the sensitivity of the Jebsen test was reasonably good though the specificity was poor. The poor specificity resulted from the relatively high number of false positive cases identified by the Jebsen test.

3.1.4 Evaluation of the Jebsen hand function test as a measure of bradykinesia using the ESRS bradykinesia item criterion

3.1.4.1 Correlation between Jebsen total time and the ESRS bradykinesia item (one-tailed significance)

N	Coefficient	Significance
40	0.233	0.074

A scatterplot of Jebsen total time against ESRS bradykinesia item may be seen below in section 3.1.4.4.

3.1.4.2 Identification of a bradykinetic group using the ESRS observer-rating criterion

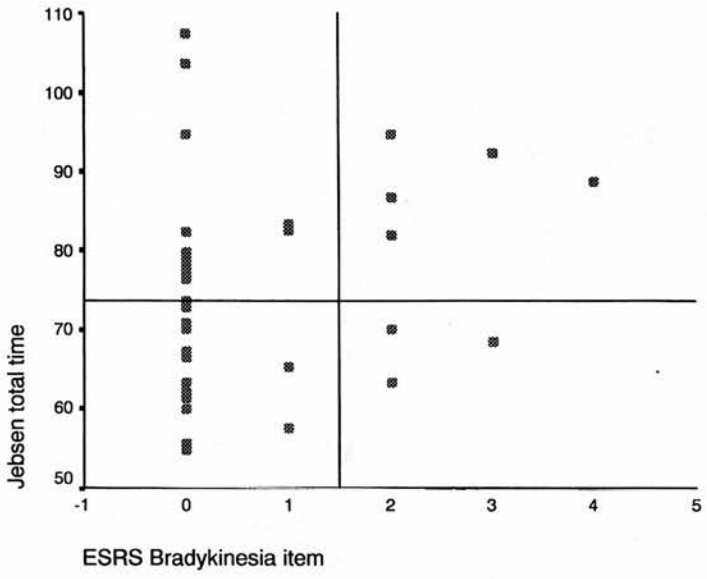
The inclusion threshold was set at a score of 2 or greater.

	Frequency	Percent
Normal/Minimal bradykinesia	33	78.6
Bradykinetic	9	21.4
Total	42	100.0

3.1.4.3 Instrumentation groups

The same Jebsen test identified groups were used as in the previous comparisons (3.1.3.3).

3.1.4.4 Scatterplot of Jebsen total time against ESRS bradykinesia item score with group thresholds marked.



3.1.4.5 Crosstabs procedure results

Measure: *Positive*	14	5
Negative	18	3
	Negative Criterion	*Positive*

Sensitivity = 62.5%

Specificity = 26.3%

3.1.4.6 Summary of the evaluation of the Jebsen hand function test using the ESRS bradykinesia item criterion

The correlation between the Jebsen test and the ESRS bradykinesia item criterion is weak and only trends towards significance. The groups identified by the measure and the criterion are less similar in size than in the previous comparison, due to the small numbers of patients identified as bradykinetic by the ESRS bradykinesia item criterion. As before, both the measure and the criterion identify the majority of patients as non-bradykinetic. Relative to this criterion, the sensitivity of the measure is good, but specificity is low.

3.1.5 Evaluation of the Jebsen hand function test using the group of ESRS bradykinesia-related items criterion

3.1.5.1 Correlation between Jebsen total time and score on group of ESRS bradykinesia-related items

One-tailed significance is quoted.

N	Coefficient	Significance
40	0.361	0.011

Scatterplot illustrated below (3.1.5.4)

3.1.5.2 Identification of a bradykinetic group using the observer-rating criterion

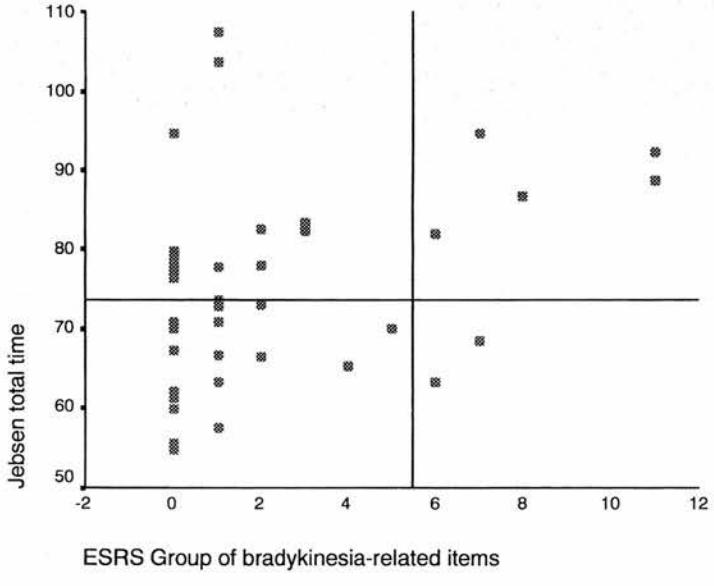
The inclusion threshold was set at a score of 2 or greater.

	Frequency	Percent
Normal/Minimal bradykinesia	34	81.0
Bradykinetic	8	19.0
Total	42	100.0

3.1.5.3 Instrumentation groups

The same Jebsen-identified group were used as in previous comparisons (3.1.3.3).

3.1.5.4 Scatterplot of Jebsen total time against score from group of ESRS bradykinesia-related items



3.1.5.5 Crosstabs procedure results

Measure:	14	5

Positive		
Negative	19	2
	Negative Criterion	*Positive*

Sensitivity = 71.4%

Specificity = 26.3%

3.1.5.6 Summary of the evaluation of the Jebsen hand function test using the ESRS group of bradykinesia-related items criterion

The strongest association between the Jebsen test and an observer-rating was found with this criterion. It is of moderate strength and a high level of significance. However, the observer-rating and Jebsen test identified groups are not very equal in size, though the majority of patients are identified as non-bradykinetic by both the measure and the criterion. Although the sensitivity of the measure is very high specificity is low.

3.1.6 Summary

Observer-ratings of bradykinesia are low in the patient group. Most patients are rated normal, those exhibiting bradykinesia are mostly rated as exhibiting only minimal or mild abnormality. The correlations between the objective measures and the ratings

criteria are mostly weak but statistically significant. The sensitivity of the measures as markers of bradykinesia is moderate to good but the specificity is low.

3.2 Bradykinesia – CANTAB reaction time test

3.2.1 Results of CANTAB instrumentation

Results are presented for the patient group only.

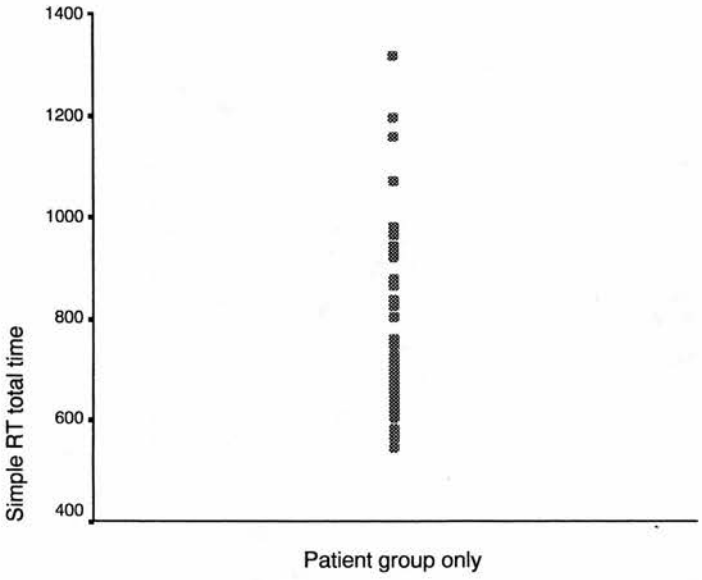
3.2.1.1 CANTAB reaction time test descriptive data for patient group only

	N	Minimum	Maximum	Mean	SD
Simple RT total time	40	547	1317	783.65	180.73
Choice RT total time	40	450	1261	783.07	169.45

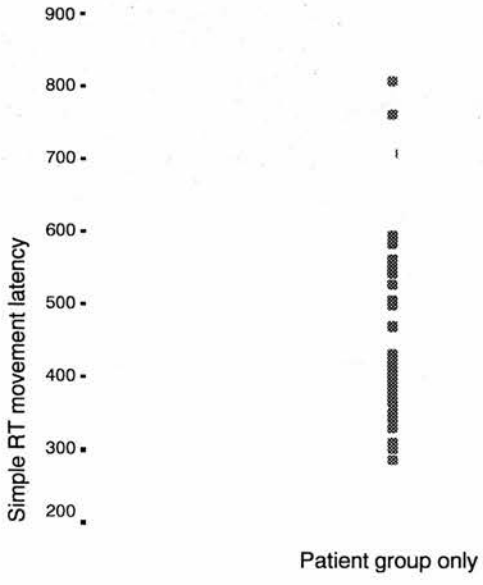
	N	Minimum	Maximum	Mean	SD
Simple RT movement latency	40	285	806	453.95	130.55
Choice RT movement latency	40	294	809	446.80	128.00

3.2.1.2 Scatterplots of CANTAB reaction time test variables for the patient group only

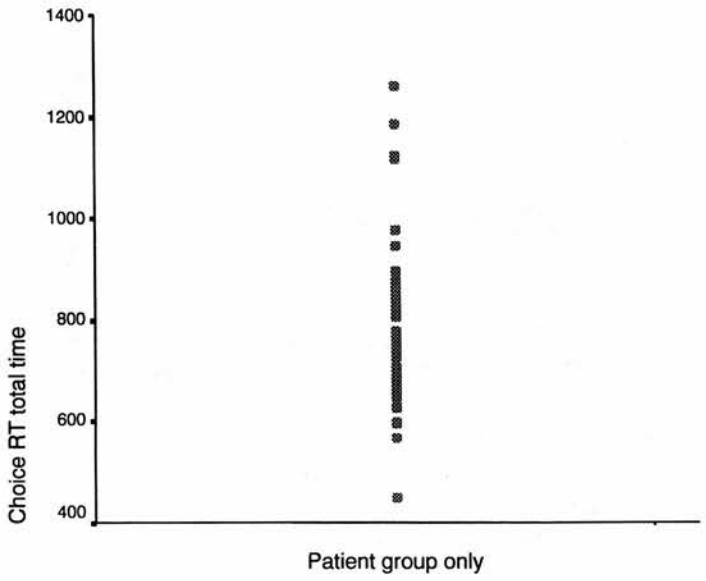
3.2.1.2.1 Simple RT total time



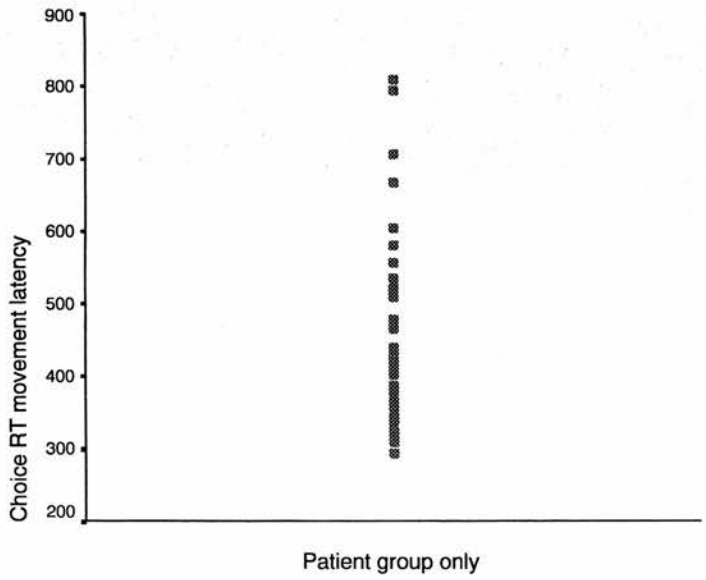
3.2.1.2.2 Simple RT movement latency



3.2.1.2.3 Choice RT total time



3.2.1.2.4 Choice RT movement latency



3.2.1.3 Summary of CANTAB reaction time test results

The pattern of results from the CANTAB reaction time variables is superficially very similar to those from the Jebsen hand function test, and the observer-ratings of

bradykinesia. The majority of patients exhibit relatively low total times and the distribution of times is non-normal.

3.2.2 Evaluation of the CANTAB reaction time test variables as measures of bradykinesia relative to the TAKE bradykinesia item criterion

Total reaction time and movement latency variables from both simple and choice RT conditions are to be used as measures.

3.2.2.1 Correlation between the CANTAB reaction time variables and the TAKE bradykinesia item (one-tailed significance)

Correlation with TAKE bradykinesia item	Simple RT total time	Choice RT total time
Coefficient	0.316	0.217
Significance	0.023	0.089

Correlation with TAKE bradykinesia item	Simple RT movement latency	Choice RT movement latency
Coefficient	0.340	0.317
Significance	0.016	0.023

3.2.2.2 The TAKE bradykinesia item groups were selected using the same inclusion threshold (score of at least 2) as in previous analyses

	Frequency	Percent
Normal/Minimal bradykinesia	28	66.7

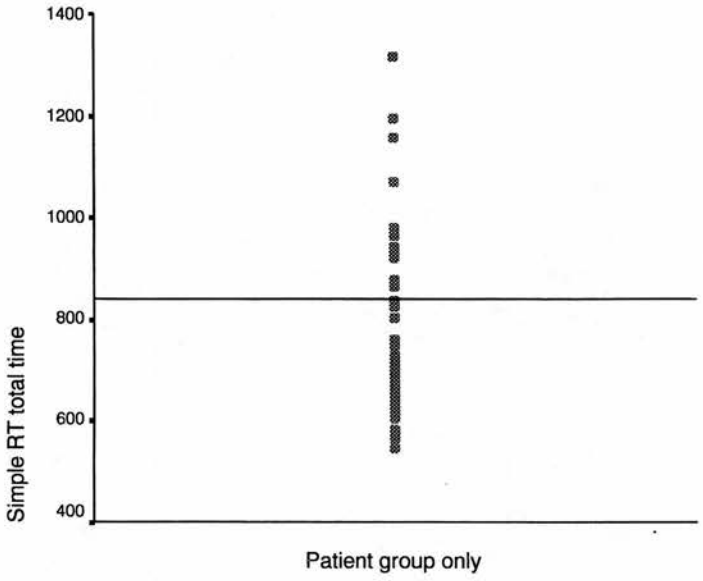
Bradykinetic	14	33.3
Total	42	100.0

3.2.2.3 CANTAB reaction time test groups identified using these data from the control group only (threshold set at 1SD above control group mean)

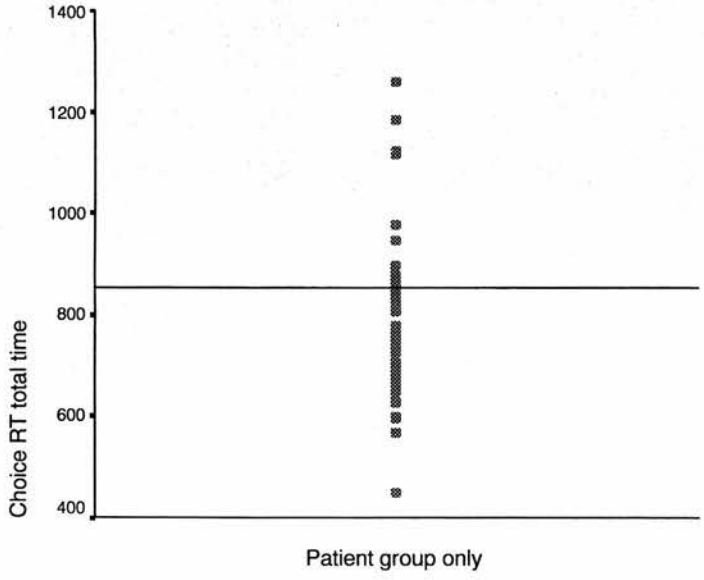
	N	Minimum	Maximum	Mean	SD	Mean +SD
Simple RT total time	17	511	933	710.94	130.97	841.91
Choice RT total time	17	549	939	737.65	116.19	853.84
Simple RT movement latency	17	253	584	414.41	108.60	523.01
Choice RT movement latency	17	290	567	414.82	86.30	501.12

3.2.2.4 Scatterplots of patient group data with group thresholds marked

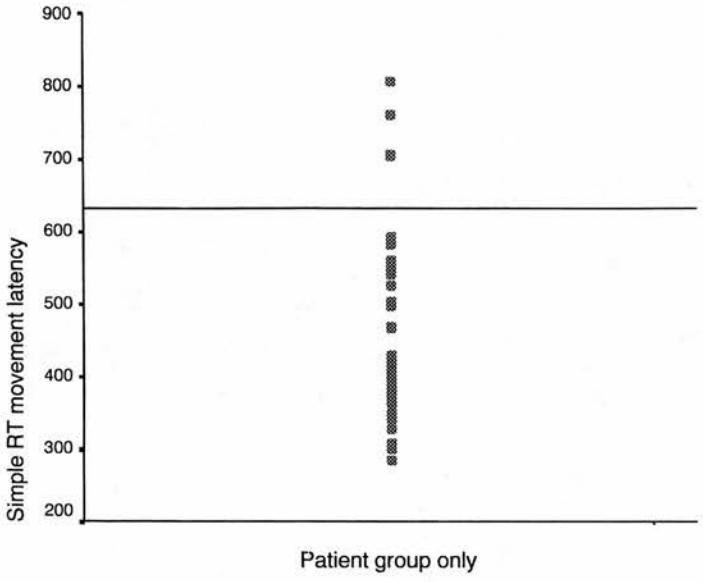
3.2.2.4.1 Simple RT total time



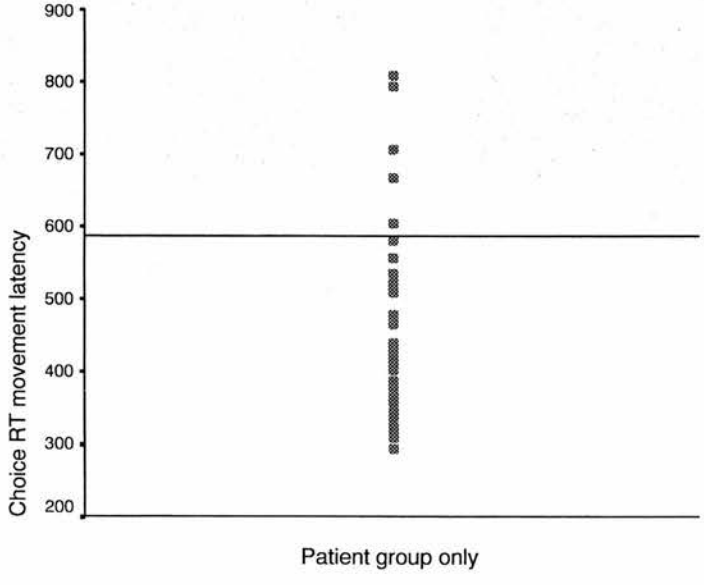
3.2.2.4.2 Choice RT total time



3.2.2.4.3 Simple RT movement latency



3.2.2.4.4 Choice RT movement latency



3.2.2.5 Frequency data for groups identified using CANTAB reaction time test variables

3.2.2.5.1 Simple RT total time

Simple RT total time	Frequency	Percent
<1SD above control group mean	28	70.0
>1SD above control group mean	12	30.0
Total	40	100.0

3.2.2.5.2 Choice RT total time

Choice RT total time	Frequency	Percent
<1SD above control group mean	30	75.0
>1SD above control group mean	10	25.0
Total	40	100.0

3.2.2.5.3 Simple RT movement latency

Simple RT movement latency	Frequency	Percent
<1SD above control group mean	28	70.0
>1SD above control group mean	12	30.0
Total	40	100.0

3.2.2.5.4 Choice RT movement latency

Choice RT movement latency	Frequency	Percent
<1SD above control group mean	29	72.5
>1SD above control group mean	11	27.5
Total	40	100.0

3.2.2.6 Sensitivity and specificity analysis of the CANTAB reaction time test variables using the TAKE bradykinesia item criterion

3.2.2.6.1 CANTAB simple RT total time against TAKE bradykinesia

Measure: *Positive*	5	7
Negative	22	6
	Negative Criterion	*Positive*

Sensitivity = 53.8%

Specificity = 58.3%

3.2.2.6.2 CANTAB choice RT total time against TAKE bradykinesia

Measure: *Positive*	4	6
	23	7

Negative		
	Negative Criterion	*Positive*

Sensitivity = 46.2%

Specificity = 60.0%

3.2.2.6.3 CANTAB simple RT movement latency against TAKE bradykinesia

Measure: *Positive*	5	7
Negative	22	6
	Negative Criterion	*Positive*

Sensitivity = 53.8%

Specificity = 58.3%

3.2.2.6.4 CANTAB choice RT movement latency against TAKE bradykinesia

Measure: *Positive*	5	6
Negative	22	7
	Negative Criterion	*Positive*

Sensitivity = 46.2%

Specificity = 54.5%

3.2.2.7 Summary of the evaluation of the CANTAB reaction time test variables using the TAKE bradykinesia item criterion

The correlation between the criterion and simple reaction time is statistically significant but weak. The same is true of the correlations between the criterion and the movement latencies in both the simple and choice reaction time conditions. However, the association between the criterion and total reaction time in the choice condition is very weak and non-significant. In all four analyses, the groups defined by the measure were of similar size to those identified by the criterion. Sensitivity and specificity were moderate in most cases. However, the two choice reaction time variables, especially total choice reaction time, demonstrated relatively higher specificity and lower sensitivity.

3.2.3 Evaluation of the CANTAB reaction time variables relative to the ESRS bradykinesia item criterion

3.2.3.1 Correlation between the CANTAB reaction time variables and the ESRS bradykinesia item

Correlation with ESRS bradykinesia item	Simple RT total time	Choice RT total time
Coefficient	0.295	0.161
Significance	0.032	0.160

Correlation with ESRS bradykinesia item	Simple RT movement latency	Choice RT movement latency
Coefficient	0.314	0.187
Significance	0.024	0.124

3.2.3.2 Frequency scores for normal and bradykinetic groups identified using the ESRS bradykinesia item (threshold value of 2)

	Frequency	Percent
Normal/Minimal bradykinesia	33	78.6
Bradykinetic	9	21.4
Total	42	100.0

3.2.3.3 Instrumentation groups

Group numbers are as before (3.2.2.3):

Simple RT total: 12 above, 28 below

Choice RT total: 10 above, 30 below

Simple RT movement: 12 above, 28 below

Choice RT movement: 11 above, 29 below

3.2.3.4 Sensitivity and specificity analysis for the evaluation of the CANTAB reaction time variables using the ESRS bradykinesia item criterion

3.2.3.4.1 CANTAB simple RT total time against ESRS bradykinesia item

Measure: *Positive*	8	4
Negative	24	4
	Negative Criterion	*Positive*

Sensitivity = 50.0%

Specificity = 33.3%

3.2.3.4.2 CANTAB choice RT total time against ESRS bradykinesia item

Measure: *Positive*	6	4
Negative	26	4
	Negative Criterion	*Positive*

Sensitivity = 50.0%

Specificity = 40.0%

3.2.3.4.3 CANTAB simple RT movement latency against ESRS bradykinesia item

Measure: *Positive*	8	4
Negative	24	4
	Negative	*Positive*

	Criterion	

Sensitivity = 50.0%

Specificity = 33.3%

3.2.3.4.4 CANTAB choice RT movement latency against ESRS bradykinesia item

Measure: *Positive*	7	4
Negative	25	4
	Negative Criterion	*Positive*

Sensitivity = 50.0%

Specificity = 36.4%

3.2.3.5 Summary of the evaluation of the CANTAB reaction time variables using the ESRS bradykinesia item criterion

The correlations between the CANTAB reaction time variables and the ESRS bradykinesia item are weaker than those using the TAKE bradykinesia item criterion. Both simple reaction time and the SRT movement latency correlate significantly with the criterion; the relationship is of weak to moderate strength. Neither of the choice reaction time variables correlate significantly with the criterion. The groups identified by the measure are of similar size to those identified by the criterion. In all cases, sensitivity is of moderate strength but specificity is low.

3.2.4 Evaluation of the CANTAB reaction time variable using the group of ESRS bradykinesia-related items criterion

The CANTAB variables were also evaluated using the group of ESRS bradykinesia-related items as the criterion.

3.2.4.1 Correlation between the CANTAB reaction time variables and the criterion

Correlation with ESRS group of bradykinesia-related items	Simple RT total time	Choice RT total time
Coefficient	0.297	0.210
Significance	0.031	0.096

Correlation with ESRS group of bradykinesia-related items	Simple RT movement latency	Choice RT movement latency
Coefficient	0.303	0.296
Significance	0.029	0.032

3.2.4.2 Frequency scores for criterion-identified groups

	Frequency	Percent
Normal/Minimal bradykinesia	34	81.0
Bradykinetic	8	19.0
Total	42	100.0

3.2.4.3 Instrumentation groups

The instrumentation group numbers are as before (3.2.2.3):

Simple RT total: 12 above, 28 below

Choice RT total: 10 above, 30 below

Simple RT movement: 12 above, 28 below

Choice RT movement: 11 above, 29 below

3.2.4.4 Results of sensitivity and specificity analysis of the CANTAB reaction time variables using the ESRS group of bradykinesia-related items criterion

3.2.4.4.1 CANTAB simple RT total time against ESRS group of bradykinesia-related items

Measure: *Positive*	8	4
Negative	25	3
	Negative Criterion	*Positive*

Sensitivity = 57.1%

Specificity = 33.3%

3.2.4.4.2 CANTAB simple movement latency against ESRS group of bradykinesia-related items

Measure: *Positive*	8	4
Negative	25	3
	Negative Criterion	*Positive*

Sensitivity = 57.1%

Specificity = 33.3%

3.2.4.4.3 CANTAB choice RT total latency against ESRS group of bradykinesia-related items

Measure: *Positive*	4	6
Negative	23	7
	Negative Criterion	*Positive*

Sensitivity = 46.2%

Specificity = 60.0%

3.2.4.4.4 CANTAB choice RT movement latency against ESRS group of bradykinesia-related items

Measure: *Positive*	5	6

Negative	22	7
	Negative Criterion	*Positive*

Sensitivity = 46.2%

Specificity = 54.5%

3.2.4.5 Summary of the evaluation of the CANTAB reaction time variables using the ESRS group of bradykinesia-related items criterion

The relationships between the CANTAB variables and the criterion are weak; all but that involving the choice reaction time variable are statistically significant. The groups identified by the criterion and the measures are of similar size. In the simple reaction time condition sensitivity is reasonably good though specificity is poor. In the choice reaction time condition sensitivity is lower though specificity is correspondingly higher.

3.2.5 Summary

The correlations between the CANTAB reaction time test variables and the bradykinesia criteria are weak to moderate in strength but statistically significant. As markers of bradykinesia, the CANTAB variables exhibit moderate sensitivity and specificity.

3.3 Rigidity

Observer-ratings of rigidity were made as part of the TAKE and ESRS scales. The TAKE item is an overall assessment for the whole body. The ESRS uses individual items for each limb; those for the upper two limbs are used in this analysis.

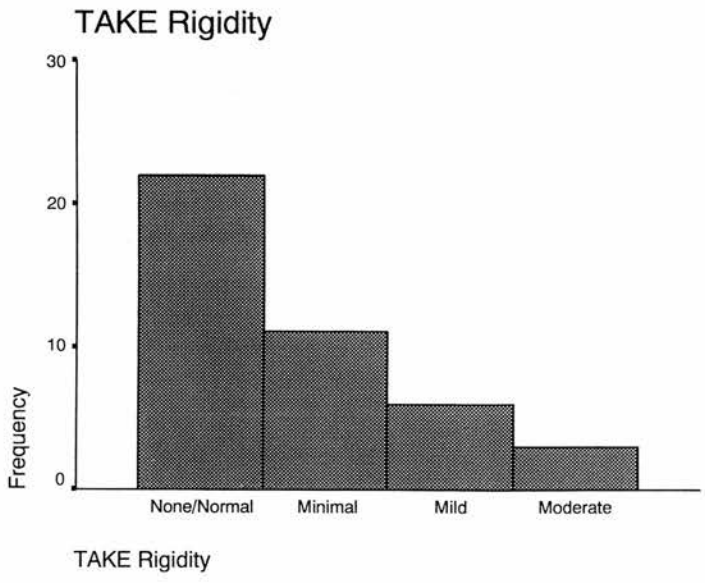
3.3.1 Observer ratings of rigidity

Ratings of rigidity are taken from the TAKE and ESRS scales. The TAKE item is a global assessment of overall rigidity. The two ESRS items used are separate items for the right and left upper limbs assessed independently.

3.3.1.1 Descriptive data for TAKE rigidity item scores in the patient group

	Minimum	Maximum	Median
TAKE rigidity	0	3	0

3.3.1.2 Bar graph of TAKE rigidity scores in the patient group



3.3.1.3 Frequency data for TAKE rigidity scores in the patient group

	Frequency	Percent
None/Normal	22	52.4
Minimal	11	26.2
Mild	6	14.3
Moderate	3	7.1
Total	42	100.0

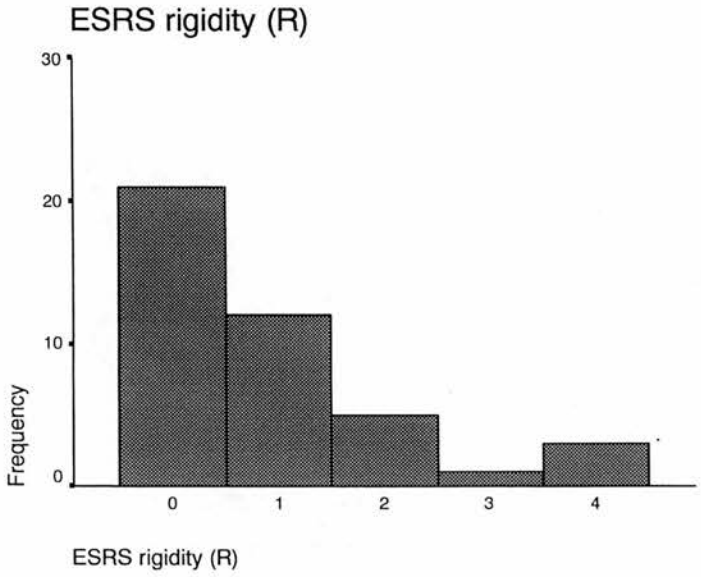
3.3.1.4 Summary of TAKE rigidity item results

Scores on this rating item are generally low. The highest score given is three (moderate rigidity) and over 50% of patients assessed were rated as exhibiting no rigidity. The scores do not exhibit a normal distribution.

3.3.1.5 Descriptive data for the ESRS rigidity (R) item

	Minimum	Maximum	Median
ESRS rigidity (R)	0	4	1

3.3.1.6 Bar graph of ESRS rigidity (R)



3.3.1.7 Score frequencies for ESRS rigidity (R) in the patient group

	Frequency	Percent
0	21	50.0
1	12	28.6
2	5	11.9
3	1	2.4
4	3	7.1
Total	42	100.0

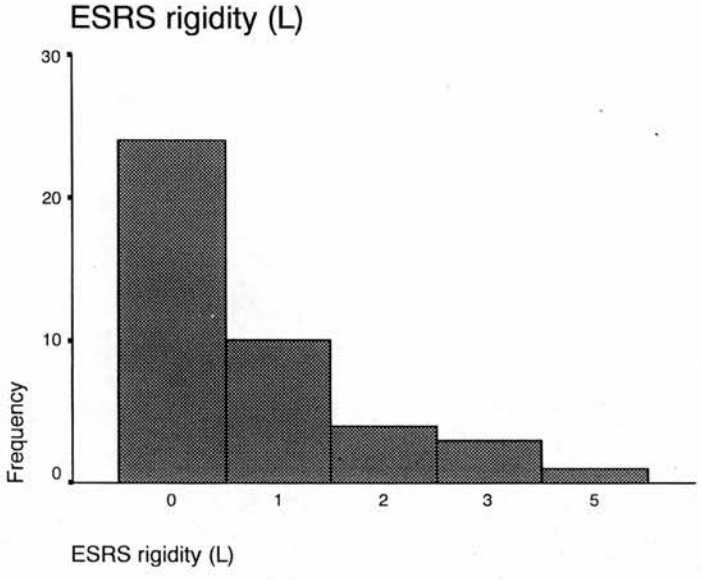
3.3.1.8 Summary of ESRS rigidity (R) ratings

Scores awarded on this rating are low. The maximum awarded was four (of a maximum of six). As with the TAKE rigidity item, over 50% of patients were rated as not exhibiting rigidity. The distribution of scores is again non-normal.

3.3.1.9 Descriptive data for the ESRS rigidity (L) item

	Minimum	Maximum	Median
ESRS rigidity (L)	0	5	0

3.3.1.10 Bar graph of ESRS rigidity (L)



3.3.1.11 Score frequencies for ESRS rigidity (L) in the patient group

	Frequency	Percent
0	24	57.1
1	10	23.8
2	4	9.5
3	3	7.1
4	0	0.0
5	1	2.4
Total	42	100.0

3.3.1.12 Summary of ESRS rigidity (L) scores

Ratings on this item are similar to those on the other rigidity items. Scores are generally low: the maximum awarded was five and the median is zero. The scores are not normally distributed.

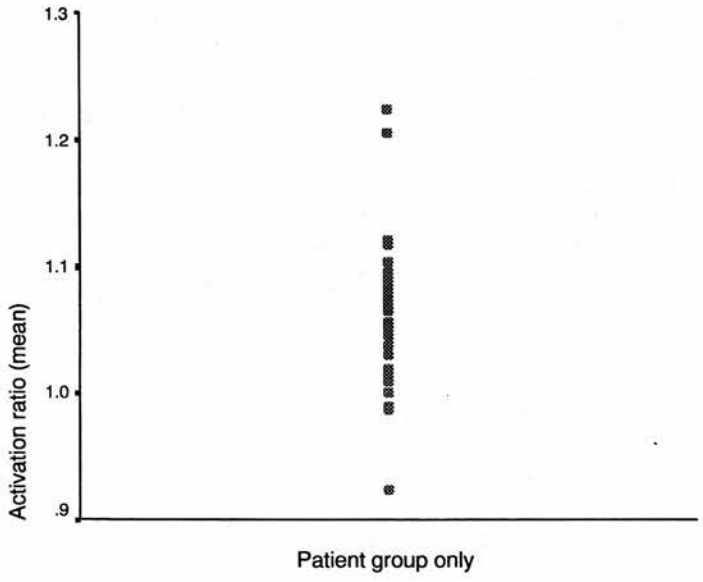
3.3.2 Instrumentation

This section comprises the results of the rigidity instrumentation in the patient group and the evaluation of the rigidity instrumentation using the observer-rating criteria.

3.3.2.1 Descriptive statistics for mean activation ratio in patient group

	N	Minimum	Maximum	Mean	SD
Mean activation ratio	32	0.92	1.22	1.0562	6.051x10-2

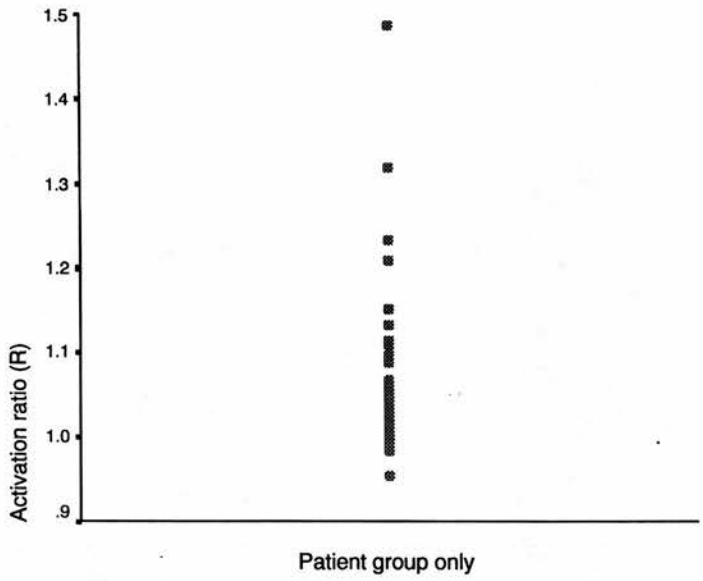
3.3.2.2 Scatterplot of mean activation ratio in patient group



3.3.2.3 Descriptive statistics for activation ratio (R) in patient group

	N	Minimum	Maximum	Mean	SD
Activation ratio (R)	34	0.95	1.49	1.0766	0.1039

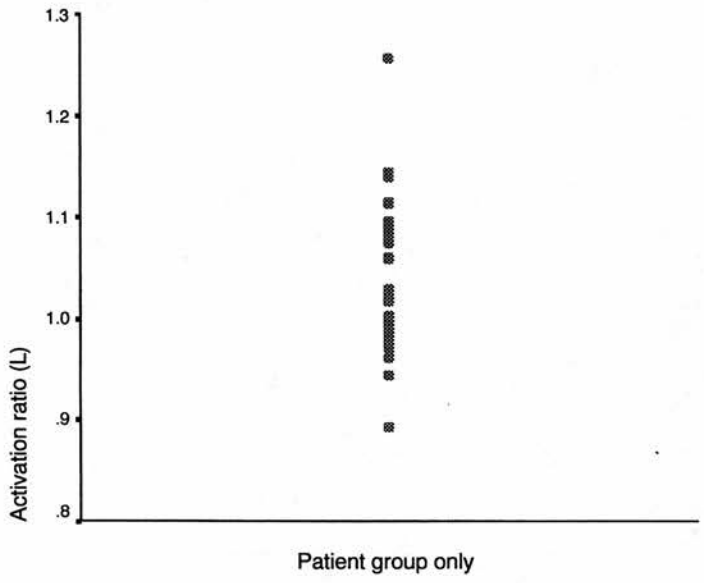
3.3.2.4 Scatterplot of activation ratio (R) in patient group



3.3.2.5 Descriptive statistics for activation ratio (L) in patient group

	N	Minimum	Maximum	Mean	SD
Activation ratio (L)	32	0.89	1.26	1.0401	7.228x10-2

3.3.2.6 Scatterplot of activation ratio (L) in patient group

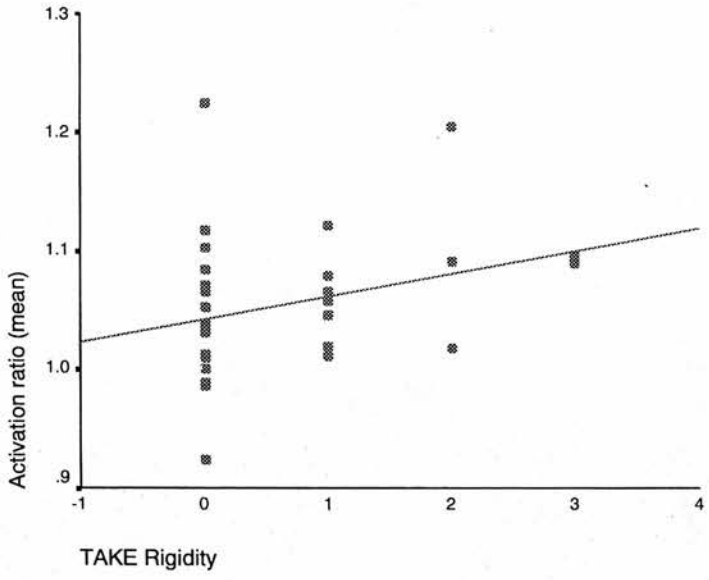


3.3.2.7 Summary of rigidity instrumentation results

As with the observer-ratings of rigidity, there is a predominance of low results on the rigidity instrumentation. However, there are outliers at the low end of the range of results as well as the top.

3.3.3 Evaluation of the mean activation ratio as a measure of rigidity severity relative to the TAKE rigidity item criterion

3.3.3.1 Scatterplot of mean activation ratio against TAKE rigidity in the patient group (a linear best fit line is marked)



3.3.3.2 Correlation between mean activation ratio and TAKE rigidity (two-tailed significance)

N	Coefficient	Significance
32	0.368	0.038

3.3.3.3 Identification of the observer-rating groups

The inclusion threshold for the observer-rating groups was set at two (mild rigidity).

	Frequency	Percent
Normal/Minimal rigidity	26	81
Rigid	6	19

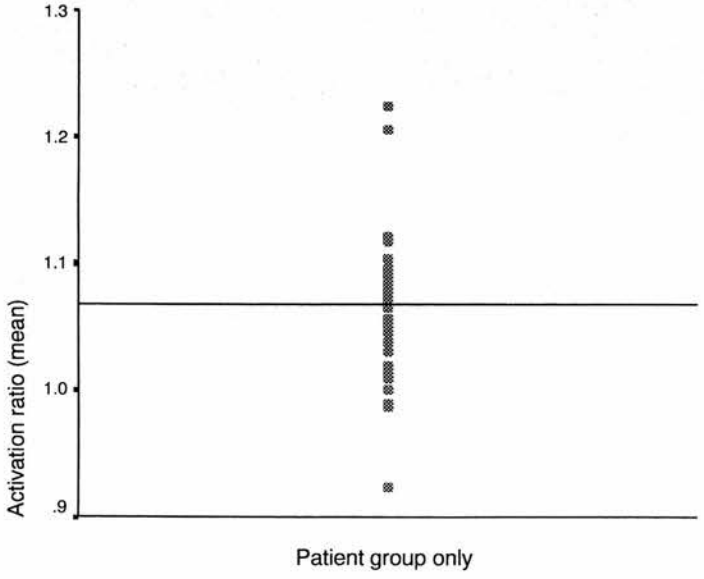
Total	32	100

3.3.3.4 Characteristics of the control group

	Mean	SD	Mean +1SD
Mean activation ratio	1.0239	4.405x10-2	1.06795

A threshold of 1SD above the mean of the control group was used to identify a group exhibiting rigidity.

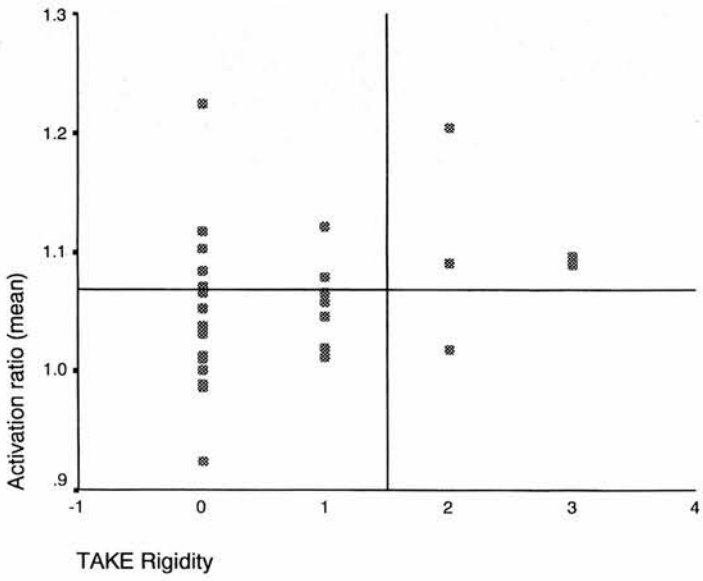
3.3.3.5 Scatterplot of mean activation ratio patient group data with group threshold (mean of the control group +1SD) marked



3.3.3.6 Frequency characteristics of the instrumentation-defined groups (mean activation ratio)

Group	Number	Percent
Less than 1SD above control group mean	19	59.4
More than 1SD above control group mean	13	40.6
Total	32	100.0

3.3.3.7 Scatterplot of mean activation ratio against TAKE rigidity with group thresholds marked



3.3.3.8 Crosstabs procedure results for mean activation ratio against TAKE rigidity

Measure:	8	5

Positive		
Negative	18	1
	Negative Criterion	*Positive*

Sensitivity = 83.3%

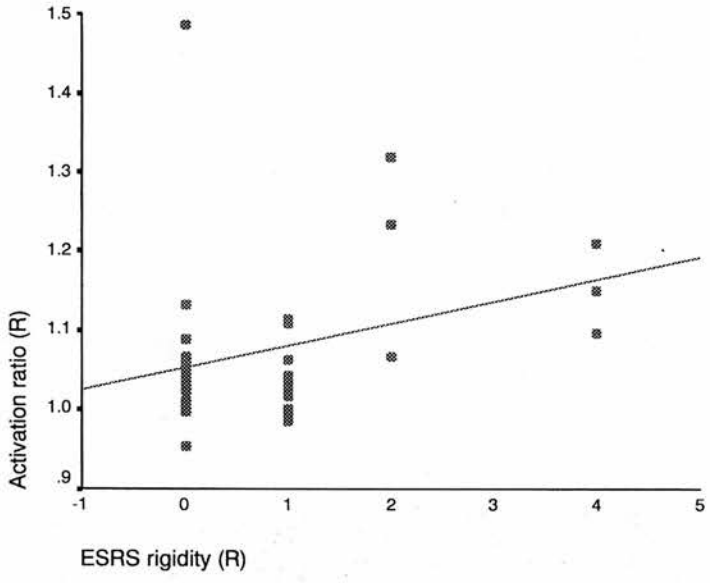
Specificity = 38.5%

3.3.3.9 Summary of the evaluation of the mean activation ratio using the TAKE rigidity criterion

The correlation between the measure and the criterion is of weak to moderate strength and statistically significant. The measure and criterion identified groups are of similar sizes. The sensitivity of the measure is very good but specificity is poor; more cases were falsely identified by the measure as rigid than were correctly identified.

3.3.4 Evaluation of the activation ratio (R) using the ESRS rigidity (R) criterion

3.3.4.1 Scatterplot of activation ratio (R) against ESRS rigidity (R) with linear best fit line marked



3.3.4.2 Correlation between activation ratio (R) and ESRS rigidity (R) (two-tailed significance)

N	Coefficient	Significance
34	0.324	0.061

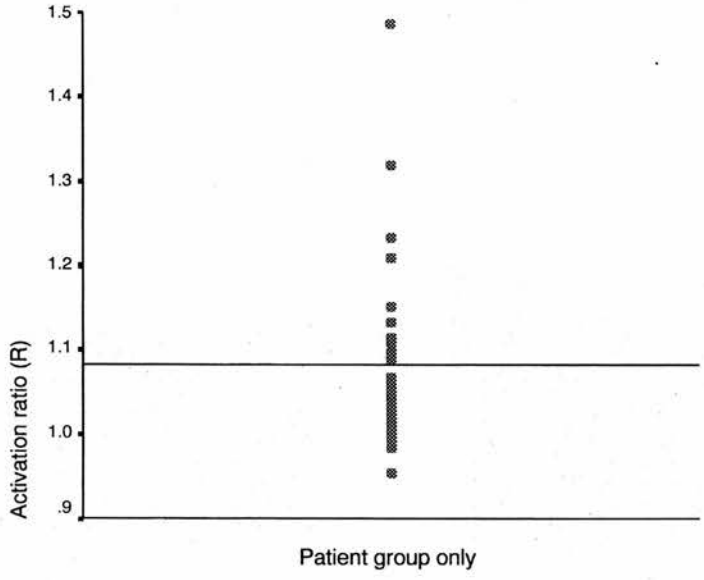
3.3.4.3 Groups were identified using the ESRS rigidity (R) observer-rating criterion

	Frequency	Percent
Normal/Minimal rigidity	28	82
Rigid	6	18
Total	34	100

3.3.4.4 Groups were identified using a threshold of 1SD above mean of control group

	Mean	SD	Mean +1SD
Activation ratio (R)	1.0394	4.270x10-2	1.08210

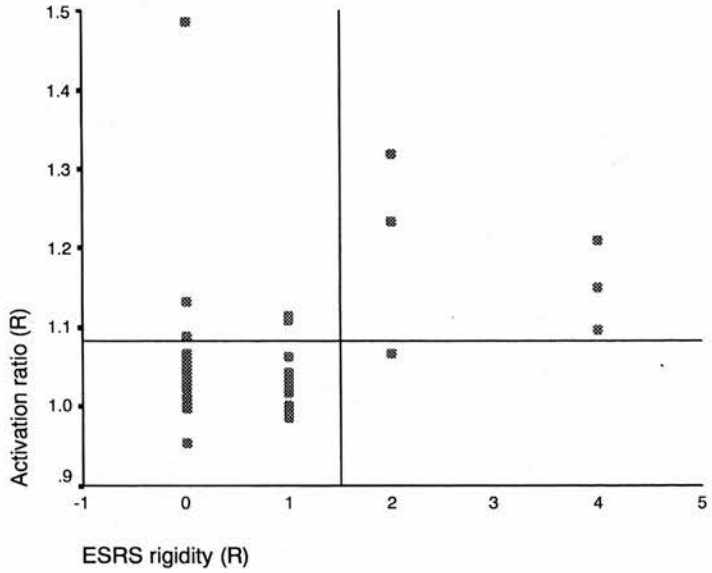
3.3.4.5 Scatterplot of activation ratio (R) with group threshold marked



3.3.4.6 Frequency data for activation ratio (R) groups

Group	Number	Percent
Less than 1SD above control group mean	24	70.6
More than 1SD above control group mean	10	29.4
Total	34	100.0

3.3.4.7 Scatterplot of activation ratio (R) against ESRS rigidity (R) with group thresholds marked



3.3.4.8 Results of the crosstabs procedure for activation ratio (R) against ESRS rigidity (R)

Measure: *Positive*	5	5
Negative	23	1
	Negative Criterion	*Positive*

Sensitivity = 83.3%

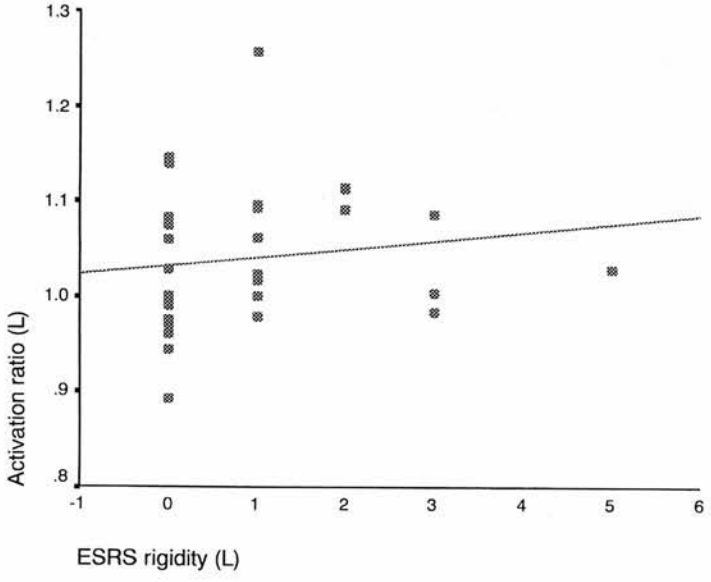
Specificity = 50.0%

3.3.4.9 Summary of the evaluation of the activation ratio (R) using the ESRS rigidity (R) criterion

The correlation between the activation ratio (R) and the ESRS rigidity (R) criterion was of moderate strength but only trends towards statistical significance. The instrumentation and observer-rating groups are of similar sizes. Sensitivity is very high and specificity is also good.

3.3.5 Evaluation of activation ratio (L) against ESRS rigidity (L)

3.3.5.1 Scatterplot of activation ratio (L) against ESRS rigidity (L) with linear best-fit line marked



3.3.5.2 Correlation between activation ratio (L) and ESRS rigidity (L) (two-tailed significance)

N	Coefficient	Significance
32	0.324	0.071

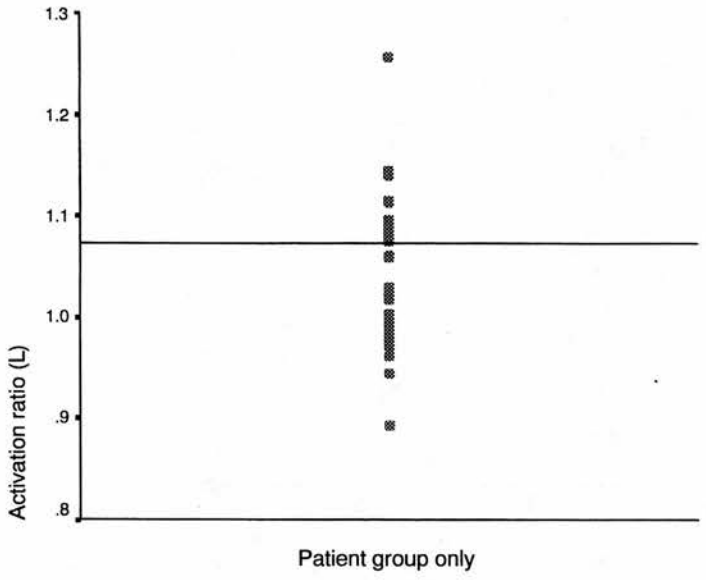
3.3.5.3 Observer-rating groups were identified using the ESRS rigidity (L) item

	Frequency	Percent
Normal/Minimal rigidity	25	78
Rigid	7	22
Total	32	100

3.3.5.4 Instrumentation groups were identified using a threshold of 1SD above the mean of the control group

	Mean	SD	Mean +1SD
Activation ratio (L)	1.0083	6.526x10-2	1.07356

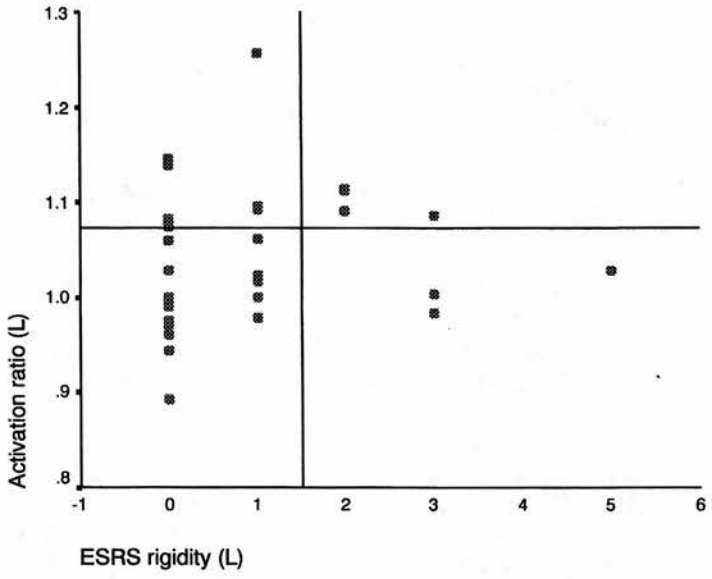
3.3.5.5 Scatterplot of activation ratio (L) with group threshold (1SD above mean of control group) marked



3.3.5.6 Frequency data for the activation ratio (L) groups

Group	Number	Percent
Less than 1SD above control group mean	20	62.5
More than 1SD above control group mean	12	37.5
Total	32	100.0

3.3.5.7 Scatterplot of activation ratio (L) against ESRS rigidity (L) with group thresholds marked



3.3.5.8 Crosstabs procedure results for activation ratio (L) against ESRS rigidity (L)

Measure: *Positive*	8	4
Negative	17	2
	Negative Criterion	*Positive*

Sensitivity = 66.7%

Specificity = 33.3%

3.3.5.9 Summary of the evaluation of the activation ratio (L) using the ESRS rigidity (L) item

The correlation between the measure and the criterion was of weak to moderate strength but non-significant. The observer-rating and instrumentation groups are of very similar size. Sensitivity is good but specificity is poor.

3.3.6 Summary

Observer-ratings of rigidity indicate that clinically significant levels of rigidity were rare in this patient group. No cases of severe rigidity were found and few reached even mild level. The correlations between the activation ratio and the ratings criteria were weak to moderate; only that for the mean activation ratio reaches significance. As a marker of rigidity the activation ratio exhibits a high level of sensitivity but its specificity is low to moderate.

3.4 Tremor – postural tremor amplitude

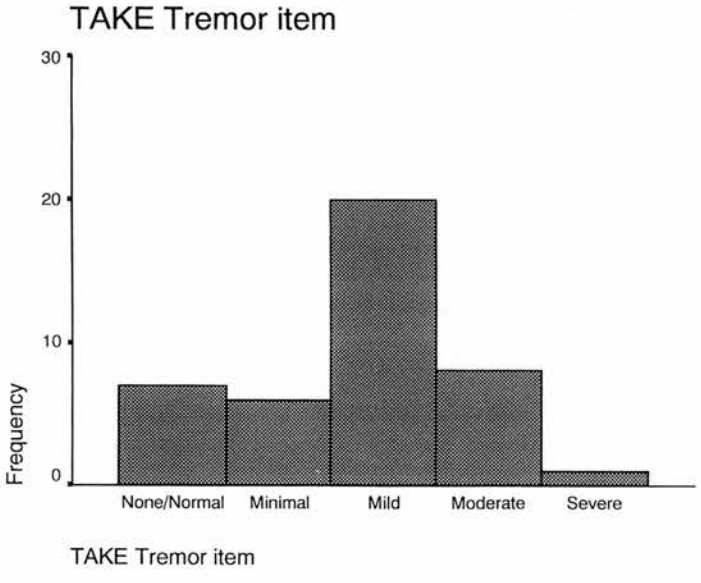
3.4.1 Observer-ratings of tremor

Observer ratings of tremor are taken from the TAKE and ESRS rating scales. These ratings are intended to be overall ratings of tremor but generally form assessments of postural tremor amplitude. The TAKE item is a global assessment of tremor severity. The two ESRS items are for the two upper limbs assessed independently (the novel scoring guidelines for the ESRS items were discussed earlier, 2.3.1).

3.4.1.1 Descriptive data for TAKE tremor in the patient group

	Minimum	Maximum	Median
TAKE tremor	0	4	2

3.4.1.2 Bar graph of TAKE tremor scores



3.4.1.3 Frequency breakdown of TAKE tremor item scores

Rating	Frequency	Percent
None/Normal	7	16.7
Minimal	6	14.3
Mild	20	47.6
Moderate	8	19.0
Severe	1	2.4
Total	**42**	**100.0**

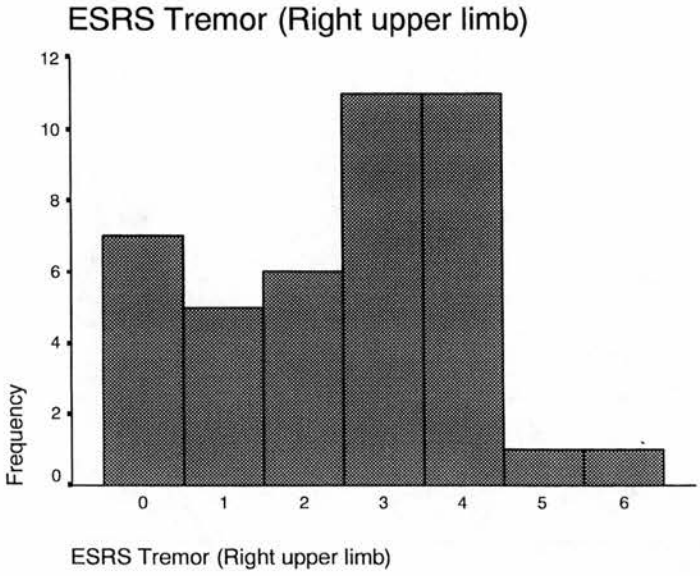
3.4.1.4 Summary of TAKE tremor item scores

The pattern of scores on this rating is very different from the other features previously considered. The majority of patients (69%) exhibit symptomatic tremor rated mild or greater. Severe tremor is rare (2.4%) but only 16.7% of patients do not exhibit greater than normal levels of tremor.

3.4.1.5 Descriptive data for ESRS tremor (RH) scores in the patient group only

	Minimum	Maximum	Median
ESRS tremor (RH)	0	6	3

3.4.1.6 Bar graph of ESRS tremor (RH) scores



3.4.1.7 Frequency breakdown of ESRS tremor (RH) scores

Score	Frequency	Percent
0	7	16.7
1	5	11.9
2	6	14.3
3	11	26.2
4	11	26.2
5	1	2.4
6	2	2.4
Total	**42**	**100.0**

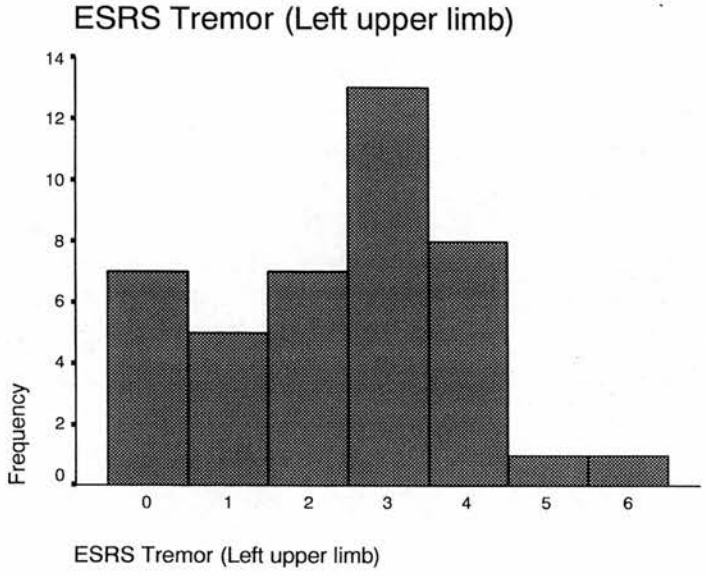
3.4.1.8 Summary of ESRS tremor (RH) scores

Though the rating scheme for the ESRS tremor items differs from that used for the TAKE tremor item, the pattern of scores is superficially very similar. Severe tremor is rare but most patients do exhibit some level of symptomatic tremor.

3.4.1.9 Descriptive data for ESRS tremor (LH) scores

	Minimum	Maximum	Median
ESRS tremor (LH)	0	6	3

3.4.1.10 Bar graph of ESRS tremor (LH) scores



3.4.1.11 Frequency break down of ESRS Tremor (LH) scores

Score	Frequency	Percent
0	7	16.7
1	5	11.9
2	7	16.7
3	13	31.0
4	8	19.0
5	1	2.4

6	1	2.4
Total	**42**	**100.0**

3.4.1.12 Summary of ESRS tremor (LH) scores

The ESRS tremor ratings for the left hand closely parallel those for the right hand: there are no apparent asymmetries. Again, severe tremor is rare but some level of symptomatic tremor is very common.

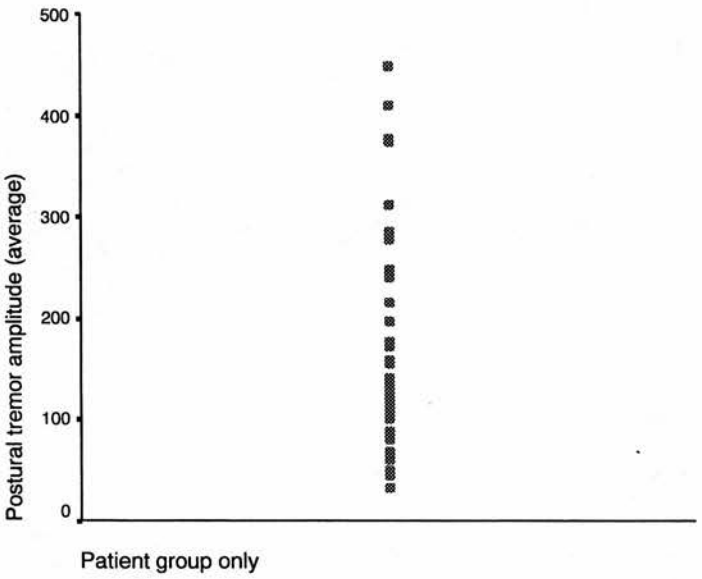
3.4.2 Instrumentation

Postural tremor amplitude was used as the primary comparative variable. Within the patient group this feature showed the following results

3.4.2.1 Mean postural tremor amplitude descriptive statistics for the patient group only

	Minimum	Maximum	Mean	SD
Postural tremor amplitude (mean)	33.34	448.80	162.47	104.64

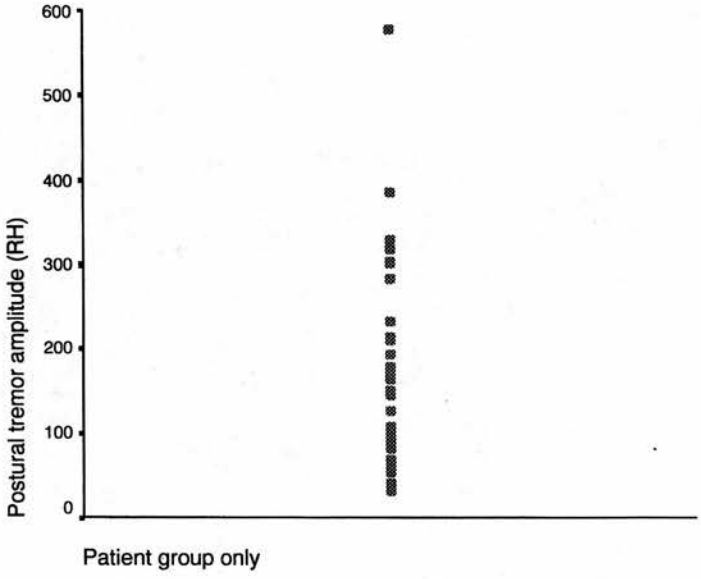
3.4.2.2 Scatterplot of mean postural tremor amplitude



3.4.2.3 Postural tremor amplitude (RH) descriptive statistics for the patient group only

	Minimum	Maximum	Mean	Std. Deviation
Postural tremor amplitude (RH)	32.29	576.74	160.73	109.77

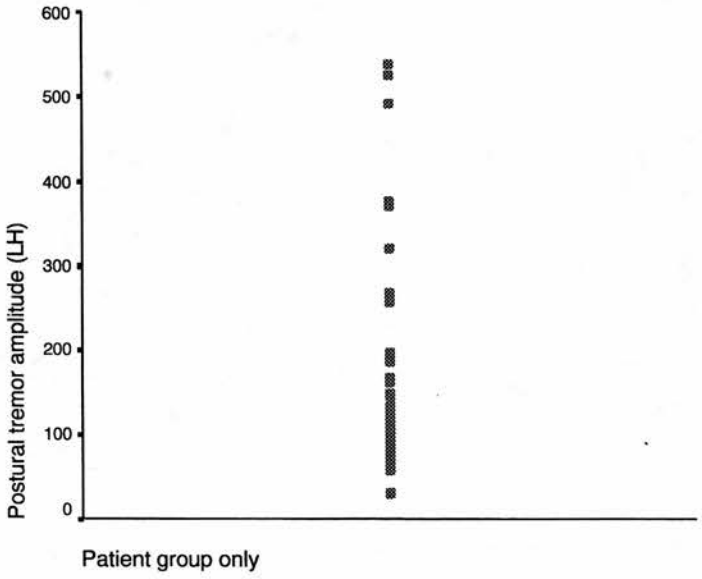
3.4.2.4 Scatterplot of postural tremor amplitude (RH)



3.4.2.5 Postural tremor amplitude (LH) descriptive statistics for the patient group only

	Minimum	Maximum	Mean	Std. Deviation
Postural tremor amplitude (LH)	30.31	538.23	164.21	131.19

3.4.2.6 Postural tremor amplitude (LH) scatterplot

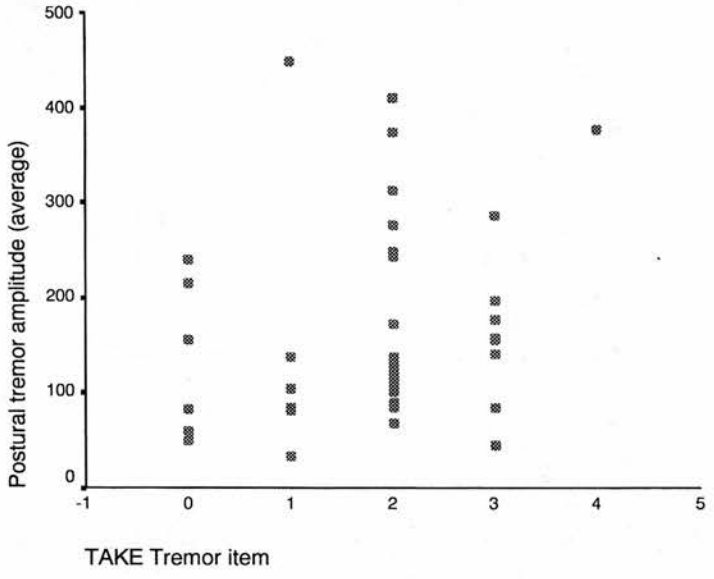


3.4.2.7 Summary of postural tremor amplitude results

All three variables indicate that a moderate degree of tremor is very common. Where tremor is severe it may be of very great amplitude but this is relatively rare. The instrumented results are superficially consistent with the observer-ratings.

3.4.3 Evaluation of the postural tremor amplitude variable as a measure of tremor severity using the TAKE tremor variable criterion

3.4.3.1 Scatterplot of mean postural tremor amplitude against TAKE tremor item



3.4.3.2 Correlation between mean postural tremor amplitude and TAKE tremor item (two-tailed significance)

N	Coefficient	Significance
42	0.301	0.053

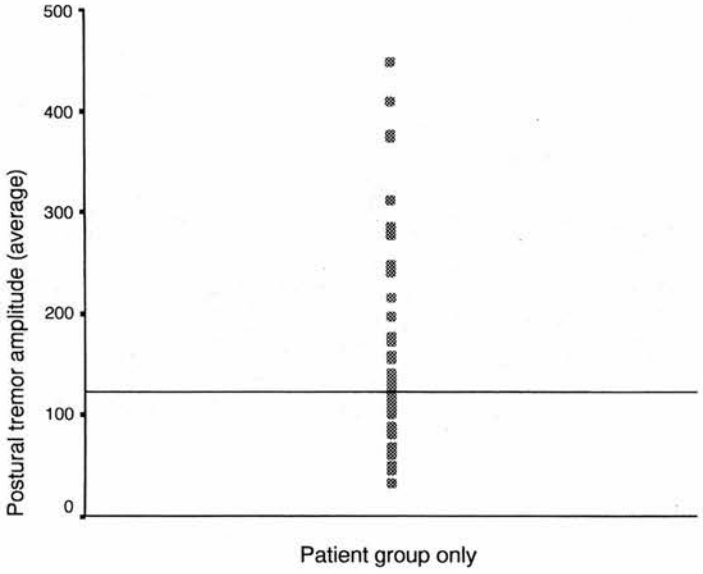
3.4.3.3 Frequency characteristics of observer-rating identified groups (using a threshold score of 2 or greater)

	Number	Percent
None/Minimal tremor	13	31.0
Tremulous	29	69.0
Total	42	100.0

3.4.3.4 Instrumentation groups were identified using a threshold of 1SD above the control group mean

	Mean	Std. Deviation	Mean + 1SD
Mean tremor amplitude	162.47	104.64	267.11

3.4.3.5 Scatterplot of mean postural tremor amplitude with the threshold value (1SD above the control group mean) marked

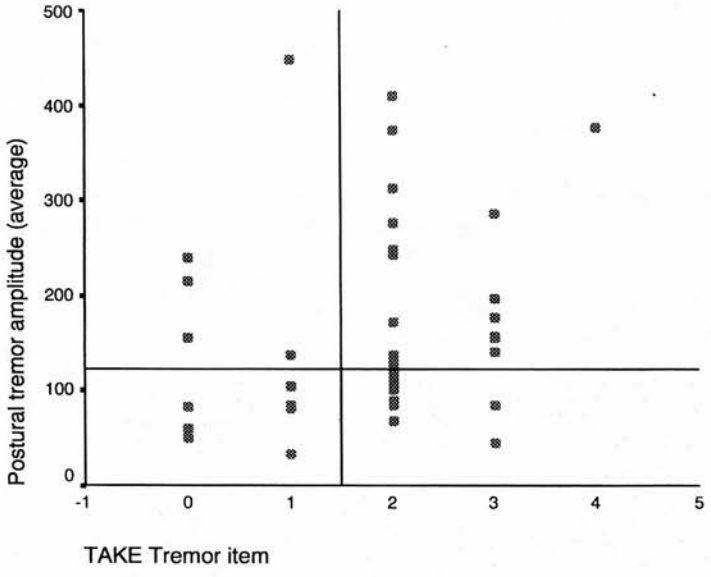


3.4.3.6 Frequency data for the patient groups identified using the instrumented mean postural tremor amplitude variable

	Number	Percent
Less than 1SD above control group mean	19	45.2
More than 1SD above	23	54.8

control group mean		
Total	**42**	**100.0**

3.4.3.7 Scatterplot of mean postural tremor amplitude against TAKE tremor item with tremulous group inclusion thresholds marked



3.4.3.8 Results of crosstabs procedure

Measure: *Positive*	5	18
Negative	8	11
	Negative Criterion	*Positive*

Sensitivity = 62.1%

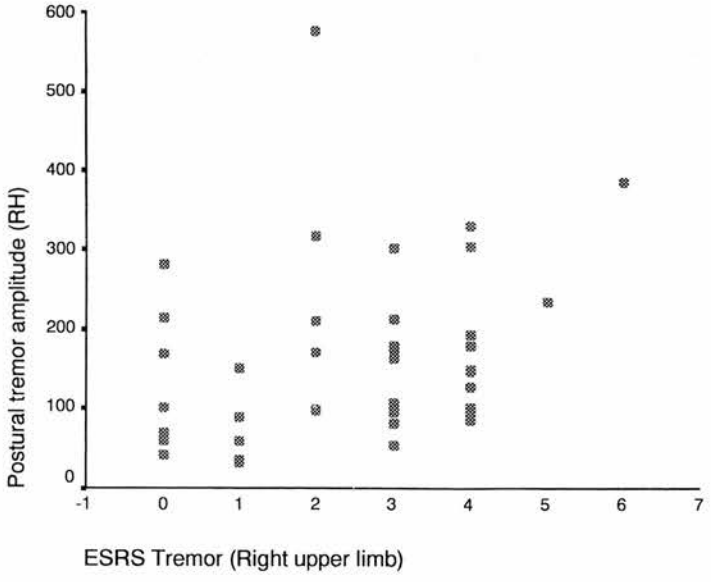
Specificity = 78.3%

3.4.3.9 Summary of the evaluation of mean postural tremor amplitude using the TAKE tremor item

The correlation between the measure and the criterion is of moderate strength and trends towards significance. The measure and criterion groups are of similar sizes. Sensitivity is good and specificity is very good. However, some patients rated as exhibiting tremor were not identified as tremulous by the measure.

3.4.4 Evaluation of the postural tremor amplitude variable (RH) using the ESRS tremor (RH) item criterion

3.4.4.1 Scatterplot of postural tremor amplitude variable (RH) against ESRS tremor (RH)



3.4.4.2 Correlation of the postural tremor amplitude variable (RH) against ESRS tremor (RH)relationship (two-tailed significance)

N	Coefficient	Significance
42	0.292	0.061

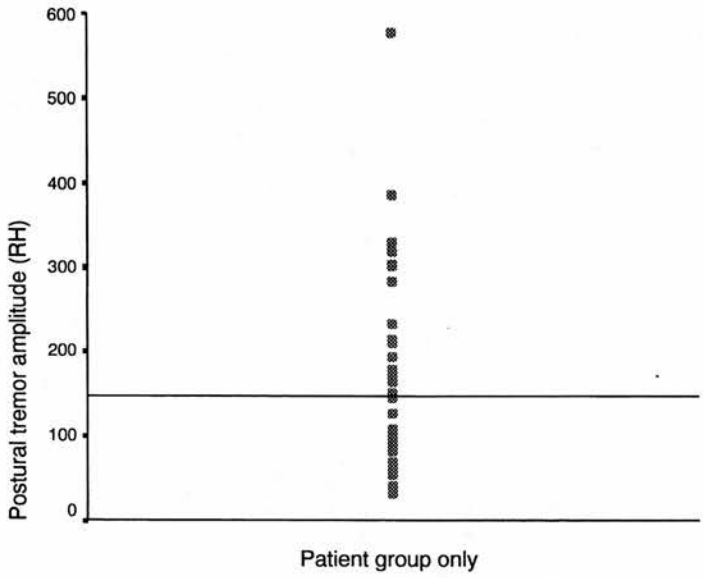
3.4.4.3 Frequency data for the observer-rating (ESRS tremor (RH)) identified groups (inclusion threshold of 2)

	Number	Percent
None/Minimal tremor	18	42.9
Tremulous	24	57.1
Total	42	100.0

3.4.4.4 Groups were identified using the postural tremor amplitude (RH) variable with a threshold value of 1SD above the mean of the control group

	Mean	**Std. Deviation**	**Mean + 1SD**
Tremor amplitude (RH)	160.73	109.77	270.50

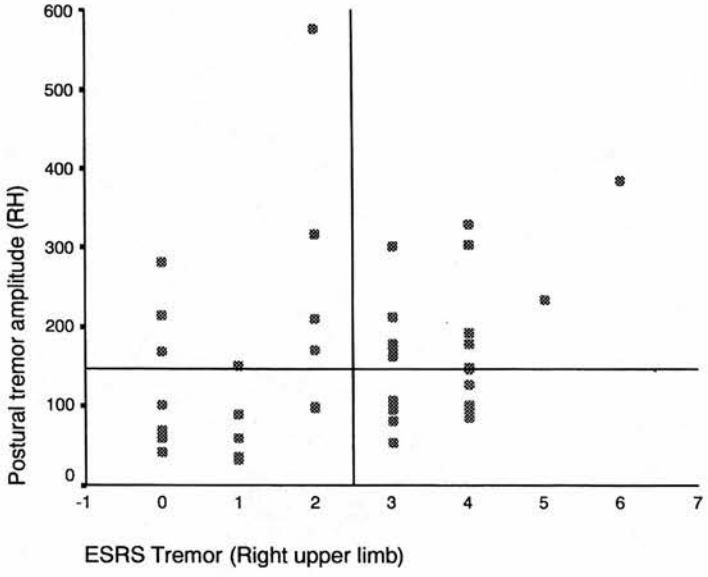
3.4.4.5 Scatterplot of postural tremor amplitude (RH) with threshold value (1SD above mean of control group) marked



3.4.4.6 Frequency data for groups identified using postural tremor amplitude (RH)

	Number	Percent
Less than 1SD above control group mean	22	52.4
More than 1SD above control group mean	20	47.6
Total	42	100.0

3.4.4.7 Scatterplot of postural tremor amplitude (RH) against ESRS tremor (RH) item



3.4.4.8 Results of crosstabs procedure

Measure: *Positive*	8	12
Negative	10	12
	Negative Criterion	*Positive*

Sensitivity = 50.0%

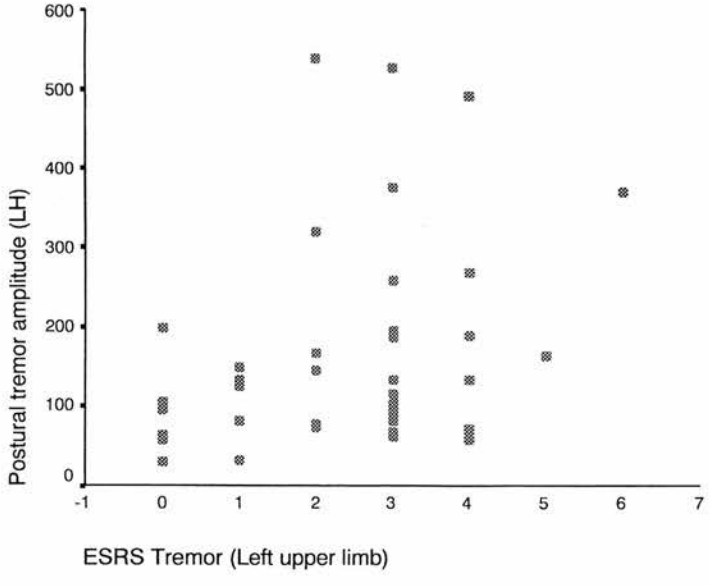
Specificity = 60.0%

3.4.4.9 Summary of the evaluation of postural tremor amplitude (RH) using the ESRS tremor (RH) item

Correlation between the measure and the criterion is moderate and trends towards significance. The criterion and measure group sizes are less similar than those used in the previous comparison (mean postural tremor amplitude and TAKE tremor item). Sensitivity is moderate but specificity is good. However, a notable number of cases positively identified by the criterion were not detected by the measure.

3.4.5 Evaluation of the instrumented postural tremor amplitude (LH) variable using the ESRS tremor (LH) observer-rating as criterion

3.4.5.1 Scatterplot of postural tremor amplitude (LH) against ESRS tremor rating (LH)



3.4.5.2 Correlation of instrumented postural tremor amplitude (LH) with ESRS observer-rating of tremor severity (LH) (two-tailed significance)

N	Coefficient	Significance
42	0.336	0.030

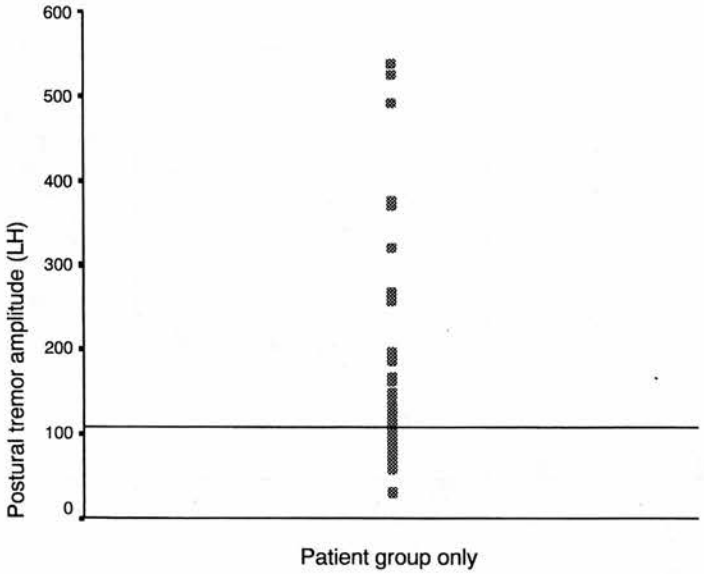
3.4.5.3 Frequency data for the groups identified by the criterion (ESRS tremor item for the LH with an inclusion threshold value of 2 or greater)

	Number	Percent
None/Minimal tremor	19	45.2
Tremulous	23	54.8
Total	42	100.0

3.4.5.4 Groups were identified using the instrumented variable postural tremor amplitude with a threshold value of 1SD above the control group mean

	Mean	**Std. Deviation**	**Mean + 1SD**
Tremor amplitude (LH)	164.21	131.19	295.40

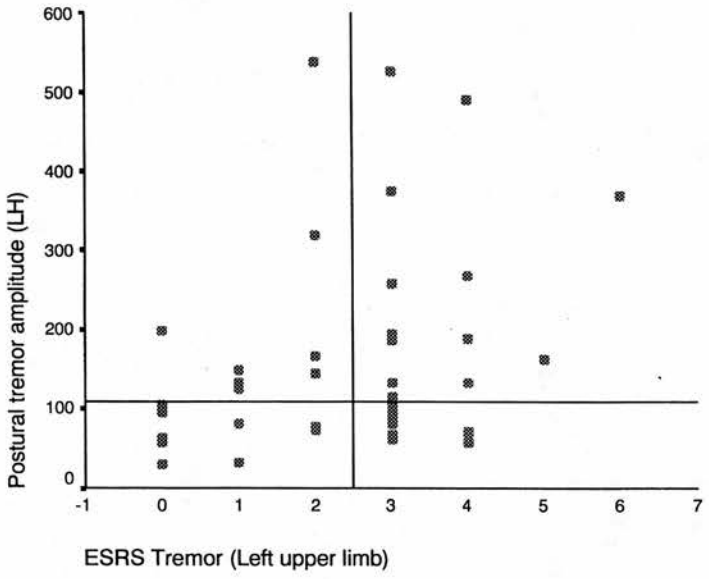
3.4.5.5 Scatterplot of postural tremor amplitude (LH) in the patient group with the inclusion threshold value marked



3.4.5.6 Frequency data for the patient groups identified using instrumented postural tremor amplitude (LH)

	Number	Percent
Less than 1SD above control group mean	20	47.6
More than 1SD above control group mean	22	52.4
Total	42	100.0

3.4.5.7 Scatterplot of postural tremor amplitude (LH) against ESRS tremor severity (LH) with threshold values marked



3.4.5.8 Results of crosstabs procedure

Measure: *Positive*	8	14
Negative	11	9
	Negative Criterion	*Positive*

Sensitivity = 60.9%

Specificity = 63.6%

3.4.5.9 Summary of the evaluation of postural tremor amplitude (LH) using the ESRS tremor (LH) item

Correlation between the measure and the criterion is of moderate strength and statistically significant. The group sizes are similar. Both sensitivity and specificity are high.

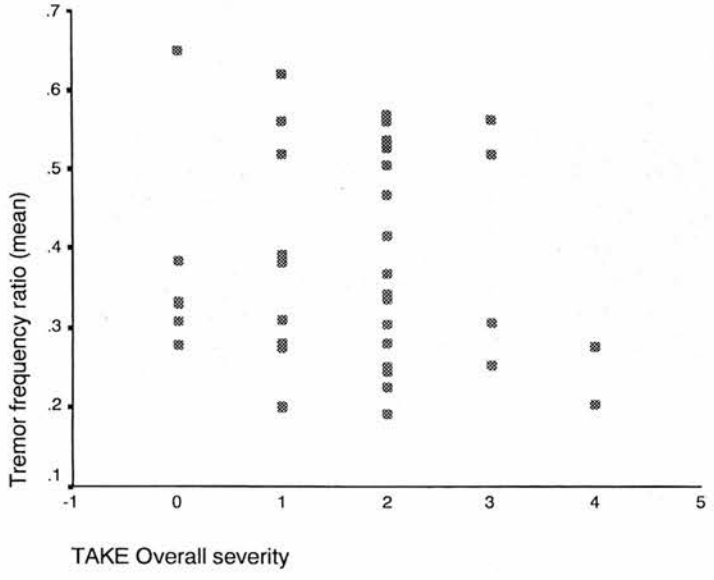
3.4.6 Summary

Observer-ratings of tremor indicated that symptomatic tremor was very common in the patient group though it was rarely of more than mild to moderate severity. Only weak to moderate correlations were found between postural tremor amplitude and the rating criteria; not all of the correlations reach significance though all trended towards significance. In contrast, the sensitivity and specificity of the variable as a marker of tremor were both high.

3.5 Tremor – resting tremor frequency ratio

The frequency ratio calculated from the FFT analysis results provided a variable which has been proposed as a marker of parkinsonism. For this evaluation the instrumentation measure will be the frequency ratio described earlier (2.1.4). The criteria will be overall ratings of parkinsonism: the overall severity item from the TAKE scale and ESRS total score.

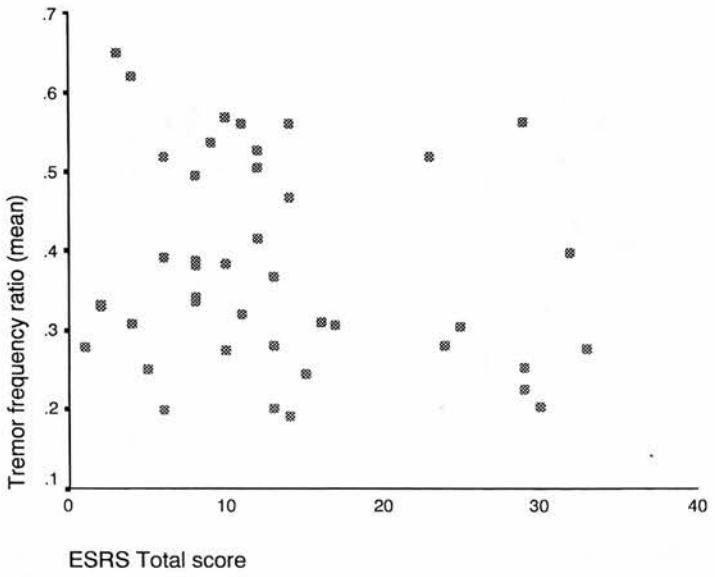
3.5.1 Scatterplot of mean tremor frequency ratio against TAKE overall severity



3.5.2 Correlation between mean tremor frequency and TAKE overall severity

N	Coefficient	Significance
39	-0.091	NS

3.5.3 Scatterplot of mean tremor frequency ratio against ESRS total score



3.5.4 Correlation between mean tremor frequency ratio and ESRS total score

N	Coefficient	Significance
42	-0.223	NS

3.5.5 Effect of Lithium tremor on analysis of tremor as a marker of parkinsonism

The negative correlations between tremor frequency ratio and the criteria indicate that the proportion of low frequency tremor (relative to overall total tremor activity) is lower in those rated as having more severe parkinsonism. Further, both correlations are non-significant.

However, in the light of the known influence of lithium therapy on tremor characteristics, the correlations were re-calculated following the exclusion of all patients receiving lithium.

	N	Coefficient	Significance
Mean frequency ratio vs. TAKE overall severity	31	-0.162	NS
Mean frequency ratio vs. ESRS total score	31	-0.313	NS

Following the exclusion of lithium-treated patients, the results are little changed. The correlations remain negative; though very slightly weaker, the correlation between mean frequency ratio and ESRS total score now trends towards significance ($p = 0.863$).

3.5.6 Summary

The resting tremor frequency ratio showed no relationship to ratings of the global severity of parkinsonism.

3.6 Motor Planning Impairment

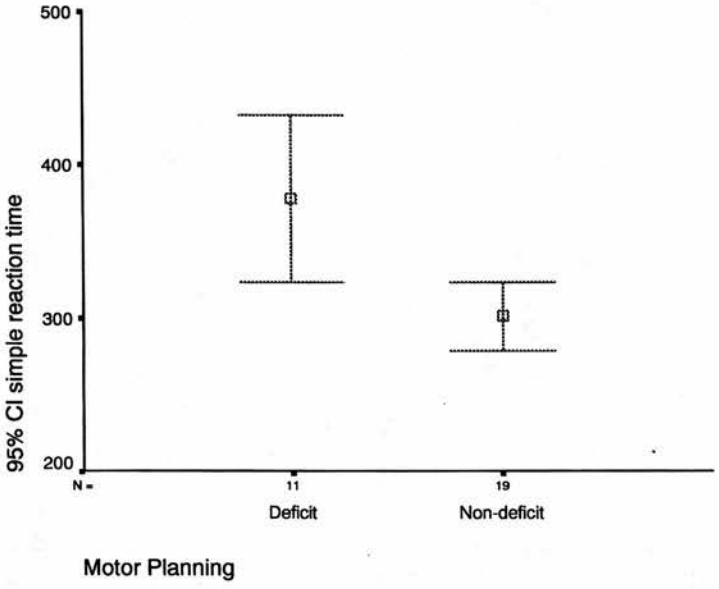
The calculation of a measure of motor planning time variable was described earlier (2.1.5). This variable may be used to identify a group who exhibit a deficit in motor planning. The members of this group are those patients who do not show the normal increase in reaction latency in the CRT condition over the SRT condition, i.e. their motor planning latency appears to be shorter than that of the control group.

With the aim of ensuring group comparability, only patients receiving antipsychotic medication were included in this analysis. From these patients, two groups were identified: a deficit group in whom motor planning time is more than 1SD below the mean of the control group, and a non-deficit group in whom motor planning time is not shorter than that of the control group.

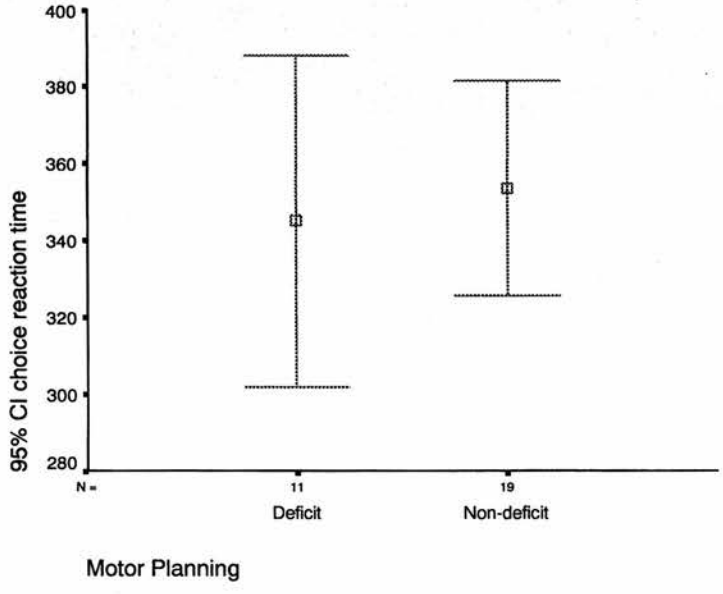
3.6.1 Reaction latencies in the deficit and non-deficit patient groups

The following error bar graph demonstrates that mean reaction latency in the deficit group is increased over the non-deficit group in the SRT condition but not in the CRT condition. The difference between the groups in the SRT condition is statistically significant ($p = 0.012$).

3.6.1.1 Graph of reaction latencies in the SRT condition



3.6.1.2 Graph of reaction latencies in the CRT condition



3.6.2 Comparison of the deficit and non-deficit groups using observer-ratings

The deficit and non-deficit patient groups were compared using the observer-rating criteria. Median scores in the two groups and the significance levels of any differences are found are presented below. In all cases, severity is greater in the deficit group than the non-deficit group.

Item	Significance
TAKE:	
Bradykinesia	p = 0.018
Rigidity	p = 0.005
Total of items 1-5	p = 0.007
ESRS:	
Group of bradykinesia-related items	p = 0.052
Total (tremor excluded)	p = 0.006

Trends were found towards findings of greater negative symptomatology and observer-rated depression in the deficit group.

Another notable finding is the lack of significant difference between the two groups using the total score of the ESRS scale ($p>0.05$) . It is only when the tremor items are excluded that the difference reaches significance. Removing the tremor item from the TAKE total of items 1-5 (i.e. considering items 1,2,4 and 5 only) increases the significance of the difference between the two groups ($p = 0.007$ to 0.002).

3.6.3 Comparison of the deficit and non-deficit patient groups and the control group using objective measures of bradykinesia

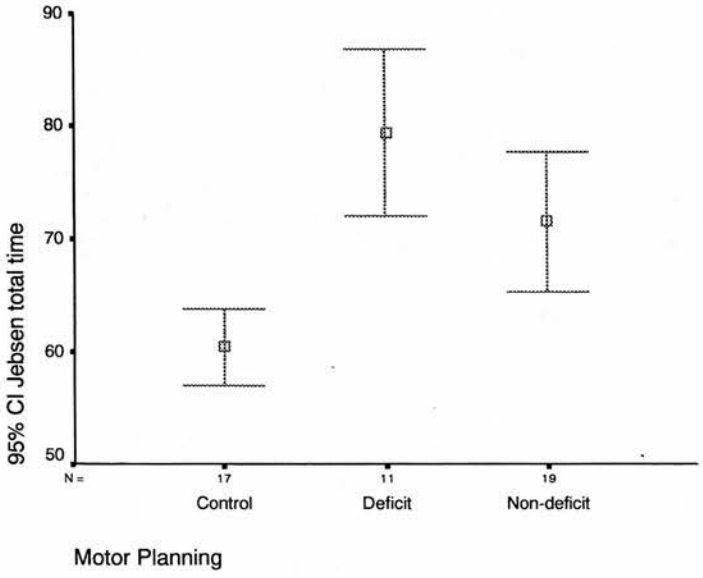
Noting the evidence reviewed previously of an association between non-physical features of parkinsonism, particularly motor control impairments, and bradykinesia, the motor planning deficit and non-deficit groups were compared with each other and with the control group using the objective measures of bradykinesia (Jebsen Hand Function Test and CANTAB reaction time test).

3.6.3.1 Jebsen Hand Function Test – results of ANOVA and post-hoc tests

An overall effect of group was found on Jebsen total time ($p = 0.000$).

	Control	Non-deficit
Deficit	$p = 0.000$	NS ($p = 0.053$ by LSD)
Non-deficit	$p = 0.003$	

3.6.3.2 Error bars of 95% confidence intervals for group means of Jebsen results

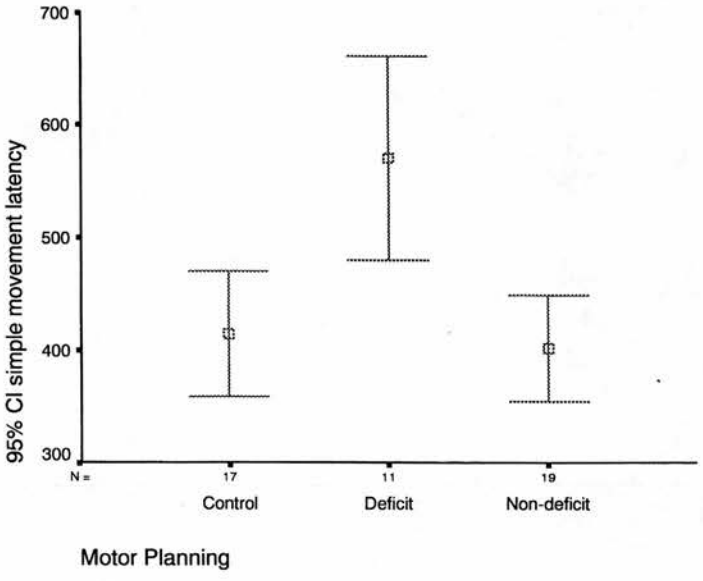


3.6.3.3 CANTAB SRT movement latency – results of ANOVA and post-hoc tests

An overall effect of motor planning group was found on CANTAB SRT movement latency ($p = 0.000$).

	Control	Non-deficit
Deficit	$p = 0.002$	$p = 0.000$
Non-deficit	NS	

3.6.3.4 Error bars of 95% confidence intervals for group means of CANTAB SRT movement latency results

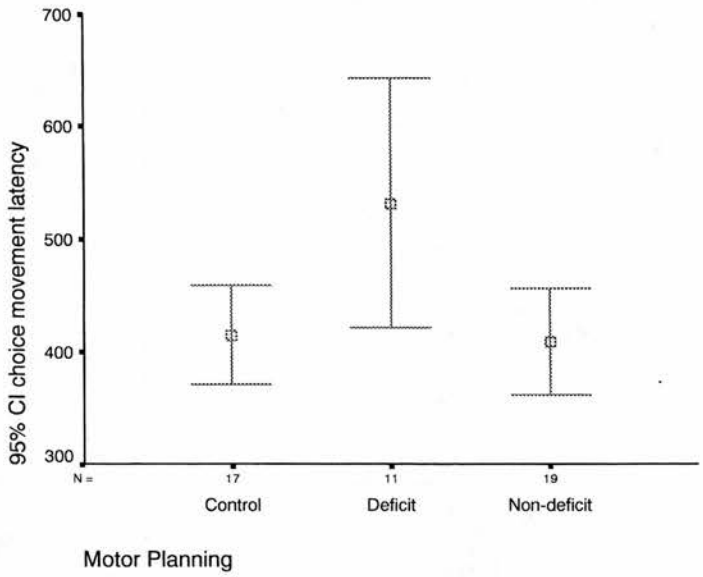


3.6.3.5 CANTAB CRT movement latency – results of ANOVA and post-hoc tests

An overall effect of motor planning group was found CANTAB CRT movement latency ($p = 0.013$).

	Control	Non-deficit
Deficit	$p = 0.032$	$p = 0.019$
Non-deficit	NS	

3.6.3.6 Error bars of 95% confidence intervals for group means of CANTAB CRT movement latency results



3.6.3.7 Summary of the comparison of motor planning groups using objective measures of bradykinesia

All three instrumented bradykinesia variables exhibit a significant effect of motor planning group. In all cases, the deficit group is slower than the control group. The non-deficit patient group is slower than the controls on the Jebsen test but not on the two CANTAB movement latency variables. The deficit and non-deficit patient groups differ on SRT and CRT movement latency, the deficit group being slower. A similar difference on Jebsen total time trends towards significance.

3.6.4 Comparison of the deficit and non-deficit patient groups and the control group using other objective measures

The deficit and non-deficit patient groups were also compared with each other and with the control group using measures of tremor and rigidity. The variables used were the mean resting tremor frequency ratio and the mean rigidity activation ratio. In both the cases the deficit and non-deficit groups did not differ significantly.

3.6.5 Summary

The presence has been demonstrated of an impairment in motor planning in a proportion of antipsychotic treated patients. This impairment is associated with increased severity of bradykinesia.

3.7 Subjective experience

This section comprises the analysis of results from the SWN and visual analogue scales, and an investigation of the validity of the scales.

3.7.1 SWN scale

3.7.1.1 Correlations between the sub-scales and the total score

SWN sub-scale	r	p
Emotional regulation	0.897	0.000
Self-control	0.900	0.000
Mental functioning	0.917	0.000
Social integration	0.908	0.000
Physical functioning	0.929	0.000

The extremely high strength of all correlations indicates that there is a great degree of correlation between all the sub-scales. Therefore it was decided to continue the analysis using only the total score.

3.7.1.2 Comparisons between treatment groups

The literature suggests that the SWN is sensitive to the subjective effects of antipsychotic medication. The total score was used in the following comparisons: patient group vs. control group, antipsychotic treated patients vs. non-antipsychotic treated patients, typical antipsychotic treated patients vs. atypical antipsychotic treated patients.

3.7.1.2.1 Patient group vs. control group

	Minimum	Maximum	Median
Patient group	27	176	120
Control group	109	185	173

The difference between the groups is highly significant ($p = 0.000$).

3.7.1.2.2 Antipsychotic treated patients vs. non-antipsychotic treated patients

	Minimum	Maximum	Median
Antipsychotic treated	56	176	120
Non-antipsychotic treated	27	171	105

The groups do not differ significantly.

3.7.1.2.3 Typical antipsychotic treated patients vs. atypical antipsychotic treated patients

	Minimum	Maximum	Median
Typical antipsychotic treated	60	165	114
Atypical antipsychotic treated	56	176	128

The groups do not differ significantly.

3.7.1.3 Correlations with BDI

The lack of discriminatory power indicated by the previous analysis may signify a lack of sensitivity in the SWN, i.e. that the SWN is a poor measure, or that the SWN is a measure of some phenomenon other than that which it purports to measure. The relationship between the SWN results and those of the BDI was examined.

	Coefficient	Significance
SWN total	-0.808	0.000

This relationship holds true for all the groups examined above though it is weaker in some groups than others.

	All cases (n = 55)	Antipsychotic treated patients (n = 28)	Non-antipsychotic treated patients (n = 10)
SWN total	r =-0.808 p = 0.000	r =-0.699 p = 0.000	r =-0.945 p = 0.000

It should be noted that this high degree of correlation holds true for all the sub-scales of the SWN, supporting the decision to focus on the total score.

	All cases (n = 55)	Antipsychotic treated patients (n = 28)	Non-antipsychotic treated patients (n = 10)
Emotional regulation	r =-0.700 p = 0.000	r =-0.560 p = 0.002	r =-0.923 p = 0.000
Self-control	r =-0.741	r =-0.550	r =-0.865

	p = 0.000	p = 0.002	p = 0.001
Mental functioning	r =-0.740 p = 0.000	r =-0.532 p = 0.004	r =-0.742 p = 0.014
Social integration	r =-0.729 p = 0.000	r =-0.647 p = 0.000	r =-0.774 p = 0.009
Physical functioning	r =-0.759 p = 0.000	r =-0.635 p = 0.000	r =-0.703 p = 0.023

3.7.1.4 Summary

A very high degree of correlation exists between all the sub-scales of the SWN and the total score. The high degree of covariance between the scales makes it appropriate to use only the total score for analysis. Comparison of different groups indicates that subjective well-being is lower in the patient group than the control group. However, no differences were found between antipsychotic and non-antipsychotic patients or between typical antipsychotic and atypical antipsychotic treated patients. A very high degree of correlation was found between scores on the SWN and BDI scales, suggesting that the construct tapped by the SWN is very similar to that measured by the BDI.

3.7.2 Visual analogue scales

Three visual analogue scales were used, for subjective sensations of sedation, slowing, and restlessness respectively.

3.7.2.1 Results of the visual analogue scales in the patient group only

	Minimum	Maximum	Mean
Sedation	0.00	10.00	7.10
Slowing	1.10	10.00	6.55
Restlessness	1.00	10.00	7.19

Results on all three scales demonstrate a wide range of scores between the minimum and maximum points on the scales. In each case the mean score is towards the unimpaired end of the scale. The scores are not normally distributed.

3.7.2.2 Evaluation of the visual analogue scales as measures of the objective signs of parkinsonism

3.7.2.2.1 Correlation between the restlessness scale and the TAKE akathisia item criterion

N	Coefficient	Significance
42	-0.276	0.077

3.7.2.2.2 Correlation between the restlessness scale and the ESRS akathisia item criterion

N	Coefficient	Significance
42	-0.035	0.824

3.7.2.2.3 Correlation between the slowing scale and the TAKE bradykinesia item criterion

N	Coefficient	Significance
42	-0.251	0.109

3.7.2.2.4 Correlation between the slowing scale and the ESRS bradykinesia item criterion

N	Coefficient	Significance
42	-0.274	0.079

3.7.2.2.5 Correlation between the slowing scale and the ESRS group of bradykinesia-related items criterion

N	Coefficient	Significance
42	-0.295	0.0573

3.7.2.2.6 Correlation between the slowing scale and the Jebsen Hand Function Test objective measure of slowing

N	Coefficient	Significance
40	-0.111	0.495

3.7.2.2.7 Summary of the visual analogue scales as measures of objective signs

The visual analogue scales for slowing and restlessness perform very poorly as indicators of objective signs of the features they purport to measure. This finding may indicate that the scales are inaccurate measures. Alternatively it may indicate that the phenomena measured by the visual analogue scales exhibit little association

with observer ratings (or objective measures) of slowing, sedation or restlessness, respectively. In other words that they do not measure the features they are intended to measure.

3.7.2.3 Associations between the visual analogue scales and other measures of subjective state

It is known that psychomotor slowing may be a feature of depression, and that feelings of restlessness may also feature in depression and anxiety so the relationship between the visual analogue scales and the BDI criterion was investigated.

3.7.2.3.1 Descriptive data for the BDI criterion

	Minimum	Maximum	Median
BDI	0	51	9

3.7.2.3.2 Correlation between the visual analogue scales and the BDI criterion

The results obtained using all three visual analogue scales were compared with those from the BDI criterion.

	Slowing	Restlessness	Sedation
All cases (n = 57)	r = -0.736 p = 0.000	r = -0.627 p = 0.000	r = -0.580 p = 0.000
Patient group only (n = 40)	r = -0.690 p = 0.000	r = -0.603 p = 0.000	r = -0.590 p = 0.000

3.7.2.3.3 Associations between the visual analogue scales

On the basis of the strength of correlation between all three visual analogue scales and the BDI criterion, the extent of correlation between the three scales was calculated.

	Sedation vs. Slowing	Sedation vs. Restlessness	Slowing vs. Restlessness
All cases	r = 0.559 p = 0.000	r = 0.445 p = 0.000	r = 0.531 p = 0.000
Patient group only	r = 0.428 p = 0.005	r = 0.302 p = 0.052	r = 0.398 p = 0.009

3.7.2.3.4 Summary of the evaluation of the visual analogue scales

Strong and highly significant correlations are found between the three visual analogue scales and the BDI. The association is strongest for the slowing scale and weakest for the sedation scale. The correlations tend to be stronger when all cases are considered than when the patient group only is considered. The consistently low scores on all three visual analogue scales and on the BDI in the control group may tend to falsely inflate the strength of correlation though the associations appear robust.

3.7.2.4 Visual analogue scales as measures of medication effects

The effects of treatment group on subjective ratings of sedation, slowing and restlessness were examined.

3.7.2.4.1 Medication effects on subjective sedation

Using analysis of variance, ratings of sedation were higher in the overall patient group ($p = 0.001$) and in most patient medication groups than in the control group. The patient medication groups did not differ significantly from each other.

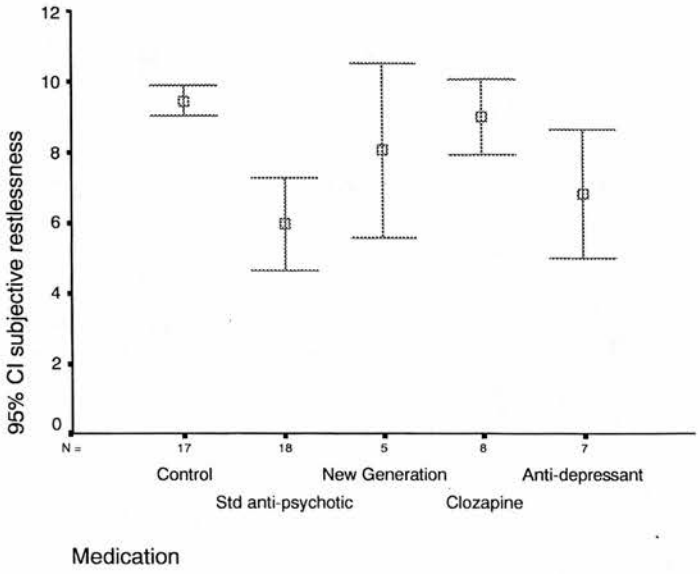
3.7.2.4.2 Medication effects on subjective slowing

Results from the slowing scale were very similar to those from the sedation scale, ratings of slowing being higher in the overall patient group ($p = 0.001$) and in most patient medication groups than in the control group. The patient medication groups did not differ from each other.

3.7.2.4.3 Medication effects on subjective restlessness

The results from the restlessness scale differed from those of the other two visual analogue scales. Though the overall patient group differed from the control group ($p = 0.000$), this was not true of all the patient medication groups. Ratings of restlessness in the clozapine group were not higher than those in the control group, and were significantly lower than those in the typical antipsychotics group ($p = 0.004$). The other patient medication groups did not differ from each other.

3.7.2.4.4 Error bars for group means of subjective restlessness in medication groups



3.7.2.4.5 Summary of the visual analogue scales as measures of medication effects

Mean ratings of sedation, slowing, and restlessness are higher in the patient group than in the control group. On the sedation and slowing scales no differences between patient medication groups were found. However, ratings of restlessness in the clozapine group do not differ from the control group and are significantly lower than in the typical antipsychotics group.

3.7.3 Summary

The results from the visual analogue scales indicate that subjective sensations of slowing and restlessness show little association with observer-ratings or instrumented measures of objective slowing. In contrast, subjective slowing was associated with depression. All patients exhibited elevated levels of subjective slowing and sedation relative to the control group. All patients except the clozapine treated medication

group exhibited elevated levels of subjective restlessness relative to the control group; levels of restlessness in the clozapine group were significantly lower than in the typical antipsychotic group.

3.8 Single case serial assessment

One patient (AG) was evaluated on six occasions over five months. During this period medication was altered (drug changed from risperidone to quetiapine and then dosage of quetiapine reduced) with the aim of reducing levels of parkinsonism. At each visit a full clinical assessment was performed and the instrumentation was used to produce an objective assessment of parkinsonism. Subjective rating scales were not used at all visits due to time constraints.

The use of a longitudinal design may allow the instrumentation to be assessed without the confounding effects that individual differences may have when a cross-sectional design is used. To this end, many of the correlations between clinical rating criteria and instrumentation measures which were calculated in previous sections using cross-sectional data from a group of patients will be calculated using the longitudinal follow-up data from this series of visits.

3.8.1 Bradykinesia

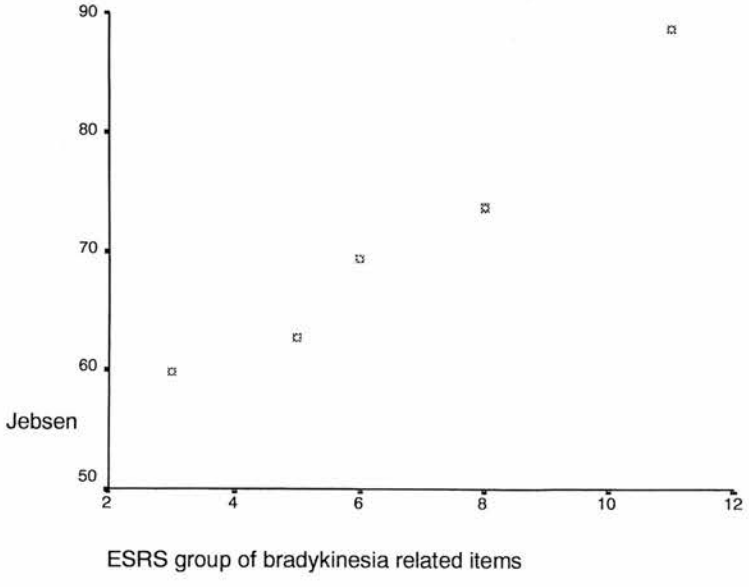
3.8.1.1 Correlation coefficients for the relationships between clinical rating criteria and instrumental measures

	r	p
Jebsen vs. TAKE bradykinesia item	0.728	0.087
Jebsen vs. ESRS group of	0.979	0.000

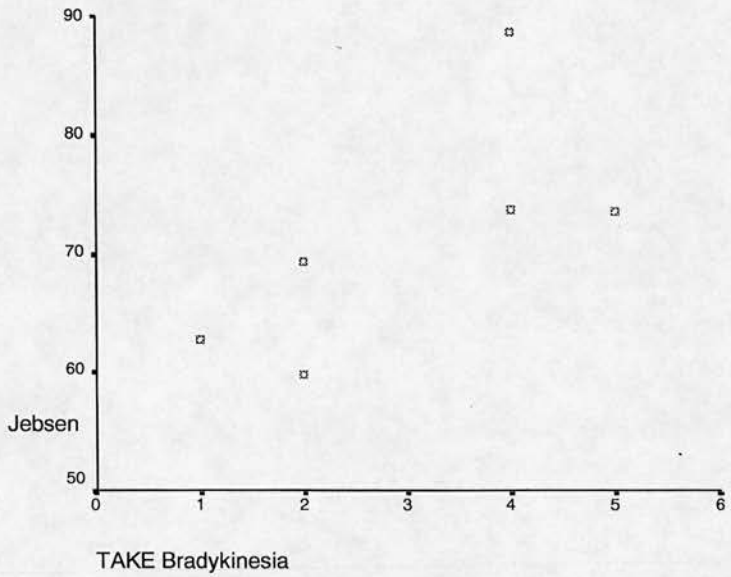
bradykinesia related items		
SRT vs. TAKE bradykinesia item	0.903	0.008
SRT vs. ESRS group of bradykinesia related items	0.815	0.037
CRT vs. TAKE bradykinesia item	0.612	0.183
CRT vs. ESRS group of bradykinesia related items	0.183	0.005

Very strong and highly significant correlations are found between Jebsen total time and the ESRS group of bradykinesia related items, between SRT and both the TAKE bradykinesia item and the ESRS group of bradykinesia related items, and between CRT and the ESRS group of bradykinesia related items.

3.8.1.2 Scatterplot of Jebsen time against ESRS group of bradykinesia related items



3.8.1.3 Scatterplot of Jebsen time against TAKE bradykinesia item



The correlations between the rating criteria and the measures of bradykinesia are stronger in the serial assessment than in the cross-sectional analysis. Though the number of observations was small, the significance levels were still high.

3.8.2 Rigidity

Very weak and non-significant correlations are found between the activation ratio and the rating criteria. This is true for both the mean activation ratio using the TAKE rigidity item and the individual arm activation ratios using the ESRS upper limb rigidity items. The correlations are of similar strength to those found in the earlier cross-sectional analysis, though with the limited number of observations these levels of correlation are non-significant.

3.8.3 Tremor

The pattern of associations found (and not found) in the cross-sectional analysis are repeated here. Clinical ratings of tremor are closely related to postural tremor amplitude

	r	p
Postural tremor amplitude (RH) vs. ESRS tremor (RH)	0.670	0.131
Postural tremor amplitude (LH) vs. ESRS tremor (LH)	0.828	0.032

As in the cross-sectional analysis, resting tremor frequency ratio is not a valid indicator of parkinsonism. However, an unexpected relationship was found between overall severity of parkinsonism and tremor characteristics. In this patient, severity of parkinsonism was inversely related to postural tremor amplitude ($r = 0.900$, $p = 0.009$).

No changes in tremor frequency composition associated with changes in severity of parkinsonism were found. Rather, tremor amplitude in the low and high frequency bands tended to correlate. This was particularly in true in posture (right hand $r = 0.848$, $p = 0.026$; left hand $r = 0.864$, $p = 0.033$) though the relationship neared significance at rest too (right hand $r = 0.750$, $p = 0.086$; left hand $r = 0.695$, $p = 0.125$).

3.8.4 Motor planning impairment

Motor planning time exhibits a tendency to increase during the course of the follow-up, indicating a reduction in parkinsonism. However, the result from visit six does not fit this pattern. If the observations from this visit are excluded from the analysis, then the motor planning variable tends to show a correlation with both the visit order and with total Jebsen time (that with visit is positive, and that with Jebsen time negative).

	r	p
Motor planning vs. visit	0.700	0.165
Motor planning vs. Jebsen	-0.857	0.045

3.8.5 Subjective experience

As the subjective ratings were not used at all visits there are too few observations for meaningful analysis.

3.8.6 Summary

The relationships between instrumental measures and clinical ratings of bradykinesia were much stronger than those found in the cross-sectional analysis. Despite the small number of observations some associations, particularly that between Jebsen

time and the ESRS group of bradykinesia related items, reached a high level of significance. The associations between measures and ratings of rigidity were non-significant. Though weak in strength the coefficients were of similar magnitude to those found in the cross-sectional analysis. The tremor results also closely parallel those found in the cross-sectional analysis. No evidence was found of a parkinsonian slow resting tremor. Tremor frequency composition remained relatively constant at all visits, with amplitude in both frequency ranges increasing as the severity of parkinsonism decreased. Motor planning time was found to increase during course of the follow-up. This change, indicative of decreasing parkinsonism, is associated with declines in ESRS total score and objectively assessed bradykinesia.

3.9 Medication-associated tremor characteristics

Clinical ratings of tremor indicate the presence of increased tremor amplitude in a majority of patients relative to the control group. However, these ratings are limited in the information they provide. In particular, they do not distinguish tremors with different frequency compositions.

Previous sections (3.4 and 3.5) have considered the results from the tremor instrumentation as they may be used to evaluate the instrumentation procedures. These data may also provide information concerning the characteristics of the tremors associated with different forms of medication. This section considers the tremor characteristics of the different patient medication groups in terms of the amplitudes of postural and resting tremor in low and high frequency ranges.

It should be noted that in all the comparisons, the variance of the patient medication groups is high. This, with the very small numbers in some groups ensures that many effects of group do not reach statistical significance, and that the likelihood of type II errors (rejection of positive finding) is elevated.

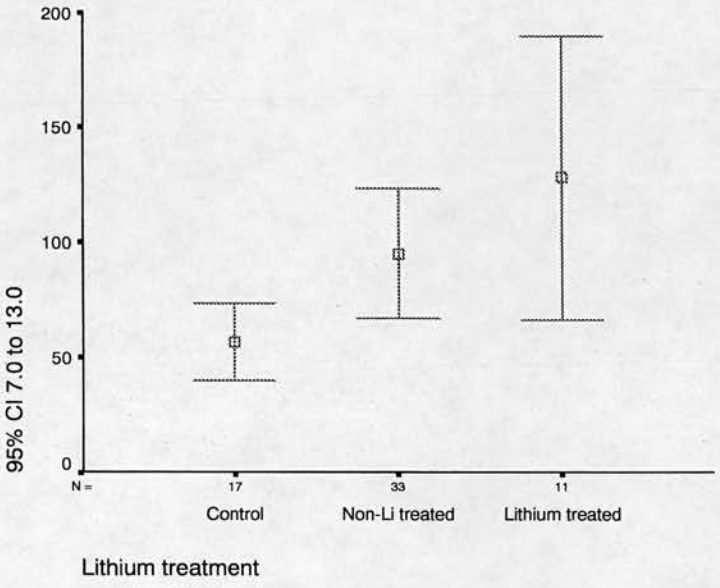
3.9.1 Lithium tremor

Increased amplitude of high frequency tremor is recognised as a side-effect of lithium treatment. The first comparisons used three groups: Lithium-treated patient group, non-Lithium treated patient group, control group.

3.9.1.1 Postural tremor (R)

A significant effect of group was found in the 7.0-13.0 Hz range ($p = 0.038$). Amplitude was greatest in the lithium group in all cases. The lithium and control groups differed significantly ($p = 0.039$).

3.9.1.1.1 Error bars for group mean postural tremor amplitude (R) in the 7.0-13.0 Hz range



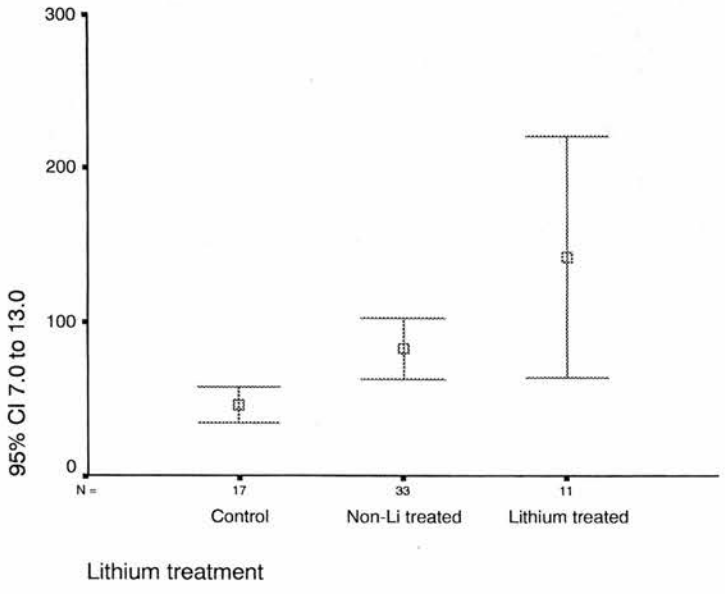
3.9.1.1.2 Post hoc comparisons for group mean postural tremor amplitude (R)

	7.0-13.0 Hz
Lithium vs. non-Lithium patients	NS
Lithium vs. Control group	p = 0.039

3.9.1.2 Postural tremor (L)

The effect of group was significant in the 7.0-13.0 Hz ($p = 0.001$) range, and there was a trend towards significance in the 3.0-7.0 Hz ($p = 0.059$) range; greater amplitude was found in the lithium group. Amplitude was greater in the lithium group than the control group in both high ($p = 0.001$) and low ($p = 0.059$) frequency ranges; it was greater in the lithium group than in the non-lithium patient group in the high frequency range ($p = 0.031$) only.

3.9.1.2.1 Error bars for group mean postural tremor amplitude (L) in the 7.0-13.0 Hz range



3.9.1.2.2 Post hoc comparisons for group mean postural tremor amplitude (L)

	3.0-7.0 Hz	7.0-13.0 Hz
Lithium vs. non-Lithium	NS	p = 0.031

patients		
Lithium vs. Control group	p = 0.059	p = 0.001

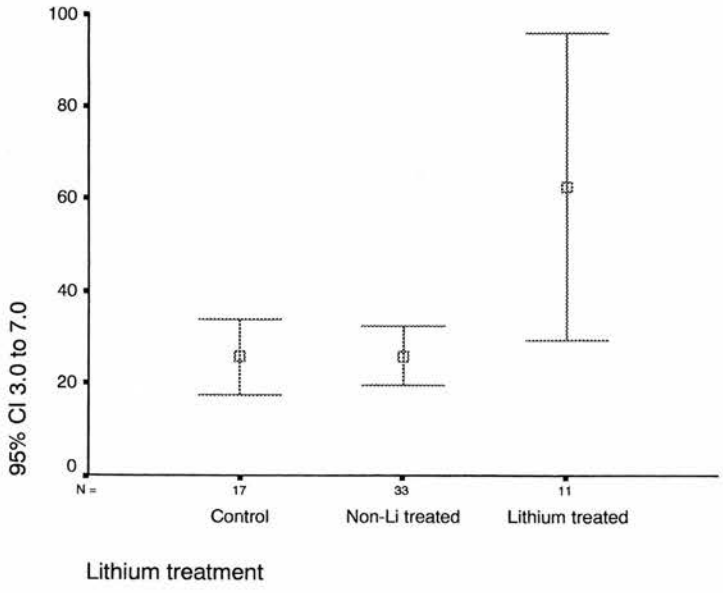
3.9.1.3 Resting tremor (R)

No significant effects of medication group were found in any of the variables tested.

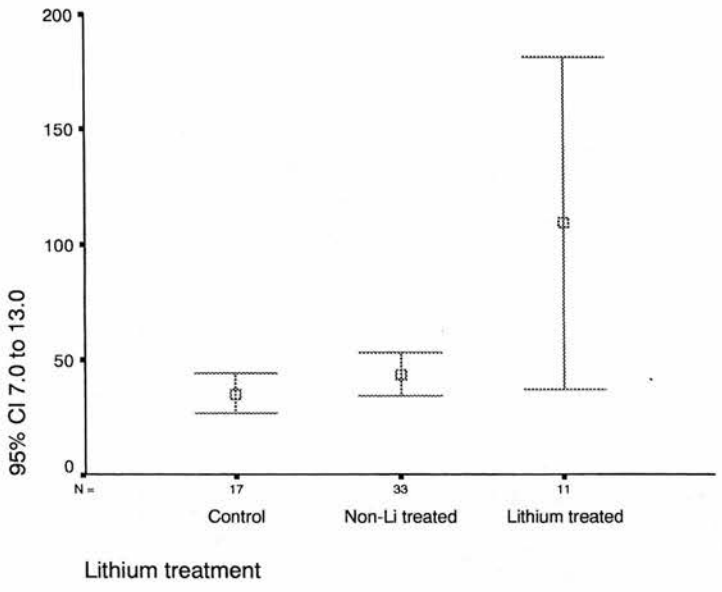
3.9.1.4 Resting tremor (L)

Significantly greater amplitude was found in the lithium group than in the other groups in both low frequency and high frequency ranges (both p = 0.000).

3.9.1.4.1 Error bar for group mean resting tremor amplitude (L) in the 3.0-7.0 Hz range



3.9.1.4.2 Error bar for group mean resting tremor amplitude (L) in the 7.0-13.0 Hz range



3.9.1.4.3 Post hoc comparisons for group mean resting tremor amplitude (L)

	3.0-7.0 Hz	7.0-13.0 Hz
Lithium vs. non-lithium patients	p = 0.002	p = 0.001
Lithium vs. Control group	p = 0.000	p = 0.001

3.9.1.5 Summary of lithium tremor characteristics

Lithium treatment is associated with a significant increase in tremor magnitude. This increase was found in postural tremor in both hands, and at rest in the left hand only. The increase is present predominantly in the higher frequency range (7.0-13.0 Hz) though there is evidence of an increase in magnitude in the 3.0-7.0 Hz range.

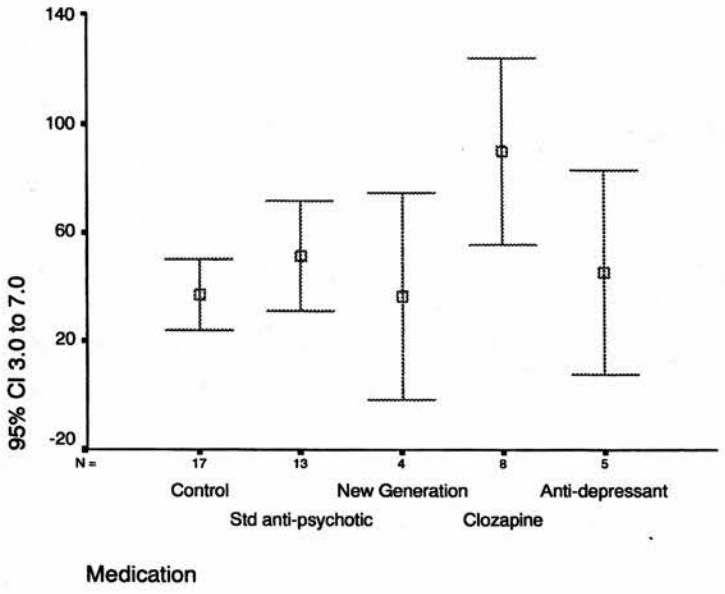
3.9.2 Other patient medication groups

Comparisons of the tremor characteristics in patient medication groups were made after the exclusion of all patients receiving lithium.

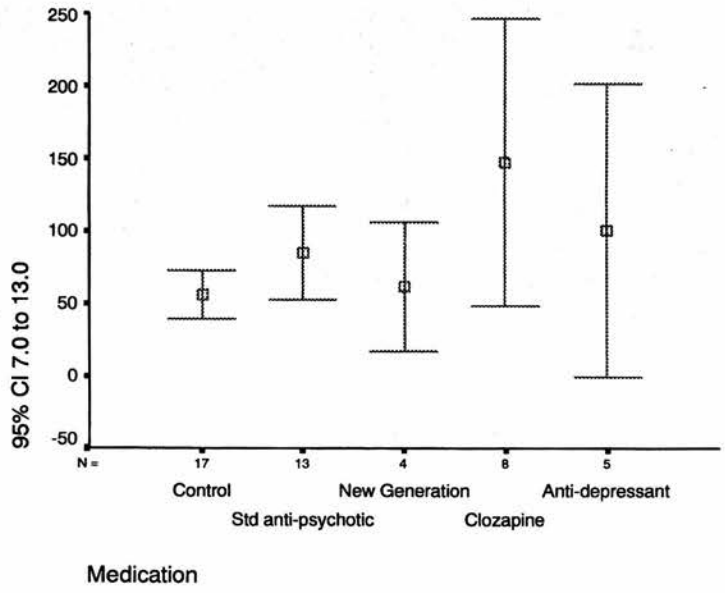
3.9.2.1 Postural tremor (R)

Significant effects of medication group were found in both low and high frequency ranges: 3.0-7.0 Hz ($p = 0.006$), 7.0-13.0 Hz ($p = 0.036$). In all comparisons amplitude is greatest in the clozapine group. The clozapine group differed significantly from the control group in both frequency ranges. The results showed trends towards greater amplitude in the clozapine group relative to the other patient medication groups in the 3.0-7.0 Hz frequency range.

3.9.2.1.1 Error bar of medication group mean postural tremor amplitude (R) in the 3.0-7.0 Hz range



3.9.2.1.2 Error bar of medication group mean postural tremor amplitude (R) in the 7.0-13.0 Hz range



3.9.2.1.3 Post hoc comparisons for medication group mean postural tremor amplitude (R)

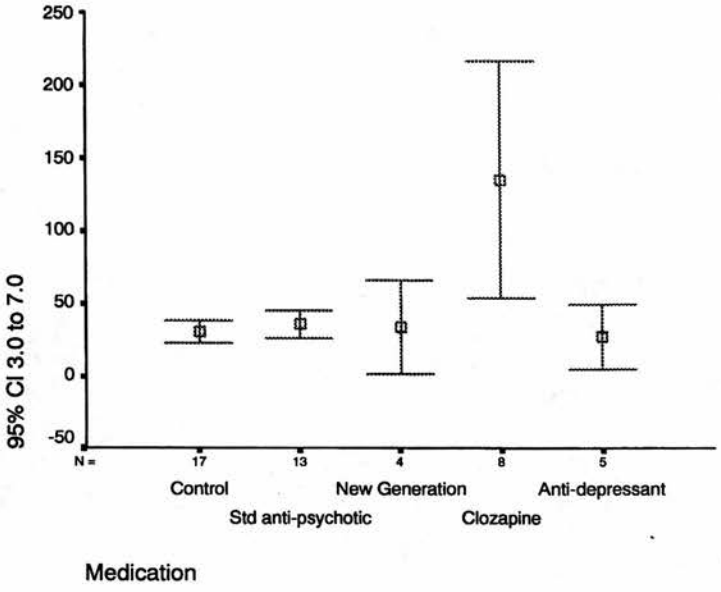
	3.0-7.0 Hz	7.0-13.0 Hz
Clozapine group vs. control group	p = 0.003	p = 0.023
Clozapine group vs. typical antipsychotic group	NS (p = 0.092)	NS
Clozapine group vs. New Generation group	NS (p = 0.080)	NS
Clozapine group vs. antidepressant group	NS (p = 0.162)	NS

The clozapine group exhibits low frequency tremor of greater amplitude than any of the other groups. Overall tremor amplitude in this group is greater than in the control group and tends to be greater than in the other antipsychotic groups.

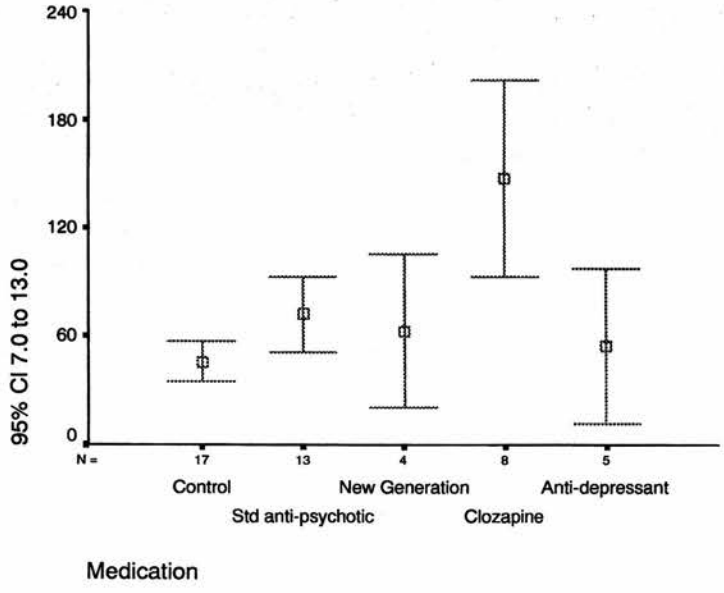
3.9.2.2 Postural tremor (L)

Significant effects of group on tremor amplitude were found in both frequency ranges: 3.0-7.0 Hz (p = 0.000) and 7.0-13.0 Hz (p = 0.000). In both ranges the clozapine group exhibited significantly greater tremor amplitude than all other groups.

3.9.2.2.1 Error bar of medication group mean postural tremor amplitude (L) in the 3.0-7.0 Hz range



3.9.2.2.2 Error bar of medication group mean postural tremor amplitude (L) in the 7.0-13.0 Hz range



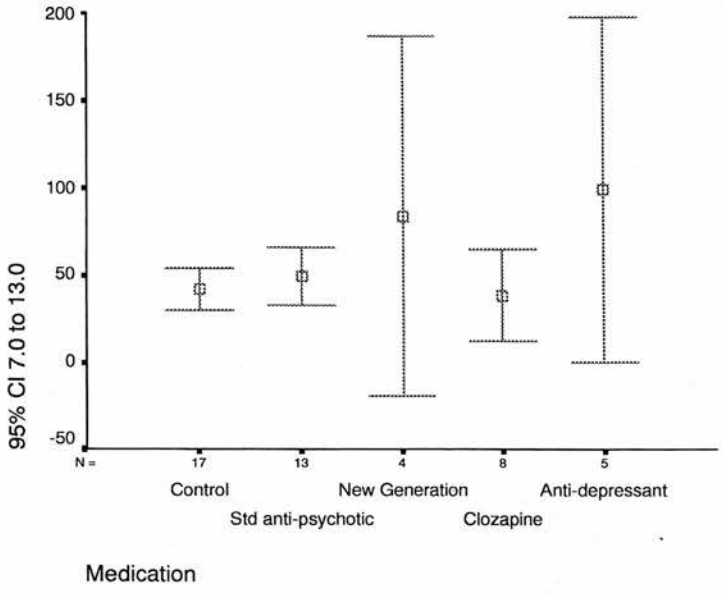
3.9.2.2.3 Post hoc comparisons for medication group mean postural tremor amplitude (L)

	3.0-7.0 Hz	7.0-13.0 Hz
Clozapine group vs. control group	p = 0.000	p = 0.000
Clozapine group vs. typical antipsychotic group	p = 0.000	p = 0.001
Clozapine group vs. New Generation group	p = 0.003	p = 0.007
Clozapine group vs. antidepressant group	p = 0.001	p = 0.001

3.9.2.3 Resting tremor (R)

None of the effects of medication group on resting tremor amplitude (R) found reached significance, however some showed trends towards significance. Differences were found, in the high frequency range, between the antidepressant group and the clozapine, typical antipsychotic, and control groups; amplitude was greater in the antidepressant group in all comparisons.

3.9.2.3.1 Error bar of medication group mean resting tremor amplitude (R) in the 7.0-13.0 Hz range



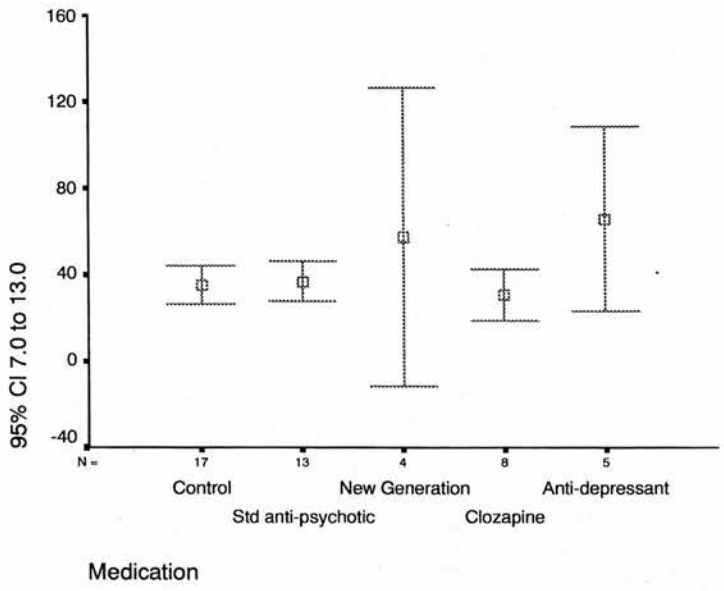
3.9.2.3.2 Post hoc comparisons for medication group mean resting tremor amplitude (R)

	7.0-13.0 Hz range
Antidepressant group vs. clozapine group	NS (p = 0.084)
Antidepressant group vs. typical antipsychotic group	NS (p = 0.178)
Antidepressant group vs. control group	NS (p = 0.056)

3.9.2.4 Resting tremor (L)

A significant effect of medication group on resting tremor amplitude (L) was found in the high frequency range (p = 0.024). Amplitude was highest in the antidepressant group, the group mean showing trends towards significant differences with the control group and the clozapine group.

3.9.2.4.1 Error bar of medication group mean resting tremor amplitude (L) in the 7.0-13.0 Hz range



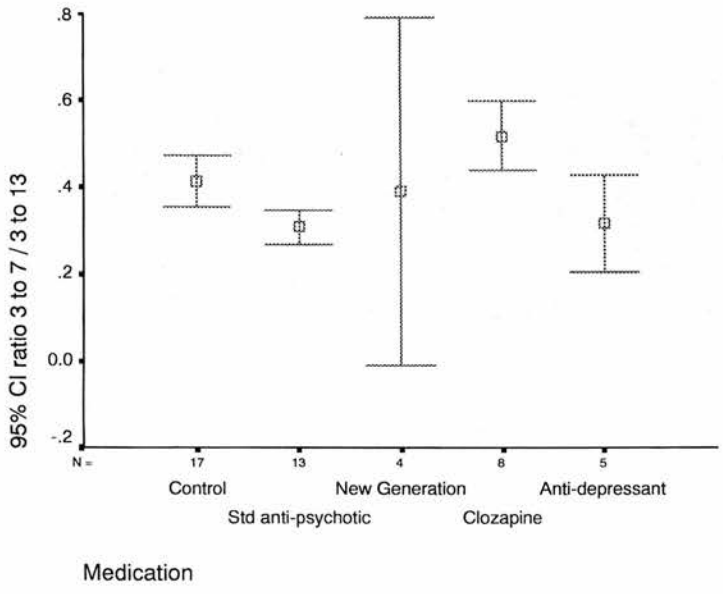
3.9.2.4.2 Post hoc comparisons for medication group mean resting tremor amplitude (R)

	7.0-13.0 Hz range
Antidepressant group vs. control group	NS (p = 0.076)
Antidepressant group vs. typical antipsychotic group	NS (p = 0.135)
Antidepressant group vs. clozapine group	NS (p = 0.061)

3.9.2.5 Proportion of low frequency resting tremor

When frequency ratios were examined directly, effects of medication group were found only in resting tremor in the left hand. In this comparison, proportion of low frequency tremor is higher in the clozapine group than in the antidepressant group and the typical antipsychotic group. There is a slight trend towards a lowered proportion of low frequency tremor in the typical antipsychotic group relative to the control group.

3.9.2.5.1 Error bars for group mean proportion of low frequency resting tremor (L)



3.9.2.5.2 Post hoc comparisons for the proportion of low frequency resting tremor (L)

	Low frequency resting tremor
Typical antipsychotic group vs. control	NS (p = 0.176)

group	
Clozapine group vs. typical antipsychotic group	p = 0.002
Clozapine group vs. antidepressant group	p = 0.038

3.9.3 Summary of medication-associated tremors

3.9.3.1 Typical antipsychotic group

The typical antipsychotic group did not exhibit any significant differences in tremor properties relative to the control group. There is a trend towards a lowered proportion of low frequency resting tremor in the left hand though this apparent finding may actually reflect a non-significant increase in high frequency activity in this condition. The error bars appear to indicate that high frequency postural tremor amplitude may be slightly increased though this was not statistically significant.

3.9.3.2 Lithium

Lithium treatment is associated with a significant increase in tremor amplitude. This increase was found in postural tremor in both hands, and at rest in the left hand only. The increase is present predominantly in the high frequency range though there is evidence of an increase in magnitude in the lower range.

3.9.3.3 Clozapine group

Postural tremor amplitude was increased; the elevation was seen in both the low and high frequency ranges in the left hand, but only in the low frequency range in the right hand. Weaker evidence was found of increased resting tremor amplitude. There is also some evidence of a relatively increased proportion of low frequency activity in resting tremor.

3.9.3.4 Antidepressant group

Relative to the control group, resting tremor in the antidepressant group was increased in amplitude; this increase occurred predominantly in the high frequency range. However, patient numbers in this medication group were very small so the results cannot be regarded as wholly reliable.

3.10 Other patient medication group comparisons

In the previous section the tremor instrumentation was used in the comparison of the patient medication groups with the control group. This section reports the results of using the other measures (instrumental measures and subjective ratings) to compare the same groups. A large number of comparisons were made in this investigation. However, the majority of findings were of non-significant differences. Only those findings which were felt to be of interest are reported. These include those in which significant differences were found between patient medication groups, and those in which patient groups did not differ from the control group.

3.10.1 Bradykinesia – Jebsen hand function test

3.10.1.1 Group mean total score

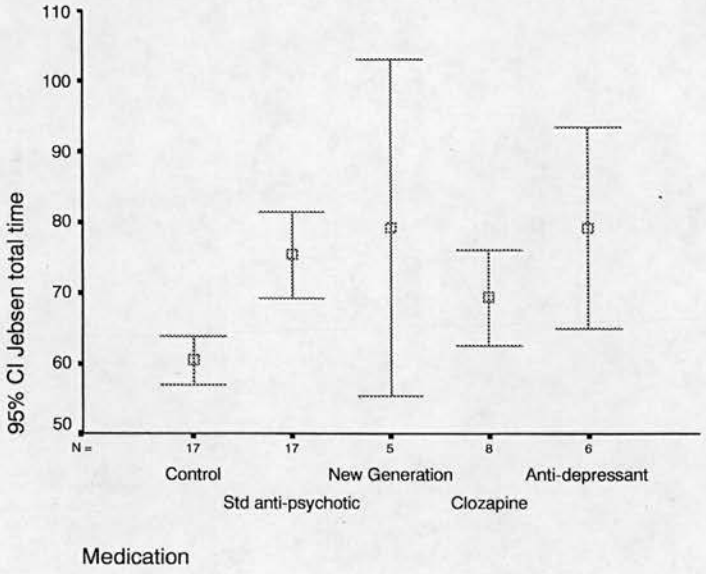
Group	Mean	SD
Control	60.39	6.95
Typical antipsychotic	75.37	11.94
New Generation AP	79.28	19.20
Clozapine	69.36	8.10
Antidepressant	79.29	13.63

3.10.1.2 Statistical comparisons

There was an overall highly significant effect of medication group ($p = 0.000$) on Jebsen test total time. Post-hoc comparisons reveal that total time was significantly higher than in the control group in the typical antipsychotic, New Generation

antipsychotic, and antidepressant patient groups. The clozapine group was not significantly slower than the control group.

3.10.1.3 Error bars for medication group mean Jebsen test total time



3.10.1.4 Comparison of medication groups using observer ratings of bradykinesia

In a further effort to evaluate the accuracy of the Jebsen test, the comparison of medication groups was repeated using observer-ratings of bradykinesia. As ratings were not performed in the control group it was decided for the purposes of this analysis to attribute ratings of zero on all items to all members of the control group.

Single item ratings of bradykinesia (TAKE and ESRS bradykinesia items) did not demonstrate significant overall effects of medication group. However, a significant effect was seen on ratings using the ESRS group of bradykinesia-related items ($p <$

0.05). A multiple comparisons procedure revealed that the only significant between-groups difference was between the typical antipsychotic group and the control group. As with the results of the Jebsen test the clozapine group was not found to be slowed relative to the control group.

3.10.1.5 Summary

When the Jebsen test is evaluated against observer-ratings of bradykinesia on its ability to detect effects of medication group on bradykinesia it exhibits greater accuracy than single-item ratings taken from both the TAKE and ESRS scales. Accuracy is comparable to that obtained by using the group of bradykinesia-related items from the ESRS scale.

3.10.2 Subjective sensations of restlessness

A significant effect of medication group was found on subjective ratings of sensations of restlessness.

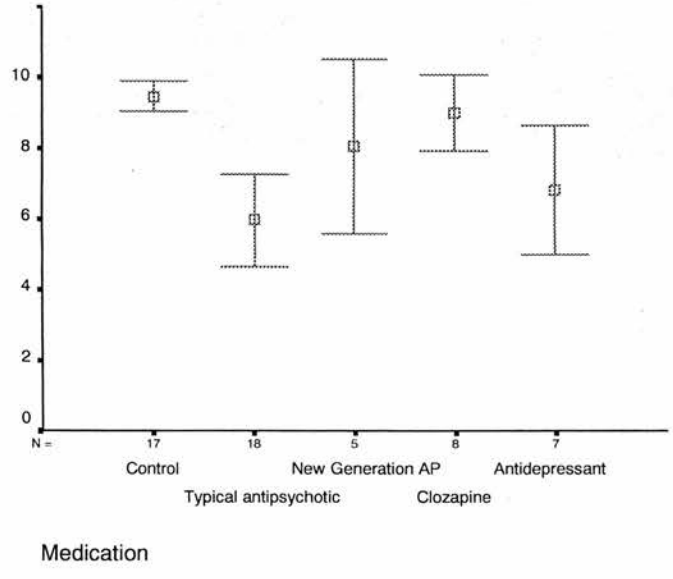
3.10.2.1 Group mean ratings of restlessness

Group	Mean rating of restlessness	SD
Control	9.5	0.8
Typical antipsychotic	6.0	2.6
New Generation antipsychotic	8.1	2.0
Clozapine	9.0	1.3
Antidepressant	6.8	2.0

3.10.2.2 Statistical comparisons

The overall effect of medication group was highly significant (p = 0.000 by ANOVA). Using post-hoc tests, subjective ratings of restlessness were significantly higher (indicated by lower scores) in both the typical antipsychotic and antidepressant groups than the control group (p = 0.000 and p = 0.031, respectively). Ratings of restlessness were higher in the typical antipsychotic group than in the clozapine group (p = 0.004); ratings in the clozapine group did not differ from those in the control group.

3.10.2.3 Error bars for medication group mean ratings of sensations of restlessness



3.10.3 Summary

The two notable comparisons described here both concerned the clozapine group contrasted with the typical antipsychotic group. In the first, the clozapine group was found to be not significantly slowed relative to the control group. All other patient medication groups were significantly slower than the control group. In the second, mean ratings of restlessness were found to be lower in the clozapine group than in the typical antipsychotic group; ratings in the clozapine group did not differ from those in the control group. Other comparisons made, such as those using the rigidity instrumentation did not demonstrate any noteworthy results.

4 Discussion

This section will assess the evidence within the context of the hypotheses presented at the end of the introduction (1.4):

1. Instrumentation has a role in the assessment of DIP
2. Bradykinesia is the predominant feature of DIP
3. Cognitive and subjective features of parkinsonism are present in DIP.

4.1 Instrumentation has a role in the assessment of DIP

The methods by which this hypothesis may be demonstrated were discussed earlier (1.4). It is necessary to show that the measures are accurate, and that they are valid measures of the features they purport to measure. Validity may be demonstrated relative to existing criteria (observer-ratings in this case) or by reference to theoretical constructs. To define a role for instrumentation, these measures must demonstrate not only validity as a means of assessment, but some advantage over observer-rating methods.

The instrumentation methods will first be considered individually, as assessors of particular features of parkinsonism. The validity of the measures will be considered relative to both the observer-rating criteria and the theoretical constructs underlying the procedures. Finally, the role of instrumentation may be considered independently of specific methods.

4.1.1 Jebsen Hand Function Test

This section will commence with an evaluation of the properties of the Jebsen Hand Function Test. Face validity of the Jebsen test, as of all performance measures, is high: the test appears to be a measure of movement speed and motor control. However, though total time taken to complete the Jebsen test correlates significantly with observer-rating criteria for bradykinesia (TAKE bradykinesia item, ESRS bradykinesia item, ESRS group of bradykinesia-related items), the association is of only moderate strength. It should be noted that in the single-case study, the degree of association found over a number of assessment visits was much greater than in the overall cross-sectional study. The sensitivity of the Jebsen test as a marker of bradykinesia was good to very good, being highest with the ESRS group of bradykinesia-related items. However specificity was low, being lowest with the ESRS group of bradykinesia-related items. This demonstration of criterion validity indicates that the Jebsen test functions as a measure of bradykinesia, though overall accuracy is not particularly high. As discussed in the introduction (1.4.1.3), an apparently lower level of accuracy in the measure than the criteria may indicate that the level of accuracy in the measure is actually higher than in the criterion.

In a supplementary evaluation of the Jebsen test, the ability of the test to distinguish medication groups was contrasted with that of the observer-ratings (3.10.1.4). The Jebsen test identified differences between medication groups not detected by single-item ratings of bradykinesia from the TAKE or ESRS scales. The accuracy of the Jebsen test was comparable to that obtained using the ESRS group of bradykinesia-related items.

It is necessary to address the relatively low degree of correlation between the Jebsen test and the observer-rating criteria. This finding may be accounted for by an examination of the constructs tapped by the measure and the criteria respectively. All of the criteria scores represent ratings of the degree of abnormality present in the patient on a particular axis at the time of assessment. In this case, this is an assessment of the severity of bradykinesia relative to a baseline (i.e. non-bradykinetic) state of normality for that particular patient. The rating is thus independent of individual differences in baseline speed of movement. However, the Jebsen test results represent a measurement of absolute speed of movement; this figure is not a measure of slowing as it cannot take into account the individual's baseline state. Logically, it would be expected that absolute speed of movement and degree of bradykinesia would correlate positively, but that the relationship would not be one of perfect correlation when the data are taken from a group of patients. In contrast, the two variables should correlate perfectly (or almost so) if all the data are from a single patient. In this situation, the single case study, the results from the Jebsen test do form an assessment of the degree of slowing relative to other visits, and the correlation coefficients for the relationships between the Jebsen test results and the observer-rating criteria are extremely high.

Thus it may be argued that the construct tapped by the Jebsen test (and by inference other performance measures too) is bradykinesia only in the case of a series of follow-up assessments of the same patient. In a one-off assessment the construct tapped is not bradykinesia, though it is closely related to bradykinesia.

The advantages of the Jebsen test over observer-ratings arise principally from its truly objective nature which removes any variation in assessments due to inter-rater differences. This characteristic contrasts with the observer-ratings which incorporate a degree of subjectivity into the assessments, despite the use of rating guidelines, and can only be valid when performed by a trained and experienced assessor. Further, the test demands little of the assessor conducting the measurements other than the ability to set up the apparatus, provide instructions and operate a stopwatch.

The results of the Jebsen test require little data analysis and are simple to interpret. However, the Jebsen test does require a fairly cumbersome array of equipment. Though the apparatus used is cheap and easily obtained, it is hardly portable. A further criticism of the Jebsen test is the time taken to complete the full battery of tasks, resulting primarily from the number of tasks included. The time taken may be excessive for inclusion of the test in regular assessments of extrapyramidal status, particularly in light of the fact that that assessment of bradykinesia forms only one part of an assessment of extrapyramidal status, itself only one component of a comprehensive clinical examination.

The construct tapped by the Jebsen test was discussed above in the context of the relationship between the Jebsen test results and the observer ratings of bradykinesia. This analysis indicated that the Jebsen test does provide a direct indicator of bradykinesia in serial follow-up assessments within a single patient, and an indirect indicator of bradykinesia when used for one-off assessments. It is also worth

considering the literature concerned with the bradykinesia construct. It is asserted that upper-body bradykinesia is the cardinal sign of parkinsonism (Quinn, 1995). This implies that the principal indicator of parkinsonism should be an assessment of upper-body bradykinesia. The use of a performance measure of upper-body bradykinesia such as the Jebsen test is consistent with this argument.

4.1.2 CANTAB

The properties of the CANTAB reaction time test as a measurement of reaction time are well established, and reaction time tests have been previously used as indicators of bradykinesia (Evarts et al., 1981). However, no evidence has been published of the CANTAB test itself being used in this fashion.

Examination of the relationships between the CANTAB variables used (SRT total time and movement latency, CRT total time and movement latency) and the observer-rating criteria for bradykinesia indicates a consistent pattern of results within the CANTAB variables. The SRT variables exhibit a moderate strength of correlation with the criteria, but the CRT variables are only weakly related to them. The sensitivity and specificity characteristics of the CANTAB tests as markers of bradykinesia reveal a similar level of overall accuracy to the Jebsen test in the CRT variables and a slightly higher level of overall performance in the SRT variables. In comparison with the Jebsen test, sensitivity is lower and specificity higher.

The issues discussed in the previous section concerning the construct validity of instrumental measures as markers of absolute speed of movement when contrasted

with observer-ratings of the degree of abnormality present are pertinent to the CANTAB instrumentation too. However, it is likely that these issues affect the simple and choice RT conditions equally, and this account does not explain the fact that the CRT results show a lower degree of correlation with the observer-ratings of bradykinesia than do the SRT results.

The CANTAB reaction time test shares many of the benefits of the Jebsen hand function test. The results are truly objective, removing any influence of rater error; the test is simple to administer; most patients found it easy to comply with. However there is the same reliance upon cumbersome equipment, though this time it is rather more expensive. Further, the CANTAB is similarly time-consuming for regular use.

Though the CANTAB provides, like the Jebsen, a measure of bradykinesia, the construct tapped is slightly different. The CANTAB movement latency provides a measure of pure speed of movement, and the reaction latency a measure of speed of motor planning and movement initiation. It is notable that all CANTAB variables (SRT reaction latency, SRT movement latency, CRT reaction latency, CRT movement latency) correlate very highly.

The movement latencies can be contrasted with the Jebsen variable which is a measure of overall performance including aspects of motor control and movement initiation. Though movement initiation is a factor in the CANTAB reaction latencies, there is an external trigger for these movements. In contrast, performance of the tasks in the Jebsen test requires a complex sequence of self-initiated goal-directed

movements. The CANTAB assessment may be viewed as assessing only a limited part of the bradykinesia complex.

4.1.3 Bradykinesia

It was noted earlier that it is not sufficient to demonstrate merely that the instrumental measures provide a valid means of assessing parkinsonism. The instrumentation must display advantages over clinical ratings. If not in terms of accuracy then in some other frame of reference, perhaps ease of use for the assessor, or by being less time-consuming. The advantages may be limited to certain situations or particular types of assessment but must be specified.

From the introduction (1.3.1.3.3) it was suggested that some instrumental measures may provide a sensitive means of detecting the early onset of EPS, by virtue of increased accuracy over observer ratings. The evidence presented does indicate that performance measures of movement speed can provide a highly accurate means of assessing the severity of bradykinesia within the individual. Further, the Jebsen test proved capable of identifying a clozapine treatment group which was not significantly slowed relative to the control group though all other patient groups were. This will be discussed in greater depth (4.5.2).

Though the procedures provide a valid means of assessing bradykinesia, they exhibit significant weaknesses. The measures do not display very high levels of sensitivity or specificity relative to the observer-rating criteria, and are not capable of functioning as one-off indicators of the presence of parkinsonism within the individual. Rather,

variations in baseline movement speed within normal and patient populations mask the minimal changes in performance due to very mild or sub-clinical bradykinesia.

Of greater promise is the very high level of validity demonstrated by the bradykinesia instrumentation in the single case study assessment. When all the observations made came from a single individual the level of correlation between the measures and the criteria was extremely high. This suggests that the benefit of bradykinesia instrumentation may be as an appropriate procedure for repeated assessments in the follow-up of individual patients on a longitudinal basis rather than for screening for the presence of bradykinesia within groups of patients.

Performance measures may thus be a means for monitoring the development and progression of bradykinesia without the need for a comprehensive examination by a doctor. They are relatively simple to administer and provide a truly objective assessment of the degree of slowing in the individual relative to a pre-treatment baseline measurement. The influence of inter-or intra-rater variability is wholly removed. Further, the assessment may be performed with the same level of accuracy and reliability without the need for a trained and experienced rater.

Within the individual, the Jebsen test proved sensitive to minor changes in condition, and it is likely that this property is not exclusive to this one particular performance measure. Though it was used in this study, the Jebsen test may not be the most practical measure for use in this fashion The criticisms made earlier, of both the Jebsen and CANTAB procedures, that they depend upon cumbersome, and in the

case of the CANTAB test expensive, equipment do not necessarily apply to all performance tests. Numerous performance measures appear in the literature, many of which use small portable pieces of equipment such as pegboards (1.3.2.1). It seems likely that many if not all of these measures may provide the same information gained from the Jebsen and CANTAB tests with a similar degree of accuracy. These measures may also be less time-consuming and thus more suitable for repeated assessments. Many of these measures could be adapted to incorporate integral automated timing devices to remove another source of confounding.

The importance of monitoring the development of bradykinesia, perhaps more so than other features of parkinsonism, is the subject of the other hypotheses and is discussed in depth (4.6).

4.1.4 Positive feedback device

The activation ratio variable derived from the results of the positive feedback device was found to correlate with the observer-rating criteria, though the association was of only moderate strength. The relationship was statistically significant for the mean activation ratio (TAKE rigidity criterion) though it was non-significant when the arms were considered separately (ESRS upper-limb rigidity criteria). As an indicator of the presence of rigidity, the activation ratio exhibited good levels of sensitivity and specificity, especially for the right arm alone.

The lack of significant correlation between the activation ratio and the ESRS rigidity item criteria raises doubts about the validity of the procedure. However, sensitivity and specificity were of a good level, suggesting that though the correlation coefficient was low, the activation ratio did provide a valid marker of rigidity.

A benefit of the activation ratio becomes apparent when the properties of the rigidity construct are contrasted with those of the constructs tapped by the instrumental measures of bradykinesia previously considered (4.1.1 – 4.1.2). It was noted that performance measures such as the Jebsen test provide an assessment of speed of movement rather than of slowing, i.e. of absolute performance rather than of the degree of abnormality. As discussed in the introduction (1.3.2.2), the activation ratio represents the degree to which stiffness in the arm increases when a reinforcing technique is used to elicit rigidity. The technique thus controls for baseline differences in rigidity; these may derive from baseline differences in muscle volume.

Though the activation ratio may demonstrate construct validity in this fashion there are difficulties with the stated construct used in the determination of the clinical rating criteria. It is assumed, as with all rating criteria, that the criterion comprises an assessment of the degree of abnormality present. However, the stated construct allowed for both increases in resting rigidity and increases in the degree to which activation increased rigidity. Within this concept, an increase in resting stiffness indicates a more severe degree of rigidity than if the rigidity is only apparent when elicited by reinforcement techniques. The rater (DGCO) stated that he was unaware of findings in the literature demonstrating baseline differences in resting rigidity due

to the gender of the patient (Walsh, 1992) though he would expect greater rigidity in a heavily muscled individual than one more slightly built. Stated in this fashion the clinical rating appears vulnerable to confounding by differences in muscle volume. However, average ratings of rigidity in male and female patients did not differ, suggesting that baseline differences in rigidity resulting from muscle volume were sufficiently accommodated.

The major benefits of the positive feedback device are those common to all instrumental methods of assessment: the assessment is objective and rater-independent. However, as a technique its uses are limited. The equipment used is specialised and cumbersome; the positive feedback device itself is the size of a small table and very heavy. Adjustment to the height of the individual patient is tricky and can be time-consuming. For smaller patients, placing the forearm in the cradle with the elbow concentric to the axle required an awkward stretch. Further equipment is needed to record the output from the device, and once data is recorded, it must be analysed to derive the activation ratio.

The problems with the rigidity instrumentation procedure are not limited to those of the device itself, the principal difficulty is one of compliance. This was not a matter of patients being unwilling to comply with the demands of the assessment, rather that some patients (and some controls too) found it impossible to relax their arm sufficiently to form part of the "torsion pendulum" combination with the rigidity device. This difficulty was most prominent in the activated condition, particularly with the "circle drawing" activation technique. Complying with the demands of the

assessment in this condition was described as being akin to patting one's head at the same time as rubbing the stomach.

The difficulties in finding a suitable reinforcement procedure were noted previously (1.3.2.2). In preliminary trials, it seemed that the different techniques produced varying degrees of activation. These could be ranked from least to greatest degree of activation as follows: very simple verbal tasks (e.g. reciting the days of the week or months of the year in reverse order), more difficult verbal tasks (e.g. serial sevens – counting downwards from 100 in sevens), gripping a rubber ball with the contralateral arm, drawing circles in the air with the contralateral arm.

The degree to which a verbal task counts as difficult rather than simple is greatly dependent upon the individual. If the task is too simple it is an inadequate stressor and will have little or no reinforcing effect. In contrast, the circle drawing task proved almost too effective, preventing some patients (who had already had a measurement of resting stiffness taken) from complying with this part of the assessment. Maintaining a grip on a rubber ball was found to have an effective reinforcement effect without being detrimental to compliance.

Some patients, perhaps among those who found it difficult to comply with the demands of the device, were also noted to be "helping" the device to move their arm. This could be detected by examining the torque trace on the record; an excess of movement in the absence of a significant torque output indicated that the work was

being done by the patient rather than the machine! Though this may be detected, it is another obstacle to obtaining a valid instrumented assessment of rigidity.

In spite of the operational difficulties, the positive feedback device provides a reasonably sensitive and specific indicator of the presence of rigidity. However there are difficulties with patient compliance with the procedure and with the equipment.

4.1.5 Rigidity

The demonstration that the positive feedback device may provide a valid marker of rigidity is not sufficient to demonstrate a role for instrumentation in the assessment of parkinsonism. It is necessary to consider other factors relating to the role of instrumentation in the assessment of rigidity, such as the need or otherwise for greater accuracy than can be obtained using observer-ratings. Though the relative importance of different features in DIP is the subject of the next hypothesis, it also has a bearing on this issue.

Observer rating criteria indicated that severe rigidity was rare in this patient group (3.3.1). The results of the instrumentation too, indicated that only a small proportion of the patient group were more than 1SD above the mean of the control group. Even using this threshold, set to ensure a relatively large "rigid" group, the numbers of patients identified as exhibiting a significant degree of rigidity was small. It is also the case that few patients complain of rigidity, in the way that they do of gross tremors, for example. Further, rigidity does not impair normal activities of daily living in the manner that bradykinesia or tremor may. Finally, in considering

hypothesis two it may become apparent that highly accurate measurement of rigidity is less important than that of other features of parkinsonism.

Though instrumentation may have accuracy and objectivity to recommend it, there is evidence for the reliability and validity of clinical ratings, demonstrating the high level of agreement found between experienced raters. And, when other rigidity instrumentation procedures from the literature are considered construct validity is often lacking (1.3.2.2), and all the procedures reviewed are dependent upon cumbersome and usually expensive equipment. Even if instrumentation has greater accuracy than clinical rating, the disadvantages of instrumentation outweigh any benefit.

In terms of practical utility in clinical practice, there is little reason to believe that possible gains in accuracy or objectivity over clinical ratings outweigh the considerable difficulties inherent in instrumental assessment of rigidity. Even in research settings, the benefits may not be sufficient to justify its use on a regular basis.

4.1.6 Postural tremor amplitude

Two forms of tremor analysis were used, producing two tremor variables: postural tremor amplitude and resting tremor frequency ratio. These variables and the procedures by which they were derived are to be evaluated separately.

The results analysis of the data from the tremor instrumentation found only weak to moderate correlations between postural tremor amplitude and clinical rating criteria (TAKE tremor item, ESRS right and left hand tremor items). Only the relationship between postural tremor amplitude in the left hand and the corresponding ESRS tremor item reaches statistical significance. In contrast, the sensitivity and specificity of this variable as an indicator of symptomatic tremor are both good, especially when mean postural tremor amplitude is evaluated using the TAKE tremor item. Overall, the accuracy of the assessment of postural tremor amplitude is good.

However, though this procedure is an accurate and valid measure of tremor amplitude, it is necessary to address the construct validity of this form of assessment. In particular, it is essential to examine what it is about tremor which is important to measure and what a tremor of abnormal amplitude signifies, whether it is an indicator of parkinsonism, a predictor of a poorer outcome, or something unrelated to any of these factors.

Severe tremor is a problem for patients; it can hinder fine motor control, affecting basic activities of daily living. Further, a very visible tremor may be highly embarrassing and have adverse effects on social competence. Less severe tremor is usually well tolerated if it is associated with otherwise effective and tolerable therapy.

In neurology there is an emphasis, in all domains of symptomatology, on the practical consequences of dysfunction. Thus tremor is considered only as it impinges

on daily life. Assessment of tremor may be via clinical ratings, but assessments of the consequences of the tremor may be made. If an objective assessment of tremor is to be made very simple techniques are most commonly be used. A typical procedure is to measure the volume of liquid spilled when lifting a glass of water which has been filled to the brim.

The evidence from this study does not indicate that increased tremor amplitude is related to any other important indicators. Amongst the patient group as a whole, the greatest individual tremor amplitudes were found in the clozapine group, a group of patients who were notably non-symptomatic on measures of other features of parkinsonism. The great variations in baseline tremor amplitude found in the normal population must also be noted. Tremor amplitude varies greatly between individuals, and from day to day within the individual. It is only when it reaches extreme levels that it is regarded as noteworthy. Further, in the single case study, tremor amplitude increased as levels of other features of parkinsonism, particularly bradykinesia decreased.

The instrumentation has advantages over observer-ratings in its objectivity and accuracy. It is simple to use and the analysis process may be configured to produce a single figure for overall tremor amplitude. In this study the analysis software ran on a standard PC. Alternatively a more simple hard-wired device could perform the same algorithms. This would be both cheaper than a dedicated PC (though the software can run on any PC) and more compact.

4.1.7 Resting tremor frequency ratio

The evaluation of the tremor frequency ratio variable indicated a level of accuracy very different to that found with the postural tremor amplitude. The resting tremor frequency ratio did not provide an indicator of the presence of parkinsonism. In fact, what correlation there was (non-significant and very weak) between the frequency ratio and the ratings of global parkinsonism was negative.

On the basis of observation there was little evidence of slow resting tremor in the patients tested. Clinical experience also suggests that a slow resting tremor is uncommon in the wider population of psychiatric patients receiving antipsychotic medication. However, it is necessary to consider whether there was a failure of the quantification procedures used.

The equipment used for this procedure was the same as that used for the instrumentation of postural tremor amplitude. The results from tremor amplitude analysis were consistent with ratings of tremor severity. It is possible to be confident that there are no inaccuracies in the raw data produced. Further, the Fast Fourier Transform method used to derive the frequency data is beyond doubt. Though this analysis is not presented here, the data from the control group were used to calculate the peak frequency of tremor activity. Results from this analysis were consistent with those from other studies of normal tremor (1.3.1.3.2), indicating a peak of activity with the 8-13 Hz frequency range.

It would have been desirable to have included a Parkinson's disease group for comparison but this did not prove possible. It was originally proposed that a group of untreated PD patients could be recruited. However, in Edinburgh there is no central service for Parkinson's disease and newly-diagnosed patients are usually prescribed L-dopa by GPs. L-dopa, like antipsychotic medication affects a multitude of different systems and disparate physical features. There is as yet no evidence concerning its specific effects on tremor characteristics. Given the difficulties involved in recruiting a group of Parkinson's disease uncontaminated by L-dopa these plans were abandoned.

The resting tremor frequency construct, like the constructs tapped by the other measures, needs to be fully investigated. Predictions were made earlier that an elevated proportion of low frequency tremor would be associated with increased severity of global parkinsonism, based on similar findings in the literature (Arblaster et al., 1993). It was hoped that the frequency ratio could function as a marker of the presence of sub-clinical parkinsonism. However, there is no evidence from this study to support this prediction. In fact, there is little evidence of slowed tremor frequency in the patients tested. The proposed association between a slow resting tremor and the presence of parkinsonism is central to the hypothesised role of instrumented tremor assessment. However, the characterisation of tremor (3.9) found that though symptomatic tremor was present in a large proportion of patients it was a postural tremor of relatively high frequency.

The failure to link changes in tremor frequency with ratings of parkinsonism or with typical antipsychotic medication brings into question the role of instrumentation, or at least of this form of instrumentation, in the assessment of tremor. If it is accepted that the failure to detect a slow resting tremor in the patients tested reflects a genuine absence of this form of tremor in the patient group tested it is necessary to consider why the studies in the literature did find this form of tremor. These studies may differ in the participating patients or in the medication received by those patients.

Some groups of patients can be more susceptible to parkinsonism, particularly older patients (Ayd, 1961), or those with chronic deficit forms of schizophrenia (Prosser et al., 1987). The mean age of the patient group in this study was 36.1 years. This was lower than the mean age in two studies which found evidence of parkinsonian resting tremor in DIP: 44.2 (Caligiuri et al., 1991) and 54.3 (Arblaster et al., 1993). In the light of findings that older patients are more susceptible to developing DIP, and that age-related changes in tremor frequency characteristics may parallel those seen in parkinsonism, it may be that these findings cannot be generalised to younger groups of patients. Further, the group assessed by Caligiuri et al. (1991) had been selected on the basis of their displaying other signs of EPS (TD) and thus cannot be regarded as being representative of all psychiatric patients.

In addition, Arblaster et al. (1993) found that in the patients exhibiting raised a low tremor frequency subsequent clinical assessment revealed previously unnoticed DIP. This parkinsonism predated the study and was apparently detectable by observer ratings though it had previously gone unnoticed. The findings indicate not so much

that the instrumented assessment had great sensitivity but that the patients involved in the study had not been adequately monitored for the development of DIP. Further, they suggest that the overall prevalence of parkinsonism in the group of Arblaster et al. was higher than in the cohort of patients involved in this study.

4.1.8 Tremor

In the literature, analysis of tremor frequency promises most as an indicator of global severity of parkinsonism. However, in this study the resting tremor frequency ratio was unrelated to ratings of global severity. If the instrumentation of resting tremor frequency cannot be validated as an indicator of parkinsonism, is there a role for other tremor quantification techniques, such as those using postural tremor amplitude (which was demonstrated to be closely correlated with ratings of tremor severity)?

An earlier section (4.1.6) considered the importance of an accurate measure of tremor amplitude. It was noted that tremor amplitude varies greatly in the normal population and is affected by a number of variables, most of them unrelated to psychosis or antipsychotic medication. Tremor amplitude is of importance only as it affects activities of daily living or becomes socially embarrassing. To a great extent both of these factors, especially the latter, depend on patient perceptions of the tremor. What is a minor inconvenience to one individual may be greatly distressing to another. There seems little need for, or justification for the use of, a highly accurate measure of tremor amplitude when it is the consequences of the tremor that are meaningful.

Overall, there is apparently little use for accelerometry techniques in clinical practice though a simple method of objectively measuring the consequences of tremor for performance of activities of daily living can have relevance (e.g. the glass of water technique noted earlier). Within the research environment, the use of accelerometry and frequency analysis techniques may have some benefits in distinguishing tremors with different frequency compositions (e.g. 3.9).

4.1.9 Summary of the role of instrumentation in the assessment of DIP

In brief, both measures of bradykinesia proved to be valid assessors of this feature. Accuracy was high, particularly in the case of the Jebsen test evaluated in a single case study. The positive feedback device provided a reasonably valid measure of rigidity though accuracy was not high. Difficulties with the method stemmed from both the awkward nature of the device itself and the problems for participants of co-operating with the procedure. A highly accurate assessment of tremor characteristics was obtained via accelerometry, quantifying tremor amplitude and frequency characteristics. However, there are problems with the constructs used. Though the instrumentation generally provided valid measures of the features concerned, in all cases questions were raised concerning the need for greater accuracy than is provided by observer ratings. This issue is addressed more closely later (4.6).

4.2 Bradykinesia is the predominant feature of DIP

The arguments to be proposed in the examination of this hypothesis were touched upon in the previous section. In addressing the first hypothesis, that instrumentation has a role in the assessment of DIP, the relative importance of measuring accurately the different features of parkinsonism was discussed. This is dependent, in part, upon the predominance or otherwise of those features, it being of greater importance to assess accurately a major feature than a minor one.

The literature suggests that upper-body parkinsonism is the cardinal sign of Parkinson's disease (Quinn, 1995), and clinical experience suggests that this is even more the case in DIP. The most obvious means of identifying a predominant feature is to note the measures, be they observer-ratings or instrumental assessments, on which the greatest proportion of patients fall outside the normal range.

On the basis of the observer-ratings, symptomatic tremor is present in a greater number of patients than is the case for any of the other features of parkinsonism. However, tremor may have many causes other than parkinsonism, including other forms of medication, such as antidepressants, lithium, or clozapine, and this analysis must exclude all those patients. Within the group of patients treated with typical antipsychotic medication only, tremor is less prevalent.

However, there are some receiving typical antipsychotic medication only do exhibit noteworthy levels of tremor. A higher frequency postural tremor has been noted in PD (Findley et al., 1981) but is less commonly described than the resting tremor.

This may account for the tremor seen in the typical antipsychotic group but without a Parkinson's disease comparison group it is difficult to be certain.

Of the other features of parkinsonism, severe levels of rigidity were rarely found and even mild rigidity was not common. Bradykinesia is exhibited by many more patients than is rigidity, and it is present to a severe degree in some. However, many other patients are rated as exhibiting only minimal levels or none at all.

When the results of instrumentation are used, comparisons may be made with a control group and it is possible to consider directly the proportion of patients who fall outside the normal range. When a bradykinetic group was identified using the Jebsen hand function test, a threshold value of 2 SDs above the mean of the control group was used to identify an 'abnormal' group of similar size to that identified using the observer ratings. In the evaluations of the other instrumental assessment procedures a threshold value of 1SD above the mean of the control group was used. This evidence may support the hypothesis that bradykinesia is the predominant feature of DIP, i.e. that severity relative to the control group is greater in this feature than in others. Alternatively, it may indicate merely that the relative sensitivity of the Jebsen test is greater than of the other tests.

Further evidence to support the hypothesis may be obtained by examining the relationship between bradykinesia and other features of parkinsonism. If one feature of a syndrome is more strongly associated with the presence of other features of the syndrome than a second, the former may be regarded as being a more central feature

of the syndrome. In terms of DIP, if bradykinesia is more strongly associated with the presence of other features of parkinsonism than is tremor, this will support the hypothesis that bradykinesia is the primary feature of DIP.

In fact, evidence does exist linking the presence of bradykinesia with other features of parkinsonism. In the literature, bradykinesia has been found to be associated with the non-physical features of parkinsonism, and to a lesser degree with rigidity (Mortimer et al., 1982). In this study the presence of the motor planning impairment was associated with significantly higher levels of bradykinesia.

On the basis of these associations, Mortimer et al. (1982) postulated the existence of two forms of Parkinson's disease: one predominantly tremulous, and characterised by preservation of intellectual abilities, the other predominantly bradykinetic in which cognitive impairments and rigidity are both present to a significant degree. This latter form more closely resembles clinical impressions of bradykinesia. However, few authors have adopted this concept, the varied clinical presentation of Parkinson's disease, and all other forms of parkinsonism, being best accommodated within a syndrome.

The finding that a particular feature of parkinsonism predominates in DIP has implications for the clinical assessment of the disorder, rating scale design etc. A number of scales, still widely used, place their emphasis on features of parkinsonism other than bradykinesia. For example, rigidity was not present to any great degree in

this patient group yet rigidity-related items form the major part of the Simpson-Angus scale (Simpson & Angus, 1970).

The ESRS (Chouinard et al., 1980) is a more modern scale, designed with input from neurology and featuring a novel dual axis rating scheme for tremor. Yet there is perhaps too great an emphasis on tremor in the context of DIP. Though there are a number of bradykinesia-related items (e.g. gait and posture) they make less contribution to total variance than the tremor items.

Though tremor is common in psychiatric patients, and a frequent source of complaints when of large amplitude, its role in DIP is unclear. It is not associated with the other physical features of parkinsonism, nor with the presence of cognitive deficits. Further, the difficulty of accommodating different forms of tremor (low frequency resting tremor vs. high frequency postural tremor) within a rating scale is yet to be satisfactorily resolved, despite the innovative rating guidelines used in the ESRS.

That tremor may result from factors other than parkinsonism is a another source of confounding. Other classes of medications such as lithium or antidepressants may exacerbate tremor, and the prevalence of multiple drug therapy in this cohort was high. The literature indicates that this is far from uncommon, many studies finding the prevalence of multiple drug therapy to be higher than was found here (Burke et al., 1996).

If it can be satisfactorily demonstrated that bradykinesia is the cardinal sign of parkinsonism, then this feature must account for the greatest proportion of variance in ratings of parkinsonism. Further reasons why bradykinesia must be emphasised in ratings of parkinsonism will follow from the next section, the consideration of the third hypothesis.

4.3 Cognitive and subjective features of parkinsonism are present in DIP

This hypothesis will be addressed under two headings: subjective features and objective impairment.

4.3.1 Objective features – motor planning impairment

The methodology for this analysis uses the results from the CANTAB reaction time test. The results of two response conditions are compared. The two conditions are similar in the form of stimulus presented and the response required. However, they differ in the extent of information-processing required. This paradigm has been used to demonstrate impairments in motor planning (Evarts et al., 1981; Flowers, 1978) and in cognition (Rogers et al., 1987).

In this study, the results from the CANTAB reaction time test are used to demonstrate the presence of an impairment in motor planning. In the SRT condition the form of response required can be known prior to the presentation of the stimulus. In the CRT condition the form of response cannot be known until the stimulus is presented. Normal subjects use this prior knowledge of the response form to pre-plan the movement to be made; this acts to decrease response latency. If SRT response latency is subtracted from CRT response latency, the remainder is the information-processing latency – the time taken to plan the response movement in the CRT condition.

It has been demonstrated that Parkinson's disease patients do not show this increase in reaction latency in the CRT condition over the SRT condition: information-processing time is apparently shorter than in normal control groups. Analysis reveals a selective increase of SRT reaction latency, attributed to a failure to make use of available information to pre-plan the response movement.

Though the motor planning impairment is present in a significant proportion of Parkinson's disease patients, its presence has not previously been demonstrated in DIP. In Parkinson's disease the presence of this motor planning impairment (and equally of other cognitive impairments) is associated with greater severity of bradykinesia. It is the association with more severe bradykinesia which can confirm that the impairment is parkinsonian in nature.

The results from the CANTAB reaction time test confirm the presence of the motor planning impairment in a group of patients. This group, termed the 'deficit group' did not show the normal increase in response latency in the CRT condition over the SRT condition, indicating a failure to pre-plan response movements. The severity of bradykinesia in the deficit and non-deficit groups was compared. Observer-ratings of bradykinesia were higher in the deficit patient group than in the non-deficit patients. Instrumented measures of slowing indicated that the deficit group was significantly slower than the non-deficit group. On one measure (Jebsen) the non-deficit group was found to be slowed relative to the control group though less slowed than the deficit group. This association between the presence of the motor planning impairment and greater severity of bradykinesia confirms that the impairment is a

feature of parkinsonism. No other significant differences were found between the two groups.

4.3.2 Subjective features

Subjective sensations were assessed primarily using the visual analogue scales and the SWN. The BDI provided a criterion of subjective depression.

4.3.2.1 Visual analogue scales

A number of notable findings were obtained from the visual analogue scales. Firstly the subjective experience of slowing was only weakly related to either observer-ratings or objective measures of slowing, and similarly, the subjective experience of restlessness showed little association with an observer rating of restlessness. These findings indicate that patients are not aware of the severity of the physical features of parkinsonism. Or, to emphasise that the subjective features may themselves be features of parkinsonism, that the subjective features of parkinsonism are not related necessarily to the objective features.

It should be noted that it is not unknown for experimenters to use patients' subjective ratings of their EPS as indicators of severity. Though subjectively slowed patients may indeed be bradykinetic the low level of correlation between subjective and slowing, consistent with other studies of the subjective experience of treatment (Gerlach & Larsen, 1999), cannot support this methodology.

The subjective ratings were not unrelated to other measures. Links were found with other indicators of subjective state and subjective slowing in particular was very closely associated with depression. As psychomotor slowing is a frequent major component of depression this is not a surprising finding.

The inclusion of the subjective sedation scale was intended to provide an insight into the extent to which patients were able to distinguish slowing from sedation, two phenomena easily misinterpreted by the observer. However, none of the observer rating scales used included an item for sedation, and this hope was further confounded by the failure of the slowing scale to provide a valid indicator of slowing. However, the results of the two scales do indicate a strong correlation between sensations of slowing and sedation. This may suggest either that the two are both induced by the medications received. Or alternatively, that patients experience the two phenomena as subjectively very similar and are as unable to distinguish them as are raters.

Significant correlation was also found between ratings of slowing and restlessness. This finding was unexpected, but may be accounted for by associations between slowing and restlessness, and depression and anxiety, respectively. The relationship between slowing and depression was noted above, subjective restlessness is a common feature of anxiety, and depression and anxiety frequently co-vary.

Perhaps the most important finding from the visual analogue scales was the lack of subjective restlessness found in the clozapine group relative to the typical antipsychotic group. Significantly increased levels of restlessness relative to the control group were found in all other patient groups. That a patient group reported normal levels of restlessness indicates that the increased levels of restlessness in the other patient groups were related to medication rather than features of psychosis. These findings will be addressed further in a later section (4.5.2).

Finally, the use of the scales must be considered. The visual analogue scales proved very simple to use. Patients found completion of these scales easy and intuitive. Responses were supplied rapidly and without the impression that patients felt the scales an imposition (not the case with all the self-ratings administered, 4.3.2.2). A wide range of scores was achieved amongst the patient group, increasing the validity of the analysis; clustering of scores was seen only in the control group.

4.3.2.2 SWN

A very high level of inter-correlation was found between the sub-scales so the SWN total score was used in analysis. The results from this analysis were disappointing. The literature states that the SWN is highly sensitive to the effects of typical antipsychotic agents, and is capable of distinguishing the effects of typical antipsychotics from those of other classes of psychoactive agents including atypical antipsychotics. However, when patient medication groups were compared, SWN total score did not indicate any differences in group mean score between the patient groups. All patient groups tended to exhibit higher mean scores than did the control

group; this tendency was not limited to groups receiving typical antipsychotic medication. The scale did not differentiate antipsychotic treated patients from non-antipsychotic treated patients or typical antipsychotic treated patients from non-antipsychotic treated patients.

The reasons for this apparent failure must be examined. There are three possible causes for this pattern of results:

1) The SWN is an inaccurate scale (i.e. criterion validity is low);
2) The SWN does not measure the phenomenon it purports to measure (i.e. construct validity is low);
3) The patient medication groups do not differ significantly on the axis measured by the SWN (the null hypothesis).

Total SWN score was found to show a very high level of correlation with BDI score. The strength of this relationship indicates that the SWN exhibits a high level of criterion validity as a measure of depression. In consequence, the first of the possible causes (that the scale is simply an inaccurate measure) may be rejected. Two alternative accounts remain: that the SWN is measuring some phenomenon other than that which it purports to measure, and that there are no differences between the patient medication groups.

The association with the BDI scores has further consequences for the evaluation of the SWN. If the SWN is a valid measure of depression it implies that, if the SWN is

indeed sensitive to the effects of typical antipsychotics, that the effects of these agents include depression.

The SWN was developed on an empirical basis, the items used being selected from a larger battery for their sensitivity to the effects of typical antipsychotics. That these effects are closely related to depression is a common theme in literature concerned with subjective experience of antipsychotics (Bandelow et al., 1992; Van Putten & May, 1978; 1.2.4). This evidence suggests that if any differences exist between the patient medication groups then they may be as likely to be detected by the BDI as the SWN. In fact, the BDI results do not show any significant differences in mean score between the patient medication groups. This may be interpreted as evidence supporting postulated cause three, that the SWN did not find significant differences between the patient medication groups because the groups do not differ on the axis measured.

It was not expected that large differences in SWN scores would be found between the groups. The severity of the physical features of DIP was low in almost all patients, and there is some evidence for associations between the physical features of EPS and subjective experiences (Casey, 1994) though it was found in this study that the relationship is not a necessary one. Though the associated features (e.g. bradykinesia and feelings of lethargy and a lack of psychic energy) do not co-vary perfectly in terms of severity, it seems unlikely that very severe negative subjective experiences would be present in a group of patients in whom severe parkinsonism was rare. However, it must be noted that differences in subjective experience of

medication between the patient medication groups were found. The visual analogue scale for restlessness found significantly lower levels of subjective restlessness in the clozapine group than in the typical antipsychotic group.

There are two remaining possible causes of the apparent failure of the SWN to detect the effects of typical antipsychotic medication, namely that the construct validity of the scale is low, and that no differences exist. The first of these accounts is supported by the finding of increased levels of subjective restlessness associated with typical antipsychotic medication. It is possible that the most comprehensive account comprises a combination of these two versions. The SWN appears to be a valid measure of the construct it taps, a construct comprising aspects of antipsychotic associated dysphoria similar in nature (subjective experience and superficial appearance) to depression. However, the construct does not include sensations of restless, a major component of the typical antipsychotic experience. Though effects of medication group were found on subjective ratings of sensations of restlessness similar effects were not found on depression-related sensations. This suggests that the 'failure' of the SWN has two contributory causes:

1) construct validity – the SWN does not tap sensations of restlessness;
2) null hypothesis – the groups do not differ in severity of depression.

In use, the SWN was unpopular with patients, unlike the visual analogue scales. Complaints focussed mainly on the questions being difficult to understand. The language of the questionnaire is complex and abstruse, perhaps a consequence of its translation from the original German. Many patients needed help from an

investigator to understand what the questions were asking of them. Further, the searching nature of the questions and the degree of introspection needed to answer them was often felt to be an imposition. Those patients being assessed on a repeat basis objected to repeating this questionnaire though not the other parts of the assessment. In addition, the sheer number of items present in the scale makes it time-consuming and arduous to complete, this being the case even for the control group.

4.3.2.3 Summary of subjective features of parkinsonism

The results of the measures of subjective experience (visual analogue scales) provide clear evidence of the presence of subjective restlessness associated with typical antipsychotic medication. However, the evidence is more equivocal on the presence of other aspects of subjective experience of antipsychotics. Though some patients' scores on the SWN indicated a lack of subjective well-being, these included patients who were not receiving antipsychotic medication, either typical or atypical. Scores were not higher in patients receiving typical antipsychotics than in other medication groups. Further, there was a very high level of correlation between scores on the SWN and self-ratings of depression, suggesting that the lack of well-being may be disease-related depression rather than medication-induced.

The results also indicate that patients are relatively unaware of the severity of their physical parkinsonism. Ratings of the subjective sensation of slowing are associated with depression rather than objective evidence of slowing, and ratings of subjective restlessness are only weakly related to observer ratings of restlessness.

In terms of the evaluation of the assessment methods the findings were similarly mixed. The visual analogue scales proved to be sensitive markers of the severity of subjective sensations. Combined with their simplicity and ease of use this makes them a powerful tool. In contrast, the SWN proved ineffective at distinguishing the effects of medication from features of psychosis. Though criticism is made of the limited construct employed in the scale, the results emphasise the similarity of the effects of medication to features of psychosis and the difficulties patients as well as physicians have in distinguishing the two types of phenomena.

4.3.3 Summary of the prsence of cognitive and subjective features of parkinsonism in DIP

The evidence examined in this section demonstrates the presence of non-physical features of parkinsonism and other subjective effects of antipsychotic medication in the patient group. It must be noted that not all the adverse effects of typical antipsychotics are parkinsonian: for example, though akathisia has been described as a component of parkinsonism, it may occur in the absence of parkinsonism.

The evidence from the CANTAB reaction time test demonstrates unequivocally the presence of a non-physical feature of parkinsonism. An impairment in motor planning was found in a significant proportion of patients. The presence of this impairment was associated with greater severity of bradykinesia; it is this association with bradykinesia which confirms that the impairment is parkinsonian. Though the presence of the impairment and the association with bradykinesia had been

demonstrated in Parkinson's disease (Evarts et al., 1981; Flowers, 1978), they had not previously been demonstrated in DIP.

Evidence was found of subjective sensations of slowing in all patients; these sensations were associated with depression. Though depression has been described in Parkinson's disease, it may be due to other causes. Some of the patients assessed in this study were treated for depression as a major affective disorder (many of whom were in the typical antipsychotic medication group) and depression has been described as a feature of schizophrenia (Knights & Hirsch, 1981).

It is difficult for the assessor to determine which of these possible mechanisms is responsible for the depression observed in an individual patient. However, when groups of patients are contrasted the comparison between patient medication groups may be informative. Despite the use of this comparison no differences were found between patient medication groups in the extent of either depression or subjective slowing. A similar pattern of results was found in ratings of subjective sedation which correlated closely with both subjective slowing and depression. Though these findings do not conclusively demonstrate the absence of medication-related sedation, slowing and depression, they provide no evidence that the presence of these sensations was related to medication.

Though the patient medication groups did not differ in the levels of slowing or sedation reported, differences were found in levels of subjective restlessness. Mean levels of restlessness were higher than in the control group in all patient groups

except for the clozapine group. The mean level in the clozapine group did not differ from that in the control group and was significantly lower than in the typical antipsychotic group.

The demonstration of subjective effects associated with medication indicates the validity of the visual analogue scales. This provides confidence that the apparent lack of differences between the patient medication groups in terms of ratings of slowing and sedation represents a genuine lack of medication group effects rather than a lack of discriminatory power in the analysis, and that the levels of slowing and sedation (and associated depression) are not related to medication factors.

The results of the SWN did not demonstrate any significant effects of medication group. Scores on the scale correlated closely with scores on the BDI, indicating that the construct tapped by the SWN is very similar to depression. The patient medication groups did not differ in subjective ratings of depression.

In summary, the results provide a demonstration of the presence of a characteristically parkinsonian motor planning impairment and of sensations of restlessness which may be a component of DIP.

4.4 Appraisal of the study

The hypotheses set out in the introduction have been addressed using the results from the clinical and instrumental assessments. However, it is necessary to evaluate how successfully the questions posed have been answered, and examine the validity of the process as a whole. Were the instrumental methods chosen, the patient group recruited, the data collected, and the forms of analysis used appropriate to allow the hypotheses to be confirmed?

4.4.1 Patients recruited

The patients were recruited mostly from acute wards and outpatients clinics. They represented a typical cross-section of patients seen in psychiatric practice. This has benefits in the extent to which results from this study may be generalised to the wider population of psychiatric patients. However, there are disadvantages in terms of research methodology in that care must be taken to avoid problems of confounding by diagnosis within medication groups or vice versa. The patient groups used in analyses are frequently heterogeneous. The medication groups tend to contain patients with varying diagnoses, diagnosis groups contain patients receiving a variety of agents.

Further, the prevalence of multiple drug therapy is high (2.5.5) with a majority of patients receiving more than one psychoactive agent. This form of prescribing may be intended to alleviate side effects of primary treatment agents, to improve response to primary treatment agents, or to treat comorbid states for which a single agent is

not effective (Dufresne, 1995). Supplementary medications in the typical anti-psychotic groups included additional typical antipsychotics, lithium carbonate, tricyclic antidepressants, SSRIs, carbamazepine, and anti-parkinsonian agents (procyclidine). Even in the clozapine group, half of the group were receiving supplementary medication including SSRIs, other antidepressants, and even a small regular dose of a typical antipsychotic (2.5.1 and 2.5.5).

In terms of research methodology, the use of treatment groups in which patients are receiving varied supplementary medications is undesirable. However, it is a common feature of prescribing practice in psychiatry and the figures for multiple drug therapy in this study were in fact lower than in other, similarly diagnosed cohorts. In the study of Burke et al. (1996) up to one third of schizophrenia inpatients were receiving 4 or more medications. Within the context of an observational study, treatment regimes cannot but reflect the high prevalence of multiple drug therapy in clinical practice.

Efforts to recruit patients receiving relatively pure treatment regimes had adverse consequences for the size of the patient groups. Though ideally larger groups with a more homogenous composition would have been recruited, it proved difficult to identify sufficient numbers of suitable patients. Limited group sizes had a detrimental effect on the analytical power of the study, reducing the significance level of group effects. It should be noted that few patients who had been asked to take part refused to do so, and no patients withdrew from the instrumented stage of the assessment.

At the top of this section it was noted that the patients recruited comprised for the most part a cross-section of patients in acute wards and outpatient clinics, and the heterogeneity of the patient groups discussed. Both of these factors may have affected the extent to which some of the results from this study could be compared with those from other studies. Failures to demonstrate the existence of certain features of parkinsonism, or other medication-related phenomena in this patient group have been described elsewhere (slow resting tremor under typical antipsychotic medication, 3.9.3.1; slow resting tremor under lithium therapy, 3.9.3.2). Possible accounts for this failure were suggested earlier, focussing on differences between the patient groups involved in this study and the others, and differences in prescribing practices. It is known that patients who are older or whose symptomatology is more chronic in nature, are more susceptible to parkinsonism. In addition, many studies are conducted using patient groups in which high potency high dose medication regimes are the norm. Both of these factors may have acted to reduce the likelihood of finding evidence of the above phenomena in this patient group.

4.4.2 Criteria

When the accuracy of a new measure is to be evaluated a Gold Standard criterion is used to establish validity. The worth of the evaluation, and hence the extent to which the validity of the measure can be established, is dependent upon the accuracy of the criterion.

The use of observer-ratings as criteria has significant weaknesses. Though intra-rater and inter-rater reliability is high for the scales used, they are dependent upon the skill and experience of the rater, and cannot be truly objective. The ordinal nature of the ratings also affects accuracy. The limited number of possible responses tends to increase the apparent reliability of the assessments when inter- or intra-rater reliability is evaluated, though the sensitivity of the measure to small changes in condition is reduced. Floor effects may also occur where severe dysfunction is rare. A further consequence of an ordinal response paradigm for the statistical analysis is a detrimental effect on the power to detect effects of medication group due to the need to use non-parametric statistical tests.

The weaknesses of some ratings were made apparent when their ability to detect minor differences between medication groups in the severity of slowing was contrasted with that of the Jebsen test (3.10.1.4). In this evaluation the accuracy of the Jebsen test proved greater than that of the TAKE and ESRS single-item ratings of bradykinesia.

Further, there are difficulties with the constructs used to assess some features. This applies particularly to rigidity (4.1.4), though to some extent tremor too (4.1.6 and 4.1.7). If the criterion to be used is felt to be inaccurate, the investigator must rely upon construct validity to demonstrate accuracy beyond that of the criterion, a far more difficult proposition.

4.4.3 Statistical analyses

The accuracy of the instrumental measures was evaluated relative to the observer-ratings criteria. The analysis included calculation of the correlation between the measures and the criteria, and the relative sensitivity and specificity of the measures. These latter two properties indicate respectively the ability of the measure to successfully identify patients who exhibit the disorder and to reject those who do not. Within the limitations of the other factors affecting the study these means of analysis were the most appropriate, allowing the accuracy of the measures to be quantified relative to that of existing criteria and the validity of the measures to be established.

4.4.4 Instrumentation selected

The instrumentation methods used in the study were selected examples of particular genres of objective assessment. It was necessary to determine whether the methods used were valid and then to consider to what extent the findings could be generalised to other instrumental measures of the same features.

4.4.4.1 Jebsen

The Jebsen test was selected because it provides a comprehensive assessment of motor performance. It has been demonstrated to be accurate and reliable, and its results do not exhibit practice effects (Jebsen et al., 1969). However, the test is dependent upon cumbersome equipment and its comprehensive nature makes it time-consuming to complete.

To investigate the role of performance measures in DIP it was necessary to use a test known to possess good properties and validated in a number of different patient groups. It could then be ensured that a failure to demonstrate validity of the measure was an indication that performance measures were an inappropriate means of assessing bradykinesia rather than simply that an inaccurate measure had been used. In order to investigate construct validity it must be ensured that a lack of criterion validity may be discounted as a possible source of invalidity.

The suggestion was made that comparable information to that obtained from the Jebsen might be achieved using simpler and quicker performance tests; such tasks would be more suitable for clinical use. Numerous suitable tasks exist, some having been previously used in Parkinson's disease where the Jebsen has not. However the accuracy and reliability of these tests have not been investigated to the same degree as the Jebsen test. Any such measures would need to be validated individually, possibly against the objective criterion of the Jebsen test. However, the validity of performance measures as assessors of bradykinesia has been demonstrated in principle.

4.4.4.2 CANTAB

The CANTAB battery is even better validated in the literature than the Jebsen test. It provides timing of movement speed to a greater accuracy than can be obtained using manual timing (e.g. in the Jebsen test). As with the Jebsen test as an example of a performance measure, it was vital to ensure that failure to demonstrate validity in the

use of a reaction time test as a measure of bradykinesia was not the result of an inaccurate measure but attributable to a lack of construct validity.

The construct validity of the CANTAB reaction time test was demonstrated though it may be of a lower level than that of the Jebsen test. As a means of evaluating the role of reaction time measures in assessing bradykinesia the CANTAB test proved very appropriate. It allowed the comparison of this type of measure with performance measures such as the Jebsen test. The very simple construct underlying reaction time tests means that the findings obtained using the CANTAB test may be generalised to any accurate reaction time test.

4.4.4.3 Rigidity

The positive feedback device had been used previously and validated in other patient groups. Unlike other methods of rigidity instrumentation reviewed in the introduction it is theoretically sound. The validity of this device in this patient group was examined earlier. The positive feedback device demonstrates a reasonable level of validity; though it is perhaps less accurate than the bradykinesia instrumentation used, this may be due in part to floor effects (severe rigidity was very rare). There were also problems with subject compliance (not specific to patient groups) with the demands of the assessment.

As with the instrumentation methods selected to assess other features of DIP it was essential to select a well validated measure of rigidity. The theoretical construct validity of the positive feedback device makes it a more appropriate choice than

other existing methods of rigidity instrumentation. However, the failings of the other methods make it difficult to generalise from the results obtained in this study.

4.4.4.4 Tremor

The principle methods of tremor transcription via accelerometry and analysis of tremor frequency composition using FFT are well established in the literature. Accelerometry is an accurate modern method of transcribing tremor activity. Unlike other older methods it is non-intrusive, relying upon a small lightweight electronic sensor fastened to the finger by a velcro strap rather than fastening the patient to a mechanical device. The interface and software used in this study were new but are standard procedures. There is little consistency in the literature over the optimal method of quantifying differences in tremor frequency composition but the algorithm used in this study had been previously validated in studies which found positive results (Caligiuri et al., 1989b; Caligiuri et al., 1991; Caligiuri & Lohr, 1993); there is no indication from other authors that it is invalid.

The two quantification procedures used must be considered separately, as they were analysed earlier. The criterion validity of the postural tremor amplitude method was demonstrated but it is unclear whether the construct has much relevance to other measures of treatment success. Specifically, the procedure may not be worthwhile to achieve the extra accuracy gained over simpler methods of assessment.

No evidence was found of an association between resting tremor frequency ratio and other measures of parkinsonism. As there were no indications of failings in the

instrumentation and the tremor frequency characterisation analysis had found evidence of medication group differences in tremor frequency composition this was attributed to characteristics of the patient group.

The methods selected for the study were highly appropriate for their purpose. Accelerometry, the FFT algorithm, and the quantification procedures used (postural tremor amplitude and resting tremor frequency ratio) are all well supported in the literature. The findings, as they pertain to the constructs examined, may be generalised to other valid instrumentation procedures. However, it is worth noting the comments made above concerning the possibility that particular characteristics of the patient group in this study may have contributed to the apparent failure to replicate previously published findings.

4.4.5 Hypotheses

This section considers the hypotheses used, how appropriate they were and how well they were addressed.

4.4.5.1 Instrumentation has a role in the assessment of DIP

This hypothesis was at the heart of the project. Instrumentation to assess the features of parkinsonism has been investigated in many other studies but, the majority of these studies have focussed on a single feature of parkinsonism (and a single method of instrumentation). All of these studies have addressed simply the validity of the particular method and not considered whether there is a meaningful role for the procedure question. The role of instrumentation as a whole was a question which had

not previously been addressed explicitly, and one which deserved this degree of attention.

4.4.5.2 Bradykinesia is the predominant feature of DIP

The predominance of bradykinesia in Parkinson's disease was noted often in this work. Clinical experience suggests that it predominates to an even greater degree in DIP. However, this question had not previously been addressed in a systematic fashion. This hypothesis was phrased in a manner that enabled the issue to be addressed using quantitative techniques.

4.4.5.3 Cognitive and subjective features of parkinsonism are present in DIP

The difficulties of distinguishing the superficially similar phenomena of DIP and deficit symptoms of psychosis were discussed in the introduction. It was also suggested that the boundary between these two domains of symptomatology was central to treatment success and a major issue in modern psychiatry. This hypothesis provided a means of empirically demonstrating that some deficits, externally similar to features of psychosis were associated with physical features of parkinsonism.

4.5 Other findings

4.5.1 Tremor characterisation

Results from the tremor instrumentation frequency analysis were used to characterise the form of tremor found in patient medication groups relative to normal physiologic tremor in the control group.

4.5.1.1 Results from this study

Medication group	Tremor properties
Typical antipsychotics	Non-significant increase in high frequency tremor activity in posture.
Lithium	Increase in postural tremor amplitude particularly in the high frequency band.
Clozapine	Increase in postural tremor amplitude in both high and low frequency bands. Non-significant increase in resting tremor amplitude particularly in the low frequency band.
Antidepressants	Increase in resting tremor amplitude particularly in the high frequency band

4.5.1.2 Results from the literature

Tremor	Qualitative properties	Quantitative properties
Parkinson's disease	Slow resting tremor. Faster postural tremor (Lance et al., 1963).	

DIP	Slow resting tremor (Arblaster et al., 1993). Faster postural tremor; slow resting tremor may be rare in DIP (Owens, 1999).	Resting tremor peak below 7 Hz (Arblaster et al., 1993), due to increased amplitude at lower frequencies (Caligiuri et al., 1991).
Lithium	Fine postural tremor, possibly related to essential tremor, distinct from resting tremors (i.e. parkinsonism) and other postural tremors (Gelenberg & Jefferson, 1995).	Acute lithium tremor is of increased amplitude; frequency composition is unchanged (Pullinger & Tyrer, 1983). Chronic lithium tremor exhibits lower peak frequency and is possibly extrapyramidal. (Tyrer et al., 1981).
Clozapine	No work concerned with the properties of tremor under clozapine could be found.	

4.5.1.3 Comparison

The comparisons will be considered by medication group commencing with the typical antipsychotic group, the primary focus of this study.

In the typical antipsychotic groups, the results of this study are inconsistent with the published literature as they do not indicate the stereotypical slow resting tremor of Parkinson's disease found by other authors (Caligiuri et al., 1991; Arblaster et al., 1993). The reasons for this result are considered earlier (4.5.1.3). In essence it is

believed to arise either from differences in the characteristics of the patients tested or the medication regimes used in the different studies.

The pattern of tremor seen in the lithium-treated patients in this study is consistent with descriptions of lithium tremor as a fine postural tremor (Gelenberg and Jefferson, 1995). Some indications were noted of increases in low frequency amplitude though not to the extent found by Tyrer et al. (1981). As with the absence of slow resting tremor in the typical antipsychotic group the lack of slow resting tremor may result from differences in the patient groups tested, some patients being more susceptible to developing EPS than others.

No studies of instrumented tremor under clozapine therapy were found. However, the increased amplitude of postural tremor is consistent with informal observations.

4.5.2 Clozapine and other atypical antipsychotics

The development of clozapine and its benefits over typical antipsychotics were examined in the introduction. It was noted that clozapine is found to be highly effective in the management of positive psychotic symptomatology yet is significantly more tolerable than typical antipsychotic agents and has increased efficacy in treatment-resistant cases. The relevant findings from this study involve both objective and subjective indicators of tolerability and re-emphasise the differences between clozapine and typical antipsychotics.

It should be noted that the clozapine group is not in way a specially selected group. Rather, this group of patients could be said to be 'negatively selected'. Clozapine is prescribed at the Royal Edinburgh Hospital to two categories of patients. The first of these categories comprises patients who have failed to respond to typical antipsychotics, those with treatment-resistant schizophrenia. The second comprises patients unable to tolerate typical antipsychotics, who exhibit severe levels of EPS even with minimal doses. It would be reasonable to expect greater levels of bradykinesia in a group consisting of patients who are likely to be receiving relatively high equivalent doses, and patients likely to exhibit more severe adverse effects relative to their antipsychotic dose. However, despite the selection bias, the clozapine group demonstrated clear advantages over typical antipsychotics.

Firstly, the Jebsen test results (3.10.1) indicated that unlike all other patient groups, and particularly the typical antipsychotic group, the clozapine group was not significantly slowed relative to the control group. Psychomotor slowing may be a feature of both DIP and psychosis (negative symptoms of schizophrenia and depression), yet this patient group were unimpaired relative to the controls.

The predominant role of bradykinesia in DIP was fully discussed earlier, as was its relationship to deficit symptoms of schizophrenia and global functioning including psychosocial competence. The value of an antipsychotic agent which is effective even in treatment-resistant schizophrenia yet has a low liability to cause DIP cannot be over-emphasised.

The second notable difference involved subjective experiences of medication. Levels of sensations of restlessness, as indicated by the visual analogue scale were significantly lower in the clozapine group than in the typical antipsychotic group, and were not significantly higher than those in the control group (3.10.2). Subjective restlessness is frequently associated with akathisia, and has been found to be a major factor in non-compliance with medication (1.2.4).

Also discussed in the introduction was the development of other atypical antipsychotics. These proposed successors to clozapine are hoped to provide all the benefits of clozapine in terms of efficacy and neurological tolerability yet without the dangerous adverse effects of clozapine. A small group of patients receiving these 'New Generation' antipsychotics were included in this study. The results from this group must be viewed with extreme caution in light of the inadequate patient numbers (see below), however the New Generation atypical antipsychotics appeared to have a degree of tolerability midway between those of typical antipsychotics and clozapine. This finding is consistent with the literature (Miller et al., 1998).

The small size of the New Generation group is in part due to the manner in which the medication groups were recruited. It was originally intended to use a single atypical antipsychotics group consisting of patients receiving either clozapine or one of the New Generation atypical antipsychotics. However, preliminary results indicated that though results of the clozapine group tended to form a cluster distinct from results of the typical antipsychotic group, results of the New Generation group were midway between the clozapine group and the typical antipsychotic group. To address this

problem attempts were made to recruit more patients to both the atypical antipsychotic groups. Though the clozapine group reached sufficient numbers it did not prove possible to recruit more patients to the New Generation group within the time available. The size of this group has adverse consequences for the power of all statistical analyses involving this group. It is also necessary to note the presence of different atypical antipsychotic agents within this group: it cannot be assumed that there are not significant differences in the tolerability profiles of these different drugs.

Despite these caveats, the findings appear to reinforce suggestions (Miller et al., 1998) that clozapine is still unique in its level of tolerability.

4.6 Implications for research and/or clinical practice

The primary thrust of this study was to investigate the role of instrumentation in the assessment of DIP, hypothesising that such a role existed. The evidence collected indicates that though instrumented assessment has a part to play this role may be limited.

The assessment methods used were generally inconvenient and impractical for regular use, relying upon awkward, cumbersome and often expensive equipment. Though the literature indicated that information concerning tremor frequency composition not available from observer ratings might mark global parkinsonism this was not the case in this patient cohort.

Further, there is little evidence to suggest that accuracy is substantially greater than with observer ratings. It was only with the use of performance measures to instrument bradykinesia that greater accuracy could be demonstrated (3.10.1). Nor is there much evidence to suggest a need for greater accuracy than can be obtained using well-designed rating scales (though a case may be made for bradykinesia, see next paragraph). The routine use of instrumentation would be justified only on finding a method of instrumentation which is not only as accurate as observer ratings but simple to use, cheap and portable.

The central role of deficit symptomatology in predicting treatment outcome was noted in the introduction (1.1.3), as was the fact that these symptoms may occur as features of both psychosis and parkinsonism. The great superficial similarities in

some aspects of these disorders make it difficult to determine whether deficits observed are features of psychosis or the adverse effects of medication (though some investigators have suggested that this is possible on a longitudinal basis (Amador et al., 1999; Carpenter et al., 1988)).

The association between cognitive deficit features of parkinsonism and bradykinesia was also noted earlier. This link has been demonstrated in Parkinson's disease (Mortimer et al., 1982) and now in DIP (3.6.2 – 3.6.3). It is proposed that assessment of the severity of bradykinesia may allow an indirect monitoring of the severity of parkinsonian deficits, free from confounding by features of psychosis.

However, as demonstrated earlier the accuracy of the performance measures of parkinsonism was only moderate in the cross-sectional analysis. The procedures are inadequate to accurately identify the presence of parkinsonism in a one-off assessment. In the single case serial assessment (3.8) a number of observations were made of clinical state using both observer-ratings and instrumental measures. In this longitudinal follow-up situation the degree of correlation between the Jebsen test and observer-ratings of bradykinesia indicated a very high level of accuracy in the results of the test.

The routine use of performance measures of bradykinesia may be recommended as a means of indirectly monitoring the development of non-physical features of parkinsonism. These deficits may be identified as resulting from adverse medication effects rather than features of psychosis and may thus be treated appropriately.

References

Allderidge, P. (1979). Hospitals, Madhouses and Asylums: Cycles in the care of the insane. British Journal of Psychiatry, 134: 321-34.

Amador, X.F., Kirkpatrick, B., Buchanan, R.W., Carpenter, W.T., Marcinko, L. & Yale, S.A. (1999). Stability of the diagnosis of deficit syndrome in schizophrenia. American Journal of Psychiatry, 56, 637-639.

Andreasen, N.C. & Olsen, S. (1982). Negative versus positive schizophrenia: definition and validation. Archives of General Psychiatry, 39, 789-794.

Andreasen, N.C. (1983). The scale for the assessment of negative symptoms (SANS). Iowa City: University of Iowa.

Andreasen, N.C. (1984). The scale for the assessment of positive symptoms (PANS). Iowa City: University of Iowa.

Andreasen, N.C., Flaum, M., Swayze, V.W., Tyrrell, G. & Arndt, S. (1990). Positive and negative symptoms in schizophrenia. Archives of General Psychiatry, 47, 615-621.

Arblaster, L.A., Lakie, M., Mutch, W.J. & Semple, E. (1993). A study of the early signs of drug-induced Parkinsonism. Journal of Neurology, Neurosurgery &

Psychiatry, 56, 301-303.

Asberg, M., Montgomery, S., Perris, C., Schalling, D., & Sedvall, G. (1978). A Comprehensive Psychopathological Rating Scale. Acta Psychiatrica Scandinavica, Supplement 272.

Austin, M.P., Mitchell, P., Wilhelm, K., et al. (1999). Cognitive function in depression : a distinct pattern of frontal impairment in melancholia? Psychological Medicine, 29: 73-85.

Awad, G.A. & Hogan, T.P. (1994). Subjective response to neuroleptics and the quality of life: implications for treatment outcome. Acta Psychiatrica Scandinavica, 89 (Suppl. 380), 27-32.

Awad, G.A. (1993). Subjective response to neuroleptics in schizophrenia. Schizophrenia Bulletin, 19, 609-618.

Ayd, F.J. (1961). A survey of drug-induced extrapyramidal reactions. Journal of the American Medical Association, 175, 1054-60.

Baddeley, A. (1990). Human Memory. Theory and Practice. Hove: Erlbaum.

Bain, P.G., Findley, L.J., Atchison, P. et al. (1993). Assessing tremor severity. Journal of Neurology, Neurosurgery and Psychiatry, 56, 868-873.

Bandelow, B., Müller, P., Frick, U. et al. (1992). Depressive syndromes in schizophrenic patients under neuroleptic therapy. European Archives of Psychiatry and Clinical Neuroscience, 241: 291-295.

Beck, A.T., Ward, C.H., Mendelson, M., Mock, J. & Erbaugh, J. (1961). An inventory for measuring depression. Archives of General Psychiatry, 4, 561-571.

Belmaker, R.H. & Wald, D. (1977). Haloperidol in normals. British Journal of Psychiatry, 131, 222-223.

Bloxham, C.A., Mindel, T.A. & Frith, C.D. (1984). Initiation and execution of predictable and unpredictable movements in Parkinson's disease. Brain, 107, 371-384.

Brébion, G., Amador, X., Smith, M.J. & Gorman, J.M. (1998). Memory impairment and schizophrenia: the role of processing speed. Schizophrenia Research, 30, 31-39.

Brown, R.A., Lawson, D.A., Leslie, G.C. et al. (1988a). Applying the Wartenburg pendulum to healthy elderly subjects. Journal of Neurology, Neurosurgery and Psychiatry, 51, 1171-1177.

Brown, R.A., Lawson, D.A., Leslie, G.C. et al. (1988b). The Wartenburg pendulum test – can it differentiate spasticity and rigidity? Journal of Neurology, Neurosurgery

and Psychiatry, 51, 1178-1186.

Brown, R.G. & Marsden, C.D. (1984). How common is dementia in Parkinson's disease? Lancet, ii, 1262-1265.

Brown, R.G. & Marsden, C.D. (1990). Cognitive function in Parkinson's disease: from description to theory. Trends In Neurosciences, 13, 21-29.

Burke, M.J., Carmichael, C., Harvey, A. & Preskorn, S.H. (1996). The relevance of polypharmacy as a potential risk for drug interactions in the outpatient treatment of major depression and the inpatient treatment of schizophrenia. Psychopharmacology Bulletin, 32, 419.

Caligiuri, M.P. & Galasko, D.R. (1992). Quantifying drug-induced changes in parkinsonian rigidity using an instrumental measure of activated stiffness. Clinical Neuropharmacology, 15, 1-12.

Caligiuri, M.P. & Lohr, J.B. (1993). Worsening of postural tremor in patients with levodopa-induced dyskinesia: a quantitative analysis. Clinical Neuropharmacology, 16, 244-250.

Caligiuri, M.P., Bracha, H.S. & Lohr, J.B. (1989). Asymmetry of neuroleptic-induced rigidity: development of quantitative methods and clinical correlates. Psychiatry Research, 30, 275-284.

Caligiuri, M.P., Lohr, J.B., Bracha, H.S. & Jeste, D.V. (1991). Clinical and instrumental assessment of neuroleptic-induced parkinsonism in patients with tardive dyskinesia. Biological Psychiatry, 29, 139-148.

Caligiuri, M.P., Lohr, J.B., Galasko, D.R. (1989b). Persistence of neuroleptic-induced parkinsonism. Abstracts of the American Psychiatric Association Annual Meeting.

Carmines, E.G. & Zeller, R.A. (1979). Reliability and validity assessment.

Carpenter, W.T., Heinrichs, D.W. & Wagman, A.M.I. (1988). Deficit and nondeficit forms of schizophrenia: the concept. American Journal of Psychiatry, 145, 578-583.

Casey, D.E. (1994). Motor and mental aspects of acute extrapyramidal syndromes. Acta Psychiatrica Scandinavica, 89 (Suppl. 380), 14-20.

Chouinard, G., Ross-Chouinard, A., Annable, L. & Jones, B.D. (1980). Extrapyramidal rating scale. Canadian Journal of Neurological Science, 7, 233.

Cole, J.O. & the NIMH Psychopharmacological Service Center Collaborative Study Group (1964). Phenothiazine treatment in acute schizophrenia. Archives of General Psychiatry, 10, 246-261.

Cooper, J.A., Sagar, H.J., Tidswell, P. & Jordan, N. (1994). Slowed central processing

in simple and go/no-go reaction time tasks in Parkinson's disease. Brain, 117, 517-529.

Creese, I., Burt, D.R. & Snyder, S.H. (1976). Dopamine receptor binding clinical and pharmacological potencies of anti-schizophrenic drugs. Science, 192, 481-448.

Crow, T.J. (1980). Molecular pathology of schizophrenia: More than one disease process? British Medical Journal, 280, 66-68.

Crow, T.J. (1995). A continuum of psychosis, one human gene, and not much else – the case for homogeneity. Schizophrenia Research, 17: 135-145.

Dalrymple-Alford, J.C., Kalders, A.S., Jones, R.D. & Watson, R.W. (1994). A central executive deficit in patients with Parkinson's disease. Journal of Neurology, Neurosurgery & Psychiatry, 57, 360-367.

Delwaide, P.J. & Gonce, M. (1988). Pathophysiology of Parkinson's signs. In: Parkinson's Disease and Movement Disorders. Eds: J. Jankovic & E. Tolosa. Urban & Schwarzenberg, Baltimore-Munich.

Dietz, V., Quintern, J. & Berger, W. (1981). Electrophysiological studies of gait in spasticity and rigidity. Evidence that altered mechanical properties of muscle contribute to hypertonia. Brain, 104, 431-449.

Dobbs, R.J., Bowes, S.G., Charlett, A. et al. (1993). Hypothesis: the bradyphrenia of Parkinsonism is a nosological entity. Acta Neurologica Scandinavica, 87, 255-261.

Dobbs, R.J., Lubel, D.D., Charlett, A. et al. (1992). Hypothesis: age-associated changes in gait represent, in part, a tendency towards parkinsonism. Age and Ageing, 21, 221-225.

Dufresne, R.L. (1995). Psychiatric pharmacy research workshop: the promise and pitfalls of polypharmacotherapy in the nineties. Psychopharmacology Bulletin 31: 789-796.

Ebmeier, K.P., Potter, D.D., Cochrane, R.H.B. et al. (1992). Event related potentials, reaction time, and cognitive performance in idiopathic Parkinson's disease. Biological Psychology, 33, 73-89.

Evarts, E.V., Teravainen, H. & Calne, D.B. (1981). Reaction time in Parkinson's disease. Brain, 104, 167-186.

Farde, L., Nordstrom, A-L., Wiesel, F-A., et al. (1992). Positron emission tomographic analysis of central D_1 and D_2 dopamine receptor occupancy in patients treated with classical neuroleptics and clozapine: Relation to extrapyramidal side effects. Archives of General Psychiatry, 49: 538-544.

Fenton, W.S. & McGlashan, T.H. (1994). Antecedents, symptom progression and

long-term outcome of the deficit syndrome in schizophrenia. American Journal of Psychiatry, 151, 351-356.

Findley, L.J., Gresty, M.A. & Halmagyi, G.M. (1981). Tremor, the cogwheel phenomenon and clonus in Parkinson's disease. Journal of Neurology, Neurosurgery and Psychiatry, 44, 534-546.

Fleming, K., Goldberg, T.E., Binks, S. et al. (1997). Visuospatial working memory in patients with schizophrenia. Biological Psychiatry 41:43-39.

Fleming, K., Goldberg, T.E., Gold, J.M., Weinberger, D.R. (1995). Verbal working memory dysfunction in schizophrenia: use of a Brown-Peterson paradigm. Psychiatry Research, 56, 155-161.

Flowers, K. (1978). Lack of prediction in the motor behaviour of parkinsonism. Brain, 101, 33-52.

Frith, C.D. (1992). The Cognitive Neuropsychology of Schizophrenia. Hove: L. Erlbaum.

Frith, C.D., Leary, J., Cahill, C., & Johnstone, E.C. (1991). Disabilities and circumstances of schizophrenic patients – A follow-up study IV. Performance on psychological tests. Demographic and clinical correlates of the results of these tests. British Journal of Psychiatry, 159 (Suppl. 13): 26-29.

Gauntlett-Gilbert, J., Roberts, R.C., Brown, V.J. (1999). Mechanisms underlying attentional set-shifting in Parkinson's disease. Neuropsychologia, 37: 605-616.

Gelenberg, A.J. & Jefferson, J.W. (1995). Lithium tremor. Journal of Clinical Psychiatry, 56, 283-287.

Gerlach, J., & Larsen, E.B. (1999). Subjective and mental side-effects of antipsychotic treatment. Acta Psychiatrica Scandinavica, 99 (Suppl. 395): 113-117.

Goldberg, T.E., Torrey, E.F., Gold, J.M., Biglow, L.B., Weinberger, D.R. (1993). Learning and memory in monozygotic twins discordant for schizophrenia. Psychological Medicine, 1: 71-85.

Goldberg, T.E., Weinberger, D.R., Berman, K.F., Pliskin, N.H. & Dodd, M.H. (1987). Further evidence for dementia of the pre-frontal type in schizophrenia – a controlled study of teaching the WCST. Archives of General Psychiatry, 44, 1008-1014.

Greenhalgh, T. (1997). How to read a paper: Papers that report diagnostic or screening tests. British Medical Journal, 315: 540-543.

Gruzelier, J., Seymour, K., Wilson, L., Jolley, A., Hirsch, S. (1988). Impairments on neuropsychological tests of temporohippocampal and frontohippocampal functions and word fluency in remitting schizophrenia and affective disorders. Archives of

General Psychiatry 45: 623-629.

Guy, W. (1976). ECDEU Assessment Manual for Psychopharmacology. Washington, DC: Department of Health, Education and Welfare.

Haase, H-J. (1985). Clinical observations on the actions of neuroleptics. In: The action of neuroleptic drugs. Eds: H-J. Haase & P.A.J. Janssen. Amsterdam: Elsevier.

Hallett, M. (1986). Differential diagnosis of tremor. In: Handbook of Clinical Neurology, vol 49. Eds: P.J. Vinken, G.W. Bruyn, H.L. Klawans. Elsevier, New York.

Hart, R.P., Wade, J.B., Calabrese, V.P., & Colenda, C.C. (1998). Vigilance performance in Parkinson's disease and depression. Journal of Clinical and Experimental Neuropsychology, 20: 111-117.

Heinrichs, D.W., Hanlon, T.E., & Carpenter Jr., W.T. (1984). The Quality of Life Scale: an instrument for rating the schizophrenic deficit syndrome. Schizophrenia Bulletin, 10, 388-398.

Hoehn, M.M., Crowley, T.J., & Rutledge, C.O. (1976). Dopamine correlates of neurological and psychological status in untreated Parkinsonism. Journal of Neurology, Neurosurgery, and Psychiatry, 39: 941-951.

Hogan, T.P. & Awad, A.G. (1992). Subjective response to neuroleptics and outcome in schizophrenia: a re-examination comparing two measures. Psychological Medicine, 22, 347-352.

Hollister, L.E. (1992). Neuroleptic dysphoria: so what's new? Biological Psychiatry 31: 531-532.

Hömberg, V., Hefter, H., Reiners, K. & Freund, H-J. (1987). Differential effects of changes in mechanical limb properties on physiological and pathological tremor. Journal of Neurology, Neurosurgery and Psychiatry, 50, 568-579.

Jaber, M., Robinson, S.W., Missale, C., & Caron, M.G. (1996). Dopamine receptors and brain function. Neuropharmacology, 35: 1503-1519.

Jebsen, R.H., Taylor, N., Treischmann, R.B., Trotter, M.J. & Howard, L.A. (1969). An objective and standardised test of hand function. Archives of Physical, Medical & Rehabilitation, 50, 311-319.

Johnstone, E.C., Crow, T.J., Frith, C.D., Carney, M.W.P. & Price, J.S. (1978). Mechanism of the antipsychotic effects in the treatment of acute schizophrenia. Lancet, i, 848-851.

Kane, J.M. (1995). Current problems with the pharmacotherapy of schizophrenia. Clinical Neuropharmacology, 18 (Suppl. 1): S154-S161.

Kay, S.R. (1991). Positive and negative syndromes in schizophrenia. New York: Brunner/Mazel.

Kay, S.R., Fiszbein, A., Opler, L.A. (1987). The positive and negative symptom scale for schizophrenia (PANSS). Schizophrenia Bulletin 13: 261-276.

Kirollos, C., Charlett, A., O'Neill, C.J.A. et al. (1996). Objective measurement of activation of rigidity: diagnostic, pathogenic and therapeutic implications in parkinsonism. British Journal of Clinical Pharmacology, 41, 557-564.

Kirollos, C., O'Neill, C.J.A., Dobbs, R.J. et al. (1993). Quantification of the cardinal signs of parkinsonism and of associated disability in spouses of sufferers. Age and Ageing, 22, 20-26.

Knights, A. & Hirsch, S.R. (1981). 'Revealed' depression and drug treatment for schizophrenia. Archives of General Psychiatry, 38: 806-811.

Kraepelin, E. (1986). Dementia Praecox. In: The clinical routes of the schizophrenia concept. Eds. and trans: J. Cutting & M. Shepherd. Cambridge: Cambridge University Press.

Lakie, M., Walsh, E.G. & Wright, G.W. (1986). Passive mechanical properties of the wrist and physiological tremor. Journal of Neurology, Neurosurgery and Psychiatry,

49, 669-676.

Lance, J.W., Schwale, R.S. & Peterson, E.A. (1963). Action tremor and the cog-wheel phenomenon in Parkinson's disease. Brain, 86, 95-115.

Lees, A.J. & Smith, E. (1983). Cognitive deficits in the early stages of Parkinson's disease. Brain, 106: 257-70.

Liddle, P.F. & Morris, D.L. (1991). Schizophrenic syndromes and frontal lobe performance. British Journal of Psychiatry, 158, 340-345.

Liddle, P.F. (1987). The symptoms of chronic schizophrenia: a re-examination of the positive-negative dichotomy. British Journal of Psychiatry, 151, 145-151.

Long, C., Thomas, D., Crochetiere, W.J. (1964). Objective measurement of muscle tone in the hand. Clinical Pharmacology & Therapeutics 5: 909-917

Lund-Johansen, M., Hugdahl, K., & Wester, K. (1996). Cognitive function in patients with Parkinson's disease undergoing stereotaxic thalamotomy. Journal of Neurology, Neurosurgery, and Psychiatry, 60: 564-571.

Marsden, C.D. (1982). The mysterious motor function of the basal ganglia: The Robert Wartenberg Lecture. Neurology, 32: 514-539.

Marsden, C.D., Meadows, J.C., Lange, G.W. & Watson, R.S. (1969). Variations in human physiological finger tremor with particular reference to changes with age. Electroencephalography and Clinical Neurophysiology, 27, 169-178.

Marsden, C.D., Meadows, J.C., Lange, G.W. & Watson, R.S. (1969). Variations in human physiological finger tremor with particular reference to changes with age. Electroencephalography and Clinical Neurophysiology, 27, 169-178.

Marshall, J. & Schnieden, H. (1966). Effect of adrenaline, noradrenaline, atropine and nicotine on some types of human tremor. Journal of Neurology, Neurosurgery and Psychiatry, 29, 214-218.

Martin, J.P. (1967). The basal ganglia and posture. Pitman, London.

Mayeux, R., Stern, Y., Sano, M., Cote, L. & Williams, J.B.W. (1987). Clinical and biochemical correlates of bradyphrenia in Parkinson's disease. Neurology, 37, 1130-1134.

McGrath, J., Scheldt, S., Welham, J., Clair, A. (1997). Performance on tests sensitive to impaired executive ability in schizophrenia, mania and well controls: acute and subacute phases. Schizophrenia Research, 26: 127-137.

Meier, M.J. & Martin, W.E. (1970). Measurement of behavioural changes in patients on L-DOPA. Lancet, i, 352-353.

Meltzer, H.Y. (1992). Dimensions of outcome with clozapine. British Journal of Psychiatry, 160 (suppl. 7): 46-53.

Meltzer, H.Y.,Burnett, S., Bastani, B., et al. (1990). Effect of six months of clozapine treatment on the quality of life of chronic schizophrenic patients. Hospital and Community Psychiatry, 41, 892-897.

Miller, C.H., Mohr, F., Umbricht, D., et al. (1998). The prevalence of acute extrapyramidal signs and symptoms in patients treated with clozapine, risperidone, and conventional antipsychotics. Journal of Clinical Psychiatry, 59: 69-75.

Mindham, R.H., Marsden, C.D., & Parkes, J.D. (1976). Psychiatric symptoms during L-dopa therapy for Parkinson's disease and their relationship to physical disability. Psychological Medicine, 6: 23-33.

Mindham, R.H.S. (1970). Psychiatric symptoms in parkinsonism. Journal of Neurology, Neurosurgery and Psychiatry, 33, 188-191.

Montgomery, S.A. & Asberg, M. (1979). A new depression scale designed to be sensitive to change. British Journal of Psychiatry, 134, 382-389.

Naber, D. (1995). A self-rating to measure subjective effects of neuroleptic drugs,

relationships to objective psychopathology, quality of life, compliance and other clinical variables. International Clinical Psychopharmacology, 10 (Suppl. 3), 133-138.

Naber, D., Walther, A., Kircher, T., Hayek, D. & Holzbach, R. (1994). Subjective effects of neuroleptics predict compliance. In: Prediction of neuroleptic treatment outcome in schizophrenia. Concepts and methods. Eds: W. Gaebel & A.G. Awad. Wien: Springer-Verlag.

Nathaniel-James, D.A., Brown, R., Ron, M.A. (1996). Memory impairment in schizophrenia: its relationship to executive function. Schizophrenia Research, 21, 85-96.

Naville, F. (1922). Études sur les complications et les séquelles mentales de l'encéphalite épidémique. La bradyphrénie. Encéphale, 17, 369-375, 423-436.

Nelson, H.E. (1976). A modified card sorting test sensitive to frontal lobe defects. Cortex, 12, 313-324.

Onuaguluchi, G. (1964). Parkinsonism. London: Butterworth.

Owen, A.M., Iddon, J.L., Hodges, J.R., Summers, B.A. & Robbins, T.W. (1997). Spatial and non-spatial working memory at different stages of Parkinson's disease. Neuropsychologia, 35, 519-532.

Owens, D.G.C. (1999). A guide to the extrapyramidal side effects of antipsychotic drugs. Cambridge: Cambridge University Press.

Pantelis, C., Barnes, T.R.E., Nelson, H.E., et al. (1997). Frontal-striatal cognitive deficits in patients with chronic schizophrenia. Brain, 120, 1823-1843.

Prosser, E.S., Csernansky, J.G., Kaplan, J., et al. (1987). Depression, parkinsonian symptoms and negative symptoms in schizophrenics treated with neuroleptics. Journal of Nervous and Mental Disorders 175: 100-105.

Pullinger, S. & Tyrer, P. (1983). Acute lithium-induced tremor. British Journal of Psychiatry, 143, 40-41.

Purcell, R., Maruff, P., Kyrio, M. & Pantelis, C. (1997). Neuropsychological function in young patients with unipolar major depression. Psychological Medicine, 27, 1277-1285.

Quinn, N. (1995). Parkinsonism – recognition and differential diagnosis. British Medical Journal, 310, 447-452.

Rabins, P.V. (1982). Psychopathology of Parkinson's disease. Comprehensive Psychiatry, 23: 421-429.

Riekkinen, M., Kejonen, K., Jakala, P., Sininen, H., & Riekkinen, P., Jr. (1998).

Reduction of noradrenaline impairs attention and dopamine depletion slows responses in Parkinson's disease. European Journal of Neuroscience, 10: 1429-1435.

Rifkin, A., Quitkin, F., Howard, A., & Klein, D.F. (1975). A study of abrupt lithium withdrawal. Psychopharmacologia, 44: 157-158.

Rogers, D., Lees, A.J., Smith, E. et al. (1987). Bradyphrenia in Parkinson's disease and psychomotor retardation in depressive illness. Brain, 110: 761-776.

Sacks, O. (1973). Awakenings. London: Macmillan.

Salamé, P., Danion, J-M., Peretti, S. & Cuerro, C. (1998). The state of functioning of working memory in schizophrenia. Schizophrenia Research, 30, 11-29.

Santamaria J., Tolosa, E. & Valles, A. (1986). Parkinson's disease with depression: a possible subgroup of idiopathic parkinsonism. Neurology, 36: 1130-1133.

Saykin, A.J., Shtasel, D.L., Gur, R.E. et al. (1994). Neuropsychological deficits in neuroleptic naïve patients with first-episode schizophrenia. Archives of General Psychiatry, 51, 124-131.

Schatz, J. (1998). Cognitive processing efficiency in schizophrenia: generalised versus domain specific deficits. Schizophrenia Research, 30, 41-49.

Schou, M. (1959). Lithium in psychiatric therapy: stock-taking after ten years. Psychopharmacologia, 1: 65-78.

Schou, M., Baastrup, P.C., Grof, P. et al. (1970). Pharmacological and clinical problems of lithium prophylaxis. British Journal of Psychiatry 116: 615-619.

Serban, G., Thomas, A. (1974). Attitudes and behaviours of acute and chronic schizophrenic patients regarding ambulatory treatment. American Journal of Psychiatry, 136: 991-995.

Shallice, T. (1988). From neuropsychology to mental structure. Cambridge: Cambridge University Press.

Shaw, K.M., Lees, A.J., Stern, G.M. (1980). The impact of treatment with levodopa on Parkinson's disease. Q J Med, 49: 283-93.

Simpson, G.M. & Angus, J.W.S. (1970). A rating scale for extrapyramidal side effects. Acta Psychiatrica Scandinavica, Suppl. 212, 11-19.

Steck, H. (1954). Le syndrome extrapyramidal et diencephalique au cours des traitements au Largatil et au Serpasil. Annales Medico-Psychologiques, 112, 737-744.

Stip, E. & Lussier, I. (1996). The heterogeneity of memory dysfunction in schizophrenia. Canadian Journal of Psychiatry, 41, 145-205.

Tandberg, E., Larsen, J.P., Aarsland, D., & Cummings, J.L. (1996). The occurrence of depression in Parkinson's disease: A community-based study. Archives of Neurology, 53: 175-179.

Taylor, A.E., Saint-Cyr, J.A., & Lang, A.E. (1986). Frontal lobe dysfunction in Parkinson's disease. Brain 109: 845-883.

Taylor, M.A., Redfield, J., Abrams, R. (1981). Neuropsychological dysfunction in schizophrenia and affective disease. Biological Psychiatry 16: 467-478.

Todes, C.J. & Lees, A.J. (1985). The pre-morbid personality of patients with Parkinson's disease. Journal of Neurology, Neurosurgery and Psychiatry, 48: 97-100.

Tyrer, P., Alexander, M.S, Regan, A. & Lee, I. (1980). An extrapyramidal syndrome after lithium therapy. British Journal of Psychiatry, 136: 191-4.

Tyrer, P., Lee, I., Trotter, C. (1981). Physiological characteristics of tremor after chronic lithium therapy. British Journal of Psychiatry, 139, 59-61.

Van Putten, T. & May, P.R.A. (1978). Subjective response as a predictor of outcome in pharmacotherapy. Archives of General Psychiatry, 35, 477-480.

Van Putten, T. (1974). Why do schizophrenic patients refuse to take their drugs?

Archives of General Psychiatry, 31, 67-72.

Vaughan, I. (1986). Living with Parkinson's disease. London: Macmillan.

Verkerk, P.H., Schouten, J.P. & Oosterhuis, H.J.G.H. (1990). Measurement of the hand co-ordination. Clinical Neurology and Neurosurgery, 92, 105-109.

Wade, P., Gresty, M.A. & Findley, L.J. (1982). A normative study of postural tremor of the hand. Archives of Neurology, 39, 358-362.

Walsh, E.G. (1992). Muscles, masses and motion: the physiology of normality, hypertonicity, spasticity and rigidity. London: MacKeith Press.

Walsh, E.G. (1992). Muscles, masses and motion: the physiology of normality, hypertonicity, spasticity and rigidity. London: MacKeith Press.

Walsh, E.G. (1996). Research into trembling fingers: a personal experience. Proceedings of the Royal College of Physicians of Edinburgh, 26, 51-72.

Ward, C.D, Duvoisin, R.C., Ince, S.E. et al. (1984). Parkinson's disease in twins. In: Advances in Neurology Vol 40. Eds: RG Hassler, Christ JF. Raven Press, New York.

Watts, R.L., Wiegner, A.W. & Young, R.R. (1986). Elastic properties of muscles measured at the elbow in man: II Patients with parkinsonian rigidity. Journal of

Neurology, Neurosurgery and Psychiatry, 49, 1177-1181.

Webster, D.D. (1966). Rigidity in extrapyramidal disease. Journal of Neurosurgery, 24 (Suppl. part 2), 299-309.

Weiden, P.J., Dixon, L., Frances, A., Appelbaum, P., Haas G., & Rapkin B. (1991). Neuroleptic noncompliance in schizophrenia. In: Advances in Neuropsychiatry and Psychopharmacology, Volume 1: Schizophrenia Research. Eds: CA Tamminga & SC Schulz, Raven Press, Ltd., New York.

Wilson, S.A.K. (1925). Some disorders of motility and of muscle tone, with special reference to the corpus striatum. Lancet, ii, 1, 53, 169, 215, 268.

Wilson, S.A.K. (1947). Neurology. London: Arnold.

Wojcik, J.D., Gelenberg, A.J., LaBrie, R.A. & Miedke, M. (1980). Prevalence of tardive dyskinesia in an outpatient population. Comprehensive Psychiatry, 21, 370-379.

Zilboorg, G., and Henry, G.W. (1941). A history of medical psychology. New York: Norton.