

**University of Stirling**  
**Department of Psychology**

**Treatment Factors and  
Neuropsychological Outcome  
in Phenylketonuria**

**Peter V. Griffiths**

**Submitted for the degree of Ph.D.**  
**September 1997**

**This dissertation is dedicated to Dr Sonja McBean, formerly of the Royal Hospital for Sick Children, Glasgow, who referred my first patient with PKU back in 1971.**

## DECLARATION

In accordance with Regulation A7.2 of Ordinance 23, I affirm that this dissertation is entirely my own work and has not been included in any other thesis.

Signed Robt Smith .

Dated 18th December 1997.

# CONTENTS

	Page
<b>Acknowledgements</b>	5
<b>Abstract</b>	6
<b>Foreward</b>	7
<b>List of Tables</b>	9
<b>List of Figures</b>	10
<b>List of Abbreviations</b>	11
<b>Chapter One - Introduction</b>	13
<b>Chapter Two - Treatment issues</b>	38
<b>Chapter Three - Influence of post-treatment hyperphenylalaninaemia on neuropsychological functioning</b>	63
<b>Chapter Four - Experimental manipulation of phenylalanine: theoretical and methodological background</b>	84
<b>Chapter Five - Experimental manipulation of phenylalanine: pilot study of dietary resumption in treated off-diet adults</b>	100
<b>Chapter Six - Manipulation of phenylalanine by amino acid supplement: cognitive and motor tests</b>	121
<b>Chapter Seven - Manipulation of phenylalanine by amino acid supplement: IQ, everyday memory and personality tests</b>	147
<b>Chapter Eight - Association between phenylalanine control during treatment and executive function</b>	175
<b>Chapter Nine - Executive function measured by variants of the continuous performance test</b>	200
<b>Chapter Ten - Overview: treatment and policy</b>	227
<b>References</b>	257
<b>Appendices</b>	285

## **ACKNOWLEDGEMENTS**

I wish to express my gratitude to the following for their generous assistance and cooperation at various stages in the realization of this project.

*At the University of Stirling:* Professor Roger Watt, Dr Ronan O'Carroll, Dr Lindsay Wilson, Dr Robin Campbell, Dr Alistair Watson, the late Mr Roy Scott, Miss Lorraine Paterson, Miss Caroline Smith, Mr Nicholas Ward and Mr Mauro Tarrini.

*At the Royal Hospital for Sick Children, Glasgow:* Professor Forrester Cockburn, Dr Ann Harvie, Dr Peter Robinson, Mrs Brenda Clark, Mrs Christina Del Priore, Mrs Karen Cameron, Miss Kerri Newlands and Mrs Myra Ferguson.

*At Fallin Primary School, Stirlingshire:* Miss M. Flower and her secretarial and teaching staff

*At Scientific Hospital Supplies, Liverpool:* Miss Katrina Roberts.

Special thanks must go to all the young volunteers with phenylketonuria on the West of Scotland Register and their families.

Finally, I should like to acknowledge the forbearance of my family - Terry, Alyn and Sian - during the period of bringing this work to fruition.

## **ABSTRACT**

Phenylketonuria (PKU) is an inherited metabolic disease that affects about one in 10,000 of the population worldwide. In the classical form of the condition, the hepatic enzyme phenylalanine hydroxylase is absent or much reduced. If untreated, severe or profound mental handicap customarily results due to the accumulation of dietary phenylalanine (phe) which is neurotoxic. The mechanism by which phe impairs growth in the immature nervous system is little understood, but myelin metabolism appears to be disturbed. Treatment is by reduction of phe in daily food intake. Treatment should ideally begin in the neonatal period if intellectual loss is to be avoided. However, the safe range of phe concentrations during treatment and the age at which treatment can be discontinued without further damage being inflicted are uncertain. The studies reported in this volume investigated neuropsychological outcomes of treatment control and cessation factors. In addition, the question of whether executive functions are especially vulnerable to elevated phe concentrations during treatment was addressed. Patient samples conformed to the practice adopted in the West of Scotland regional centre for the management of PKU of maintaining dietary treatment until age 10 or beyond. Almost exclusively, negative findings emerged. These suggested that, if control of phe intake conforms to current UK recommendations for the preschool and primary years, neither global nor specific intellectual deficit result. Furthermore, the data supported the view that cessation of treatment at 10 years of age does not have harmful consequences. These findings have direct implications for the formulation of clinical policy on the treatment of PKU, but it must be recognized that the history of the successful treatment of PKU and mass screening for the disease spans a mere three decades. Thus, treatment outcome research to date is based only on children and young adults. In future investigations, a life-span approach will be required before the issues raised in this thesis can be finally settled.

## **FOREWARD**

The empirical core of this thesis is the six studies reported in Chapters Three, Five, Six, Seven, Eight and Nine. The chapters follow the chronology of the work. I was personally responsible for posing the questions addressed by the studies in Chapters Three, Five, Eight and Nine and for devising the methodologies. I formed part of a multi-disciplinary team consisting of paediatric, biochemistry and neurology colleagues that investigated the supplementation effects reported in Chapters Six and Seven. I was entirely responsible for the selection and, in some cases design, of all the psychological tests administered and for the analyses of the results derived therefrom. Blood samples were taken and biochemical assays were conducted by medical and biochemistry staff, mostly at Yorkhill and Stobhill Hospitals, Glasgow. The interpretations of the findings obtained by the research are my own. This thesis and the journal reports it embodies were written independently by myself though earlier drafts of papers and chapters have been commented on by academic and clinical colleagues.

In the studies described in Chapters Three, Six, Seven and Eight some of the data collection was conducted under my supervision by undergraduate students pursuing their studies in the Department of Psychology, University of Stirling. I have acknowledged their help above.

At the time of writing, three reports have been published (Chapters Three, Seven and Eight), one has been accepted (Chapter Six) and one submitted (Chapter Nine).

These are the publication details to date.

### **Chapter Three**

Griffiths, P., Paterson, L. & Harvie, A. (1995) Neuropsychological effects of subsequent exposure to phenylalanine in adolescents and young adults with early-treated phenylketonuria. *Journal of Intellectual Disability Research*, **39**, 365-372.

### **Chapter Six**

Griffiths, P., Ward, N., Harvie, A. & Cockburn, F. Neuropsychological outcome of experimental manipulation of phenylalanine intake in treated phenylketonuria. *Journal of Inherited Metabolic Disease*. (Accepted, June 1997.)

### **Chapter Seven**

Griffiths, P., Smith, C. & Harvie, A. (1997) Transitory hyperphenylalaninaemia in children with continuously treated phenylketonuria. *American Journal on Mental Retardation*, **102**, 27-36.

### **Chapter Eight**

Griffiths, P., Tarrini, M. & Robinson, P. (1997) Executive function and psychosocial adjustment in children with early treated phenylketonuria: correlation with historical and concurrent phenylalanine levels. *Journal of Intellectual Disability Research*, **41**, 317-323.

### **Chapter Nine**

Griffiths, P., Campbell, R. & Robinson, P. Executive function in treated phenylketonuria as measured by the one-back and two-back versions of the continuous performance test. *Journal of Inherited Metabolic Disease*. (Accepted, October 1997.)

## LIST OF TABLES

<b>Table Number</b>		<b>Chapter</b>	<b>Page</b>
<b>1</b>	Incidence of PKU	<b>1</b>	<b>24</b>
<b>2</b>	IQ and dietary cessation	<b>2</b>	<b>52</b>
<b>3</b>	Treatment discontinuation	<b>2</b>	<b>55</b>
<b>4</b>	Test battery means and SDs	<b>3</b>	<b>72</b>
<b>5</b>	F-ratios	<b>3</b>	<b>75</b>
<b>6</b>	Demographic data	<b>5</b>	<b>103</b>
<b>7</b>	Test scheme	<b>5</b>	<b>106</b>
<b>8</b>	Phe changes off and on diet	<b>5</b>	<b>110</b>
<b>9</b>	Test battery scores	<b>5</b>	<b>111</b>
<b>10</b>	Change scores	<b>5</b>	<b>111</b>
<b>11</b>	Neurological test data	<b>5</b>	<b>114</b>
<b>12</b>	MRI summary	<b>5</b>	<b>115</b>
<b>13</b>	Manipulation: neuropsychological tests	<b>6</b>	<b>133</b>
<b>14</b>	Cross-over effects	<b>7</b>	<b>163</b>
<b>15</b>	Guesses about manipulation	<b>7</b>	<b>164</b>
<b>16</b>	Phe and test correlations	<b>8</b>	<b>189</b>
<b>17</b>	Glasgow and Munster mean IDCs	<b>8</b>	<b>197</b>
<b>18</b>	Treatment parameters	<b>9</b>	<b>217</b>

## LIST OF FIGURES

<b>Figure Number</b>		<b>Chapter</b>	<b>Page</b>
<b>1</b>	Structure of phenylalanine and tyrosine	<b>1</b>	<b>26</b>
<b>2</b>	Metabolism of phenylalanine	<b>1</b>	<b>28</b>
<b>3</b>	Decline in IQ	<b>2</b>	<b>40</b>
<b>4</b>	IQ distributions	<b>2</b>	<b>42</b>
<b>5</b>	IQ and historical phenylalanine level	<b>2</b>	<b>49</b>
<b>6</b>	Null hypothesis model	<b>3</b>	<b>64</b>
<b>7</b>	Alternative hypothesis model	<b>3</b>	<b>65</b>
<b>8</b>	Index of dietary control	<b>3</b>	<b>69</b>
<b>9</b>	Test battery results	<b>3</b>	<b>74</b>
<b>10</b>	Neonatal phe control	<b>3</b>	<b>82</b>
<b>11</b>	First decade IDCs	<b>5</b>	<b>104</b>
<b>12</b>	Phe manipulation effects	<b>6</b>	<b>131</b>
<b>13</b>	Single crossover model	<b>6</b>	<b>144</b>
<b>14</b>	Groups One and Two IDCs	<b>7</b>	<b>157</b>
<b>15</b>	Dietary control, years 1 to 9	<b>8</b>	<b>185</b>
<b>16</b>	Historical phe-level profile	<b>9</b>	<b>207</b>
<b>17</b>	Hit rate per block	<b>9</b>	<b>212</b>
<b>18</b>	Dot pattern task	<b>9</b>	<b>223</b>
<b>19</b>	Actual vs recommended phe levels	<b>10</b>	<b>251</b>

## LIST OF ABBREVIATIONS

<b>Abbreviation</b>	<b>Meaning</b>
<b>ANCOVA</b>	analysis of covariance
<b>ANOVA</b>	analysis of variance
<b>BMDP</b>	Bio-Medical Data Package
<b>BPVS</b>	British Picture Vocabulary Scale
<b>CA</b>	chronological age
<b>CNS</b>	central nervous system
<b>CPT</b>	Continuous Performance Test
<b>CRT</b>	choice reaction-time
<b>DHSS</b>	Department of Health and Social Security
<b>DoB</b>	date of birth
<b>EMG</b>	electromyogram
<b>hyperphe</b>	hyperphenylalaninaemia
<b>IDC</b>	index of dietary control
<b>IQ</b>	intelligence quotient
<b>ISI</b>	inter-stimulus interval
<b>ITI</b>	inter-test interval
<b>MFFT</b>	Matching Familiar Figures Test
<b>MRC</b>	Medical Research Council
<b>MRI</b>	magnetic resonance imaging
<b>PAH</b>	phenylalanine hydroxylase
<b>phe</b>	phenylalanine
<b>PKU</b>	phenylketonuria

## **LIST OF ABBREVIATIONS (Continued)**

<b>Abbreviation</b>	<b>Meaning</b>
<b>RBMT</b>	Rivermead Behavioural Memory Test
<b>RHSC</b>	Royal Hospital for Sick Children (Glasgow)
<b>RVLT</b>	Rey Verbal Learning Test
<b>TMS</b>	transcranial magnetic stimulation
<b>tyr</b>	tyrosine
<b>VRT</b>	Visual Retention Test
<b>WISC</b>	Wechsler Intelligence Scale for Children
<b>WMS</b>	Wechsler Memory Scale
<b>WRAT</b>	Wide Range Achievement Test

## **Chapter One**

### **INTRODUCTION**

#### **Phenylketonuria: the disease and its treatment**

Phenylketonuria (PKU) is a recessively inherited metabolic disease that falls into the category of aminoacidopathies. In PKU the conversion of the essential amino acid phenylalanine (phe) to tyrosine is impaired because of a deficiency of the hepatic enzyme phenylalanine hydroxylase (PAH).

Consequently, abnormally high levels of phe accumulate in the blood and body tissues. Excess, unmetabolized phe is toxic to the immature central nervous system (CNS). Left to run its natural course, PKU typically results in severe or profound mental handicap and associated motor dysfunction. However, if detected in the neonatal period, the disorder can be treated and intellectual deficiency prevented (Thompson & O'Quinn, 1979).

In its untreated form, PKU is characterised by physical features such as fair skin and hair, blue eyes and a mouse-like body odour. The urine of the sufferer contains phenylketones (Cleland, 1978). These aromatic compounds give the urine a distinctive smell from which the name phenylketonuria originated, though terms like hyperphenylalaninaemia (abnormal level of phe in blood) or PAH deficiency more accurately describe the essential condition. Nowadays, in the Western world, children with untreated PKU are almost never seen, though adults can be found throughout institutional and community facilities for the learning disabled.

PKU is thus an enzyme deficiency and belongs within the category of 'inborn errors of metabolism', an expression coined by Garrod (1909). It is the commonest disease entity within this nosological grouping, affecting on average about one in 8,000 live births in Western Europe (Smith & Brenton, 1995). PKU was the first demonstration of a metabolic disease that could adversely affect intellectual ability.

PKU is a 'silent' disease (Centerwall & Centerwall, 1961). The baby with PKU appears normal and asymptomatic. Before birth, excess phe crosses the placental membrane where it is metabolized by the normal maternal enzyme. Only later in post-natal life, when the unusual body odour, depigmentation of the skin and signs of deviant development become apparent, will clinical suspicion be raised. By this time effective treatment is impossible as the neuropathological process can only be arrested, not reversed. Biochemical analysis of a capillary blood sample, usually taken around the end of the first week of life, will indicate whether there are abnormally high levels of phe circulating in the bloodstream. Persisting hyperphenylalaninaemia confirms the disease and treatment must be initiated immediately to prevent progressive brain damage.

Treatment of PKU is by dietary prophylaxis and consists of severely restricting foodstuffs that contain phe. As yet, enzyme replacement and gene therapy are only theoretical possibilities. Phenylalanine is present in all protein; therefore meat, eggs, fish, poultry and dairy products are virtually excluded; as are vegetables and cereals that have a high protein content, such as pulses and wheat. To compensate for the protein deficiency that would result from such a narrow food intake, a dietary supplement containing essential amino

acids other than phe has to be administered. This formula is manufactured by firms specialising in dietary products, nowadays usually from synthesized chemicals.

The protein restricted diet and supplement are normally managed on a day-to-day basis by the PKU child's caretakers under the supervision of a professional dietician. The patients' and parents' organisation in the UK, the National Society for Phenylketonuria, publishes menus and recipe books to help carers provide variety at mealtimes. Blood phe levels are customarily monitored weekly to ensure maintenance within the therapeutic range and to ascertain the degree of compliance with the diet. The UK currently has around 38 laboratories providing population screening tests for newborns and biochemical assays for identified patients, this service having been established as a result of Government recommendations for nationwide coverage in 1979 (Wolff, 1981). In Scotland, a single centre at Stobhill Hospital, Glasgow, serves the entire region, screening roughly 60,000 infants a year.

Psychological interest has centred principally on developmental and cognitive outcome variables in relation to treatment factors such as age of commencement and age of cessation of dietary therapy. Severity is also a factor. PKU is a genotypically heterogeneous condition and, in the phenotype, phenylalanine hydroxylase activity in the liver can vary from mild deficiency to nil. Neuropsychological research continues to focus on issues such as the effects on brain growth and cognitive function of dietary compliance and dietary continuation. A core issue that has direct implications for treatment policy is whether an age is ever reached at which normal eating

habits can be allowed without fear of damaging the CNS and impairing mental and behavioural functioning.

PKU is a Mendelian recessive disorder. The carrier frequency in temperate-zone populations is 2% and both parents have to be carriers for the genetic mutation to express itself. Even so, the probability is only one in four that a child born to two carriers will have the disease (Harris, 1970). Children born to individual mothers and fathers with PKU will not have the condition unless the partner is a carrier. Genetic counselling of parents identified as carriers following the birth of a PKU child and older PKU patients planning a family forms an integral part of contemporary treatment. The economics of mass genotyping precludes population screening for carriers at the present time. It is imperative that off-diet girls with PKU who are contemplating pregnancy are retreated with a supplemented, phe-restricted diet pre-conceptually and during gestation. Unless maternal blood phe levels are reduced, foetal CNS damage and malformation of other organs will almost inevitably result from the passage of phe-laden blood across the placenta, even though the unborn child does not itself have PKU (Smith *et al.*, 1990b).

In summary, PKU is a genetic disorder inherited by a child from two apparently healthy parents that prevents the normal utilization of protein food by the body. During digestion, protein is broken down into its constituent parts, one of which is phenylalanine. However, the liver in PKU is deficient in the enzyme that converts phe into tyrosine which is the next step along the metabolic pathway that would otherwise lead to the formation of, amongst other biochemically active substances, the catechol hormones adrenalin and noradrenaline. Unmetabolized phe is neurotoxic to the growing CNS.

Although the precise mechanism by which it affects neural structures is unknown, the overall process by which the genetic aberration becomes translated into brain damage in the immature organism is indirect, multi-stage and insidious. It can be conceptualised simply as follows.

Genetic mutation → Phenylalanine hydroxylase deficiency → Metabolic block → Biochemical imbalance → Brain damage

## **Early history**

The history of present-day knowledge of PKU is founded on the work of three men: Asbjorn Folling, a Norwegian physician and biochemist; Horst Bickel, a German paediatrician; and Robert Guthrie, an American cancer specialist. Guthrie switched to research into the prevention of mental retardation resulting from inborn errors of metabolism following the birth of a mentally handicapped son. Respectively, they formulated solutions to the pathogenesis of PKU, its treatment, and its detection in the newborn population.

PKU was discovered by Asbjorn Folling, Professor of Nutritional Research at the Physiological Institute of the Veterinarian University of Oslo in 1934 (Folling, 1971). In Norway it is still referred to as Folling's disease, in recognition of Folling's pioneering work, though he gave it the name 'oligophrenia phenylpyruvica'. This Latinized term and its alternative 'phenylpyruvic oligophrenia', with their allusions to feeble-mindedness, only enjoyed a brief period of currency as diagnostic labels. Shortly before World

War Two, Penrose and Quastel (1937) suggested the description 'phenylketonuria' and this is the term in use today.

The story of the discovery of PKU is a tale of scientific persistence and chance favouring the prepared mind. A mother brought her two developmentally retarded children, a boy of four and a girl of six, to Folling's clinic for an opinion about their peculiar odour (Folling, 1994). The children's father was asthmatic and so distressed by their smell that he was unable to tolerate their presence close by. A colleague of the father recommended consulting Folling because of his earlier postgraduate work on acidosis.

Folling found no clinical signs in the children other than mental retardation and the distinctive, acetone-like aroma of their urine. Because of the strong suspicion of ketones being excreted, he tested their urine for diabetes by adding ferric chloride. Normally, ferric chloride turns urine purple in the presence of diacetic acid, but instead it turned green. This reaction had never been seen by Folling and it was not described in the literature (Folling, 1971). Folling used classical organic chemistry and crystallography to determine the molecular composition of the unknown substance and established it as phenylpyruvic acid, a substance not normally present in urine.

Folling went on to survey 430 residents in hospitals for the mentally retarded and found the green reaction in the urine of eight. All these individuals were excessively fair in colouring and had eczema and spasticity. He correctly hypothesized that the syndrome was an inherited disorder of phenylalanine metabolism (Guttler, 1984). In exploring this idea, he was the first to use bacteria as a tool to measure the amount of phe in blood and urine and found

high levels of phe in the blood of patients diagnosed clinically and by the ferric chloride test. He suggested a block in the normal metabolism of phe which, on the basis of his observations of the disease in siblings and consanguineous families, he considered was inherited as a recessive trait (Bickel, 1980).

Between 1945 and 1953, biochemical research on normal aromatic amino acid metabolism clarified that the main catabolic pathway of phenylalanine was its conversion to tyrosine in the presence of PAH. Jervis (1953) showed how biopsied liver samples from controls could catalyze this reaction *in vitro* but liver samples from individuals with PKU could not, suggesting the absence of hepatic PAH in PKU.

Almost twenty years after Folling's accidental discovery of a positive urine test for PKU, the first report of an experimental treatment was published as a case study (Bickel *et al.*, 1953). In 1949, Horst Bickel moved from Zurich to the Birmingham Children's Hospital, as a research fellow, where routine ferric chloride tests were being applied to the urine of all mentally handicapped children attending outpatient clinics. In 1951, a colleague, John Gerrard, diagnosed PKU in a two-year-old, retarded Irish girl. Bickel confirmed the diagnosis by paper chromatography. In addition to the green phenylpyruvic acid reaction, the girl had fair hair, eczema and a mouse-like smell. Her mother pressed for treatment. Bickel wondered if excess phenylalanine itself had caused the child's brain damage and whether restricting its intake would improve the condition.

Bickel and his colleagues treated the girl by reducing phe with a protein-

restricted diet supplemented with a protein formula free of phe, casein hydrolysate, first mentioned to Bickel by Louis Woolf, a biochemist at Great Ormond Street Hospital, London. The therapeutic trial began in 1951, when the girl was 2 years 2 months old. Her behaviour improved in a few months. She sat upright, walked, interacted with people and showed interest in her surroundings; her hair darkened and her ketotic smell disappeared; she made up six months of development in a year as measured by the Terman-Merrill. Temporarily loading her diet with phe caused her to revert to her pre-treatment state. When discharged, sadly, her mother was unable to sustain the diet. The child was thus the first example of a treatment success and a compliance failure. Now in her mid-forties, she is severely mentally retarded and institutionalized (Bickel, 1996). Shortly after Bickel's publication, Woolf *et al.* (1955) observed that treatment with a low phenylalanine diet led to a reduction in neurological symptoms such as seizures and gait disorder but that IQ improvements were less dramatic.

The findings of these early reports pointed to the conclusion that the best results could probably be achieved by early detection and implementation of the preventative diet. This implied screening every newborn for the metabolic error. Early attempts employed a development of Folling's ferric chloride test: the Phenystix method. Phenystix kits contained strips impregnated with ferric chloride which, if soaked with urine from the infant's nappy, would show the green reaction if phenylpyruvic acid was present in large quantities (Heatonward & Wiley, 1984). Kits were generally issued to district nurses and health visitors for use during post-natal follow-up.

The measure yielded too many false negatives (Medical Research Council,

1981). As PKU babies were not always detected, and hence suffered the ravages of the disease, the shortcomings of the Phenistix method stimulated a search for a way of quantifying phe in the blood. A further impetus to finding a simple method for measuring blood phe was the increasing need to monitor blood levels of patients during dietary treatment to ascertain the degree of compliance.

With a background of many years in cancer research, Robert Guthrie surmised that the technique of competitive inhibition, used for screening the blood of patients undergoing cancer treatment, might be modifiable for blood screening in PKU (Guthrie, 1996). The principle underlying the technique is that an inhibitor compound, which normally prevents growth of *Bacillus Subtilis*, is neutralised by the presence of abnormal amounts of phe in culture, thereby allowing bacterial proliferation and showing a visible change in the appearance of the medium. Guthrie collected blood from a developmentally delayed niece of his who, at 15 months of age, was diagnosed late as having PKU. He pricked the heel, collecting the blood drops on filter paper for later assay in the laboratory. Apart from the procedure's practicality and simplicity, the appeal of collecting blood on absorbent paper lay in its being widely accepted by professionals and parents. Neonatal screening for PKU effectively began in 1961 when Guthrie started receiving filter paper specimens of blood from newborns in two Jamestown hospitals in New York (Guthrie & Susi, 1963). In the next few years, almost half a million babies in 29 states were screened for PKU by means of the now-familiar 'Guthrie cards' - and 39 cases were found. Today, all 50 states require newborn screening for PKU by law and mass screening programmes are in place throughout the North American continent, Europe, Australasia and the developed countries of

the world.

In the UK, routine screening for PKU by means of the Phenystix method was introduced in the late 1950s and conducted mainly by health visitors.

Between 25 and 50 per cent of PKUs were missed either because of false negatives or the test not being performed and, subsequently, the Phenystix approach to diagnosis was abandoned in favour of the microbiological inhibition assay (Guthrie testing). In Scotland, Phenystix were replaced by Guthrie tests as early as 1965. In the decade from 1968 to 1978, the detection rate for PKU improved radically with neonatal miss-rates falling to around one per cent for the UK as a whole (Wolff, 1981).

Research into PKU treatment paralleled the inauguration of screening programmes and the need for a concerted approach to treatment issues was recognised early in Europe and the USA. In the UK, a National Register for PKU was set up in Liverpool by Freddie Hudson under the auspices of the Medical Research Council and the, then, Department of Health and Social Security. The date was 1964. Its purpose was 'to document the clinical and intellectual progress of patients treated by means of a low phenylalanine diet' (Smith & Wolff, 1978). After Hudson's death in 1976, the Register moved to London, Isabel Smith taking over the directorship. Biochemical and intelligence test results from centres throughout the British Isles, including Eire, were sent to London and entered into a national database.

A similar project arose in the USA. The United States National Collaborative Study of Children Treated for Phenylketonuria was initiated in 1967, though at that time only 19 centres in 12 states actually submitted returns to the Los

Angeles headquarters (Koch *et al.*, 1985). The purpose of the USA Collaborative Study was explicitly stated in terms of two questions: 1) 'is dietary restriction of phenylalanine intake an effective treatment for phenylketonuria during the first six years of life?' and 2) 'is dietary restriction of phenylalanine necessary in the treatment of phenylketonuria after six years of age?' (Williamson *et al.*, 1985). The answer to the second question remains incomplete. Research in both the USA and UK was organised on a cohort basis with all children detected and treated during the first five years of each project being followed up for the first six years of life before subsequent analysis of IQ outcome data was undertaken.

Following his demonstration of the treatability of PKU, Bickel returned to Germany from Birmingham. In 1978 he inaugurated the German Collaborative Study of PKU, a multicentre project based in Heidelberg collating data from paediatric departments in Berlin, Dusseldorf, Gottingen, Hamburg, Heidelberg, Munich, Munster and Ulm. Almost forty years after Folling's discovery, Bickel published results of intelligence testing from the first longitudinal study of PKU in Western Europe based on mass screening (Bickel *et al.*, 1973), although small sample research findings of IQ outcome had been reported earlier (*e.g.* Berman *et al.*, 1966; Kang *et al.*, 1970).

### **Incidence of phenylketonuria**

Epidemiological findings emerged from work on population screening well in advance of treatment results as data suitable for determining incidence could be collected immediately following successful establishment of national

programmes. It soon became apparent that the geographical distribution of PKU was variable both within and between countries.

Between 1969 and 1970, 63 per cent of newborns in the UK were being screened by Guthrie testing. By 1974, this figure had risen to 99 per cent of live births. Thus, incidence figures virtually equate with prevalence. The following table is adapted from frequency data derived from the MRC/DHSS Phenylketonuria Register during the period 1974 to 1976 inclusively for England, Scotland, Wales and Northern Ireland and provides a comparison with 1975 data, in brackets, for Eire (Smith & Wolff, 1978).

	<b>Number screened</b>	<b>PKU cases identified</b>	<b>Incidence per 100,000</b>
England and Wales	1,809,720	174	9.61
Scotland	201,119	18	8.95
Northern Ireland	79,065	22	27.83
(Eire	67,269	35	26.90)

**Table 1.** Annual Incidence of PKU in the British Isles from 1974 to 1976.

The above incidence figures relate to all cases of infants with persistently elevated phe levels. This early survey confirmed the now accepted average prevalence estimate of about 1 per 10,000 for the UK, but revealed an

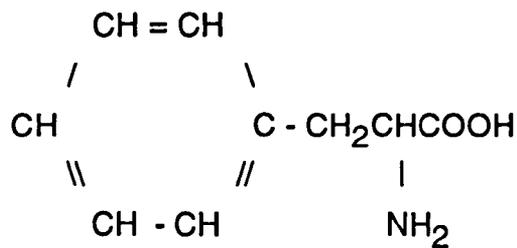
imbalance between mainland Britain and Ireland - both the North and the Republic - with the Irish figures being around three times greater than those found on the mainland. A breakdown of the figure for England and Wales showed substantial regional differences, with a low rate in East Anglia (5.12 per 100,000) and high rates in North West and South West England (14.62 and 12.15 respectively). Modern population genetic studies suggest that mutations of the gene coding for PAH fall into two broad locational categories in the UK: an eastern, probably reflecting the migration and settlement patterns of the Scandinavian races; and a western, probably reflecting those of the Celtic peoples (Scriver *et al.*, 1996). The mixed ancestry of inhabitants of the British Isles rather than intragenic crossover more likely accounts for the heterogeneity of genotypes in the UK and Ireland (Tyfield *et al.*, 1991).

By 1971, mass screening of over five million neonates in eight European countries had been achieved, 668 cases being identified as having classical PKU (persisting blood phe levels over  $1200\mu\text{mol/l}$ ). The incidence rate was found to vary considerably in different countries from 1 per 3000 in Ireland to 1 per 28,000 in Belgium (Guttler, 1984). World-wide, Northern Ireland has the highest rate, Japan the lowest. The ratio for Scotland is 1 per 8200 (Guthrie, 1980), the preponderance of cases being in the West. Figures as high as 1 per 40 have been found in populations where there is a high degree of consanguinity such as Gypsies (Nyhan, 1981). About 70 children with PKU are born in the UK each year (Clark *et al.*, 1995).

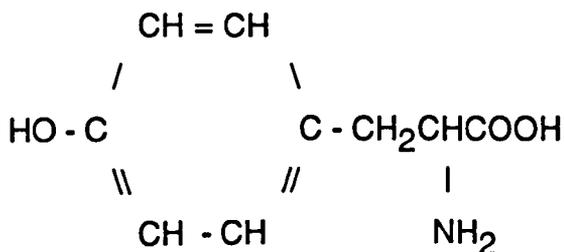
## Biochemistry

Phenylalanine is an essential amino acid: that is to say, it cannot be synthesized by the body but has to be obtained from food intake. All natural protein contains phe which is found linked to other amino acids like glycine, valine, leucine and cystine in polypeptide chains. These giant, linear polymers consist of concatenations of up to 300 amino acid molecules. During digestion, peptidases in the small intestine attack peptide links, breaking down proteins into 20 single amino acids. These are then transported to conversion sites such as the liver by the bloodstream (Brody, 1994).

The molecular structure of phenylalanine is a phenyl radical with alanine (methyl, carboxyl and amino groups attached to a CH hub):



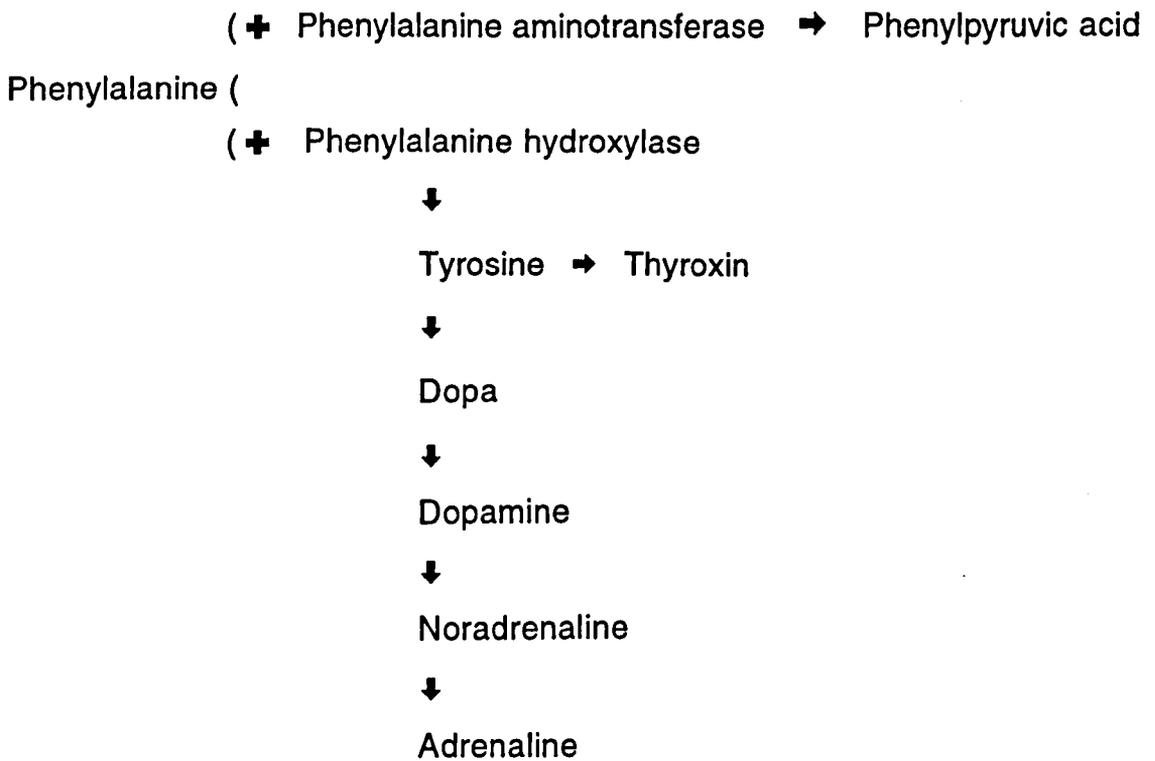
When hydroxylated by PAH, tyrosine is formed:



**Figure 1.** Molecular structures of phenylalanine and tyrosine.

In normal phenylalanine metabolism, phenylalanine is hydroxylated to tyrosine (tyr) and transaminated to phenylpyruvic acid (phenylpyruvate) and its metabolites, phenyllactic acid (phenyllactate) and phenylacetic acid (phenylacetate). The complete phe hydroxylating system only occurs in the liver in humans and comprises the three enzymes phenylalanine hydroxylase, dihydropteridin reductase and dihydrofolate reductase; and the cofactor tetrahydrobiopterin (Scriver & Clow, 1980). The metabolites of phenylpyruvic acid are excreted in the urine, but not phenylpyruvic acid itself. The next stage along the metabolic pathway is the conversion of tyr to the three hormones thyroxin, noradrenaline and adrenaline (Morgan, 1965). Thyroxin is a compound of tyr and iodine. Formed in the thyroid gland, it governs energy metabolism. The catecholamines, noradrenaline and adrenaline, stimulate the sympathetic branch of the autonomic nervous system. Acting as neurotransmitters, they are released by catecholaminergic neurones. Dihydroxyphenylalanine (dopa) and dopamine are precursors of noradrenaline and adrenaline, dopamine being an important neurotransmitter and implicated in psychosis. At the dopaminergic synapse dopamine is either released or destroyed by monoamine oxidase (Bender, 1993).

Phenylalanine is thus necessary for the formation of key hormones and neurotransmitters, but a nutritionally varied diet ensures sufficient intake. A simplified scheme of phenylalanine metabolism is as follows.



**Figure 2.** Normal metabolism of phenylalanine.

Phenylketonuria is the commonest disorder among the group of metabolic diseases collectively known as the hyperphenylalaninaemias, accounting for about 95% of cases. The two main variants, dihydropteridin reductase and tetrahydrobiopterin deficiency, are caused by impaired coenzyme synthesis. Dihydropteridin reductase deficiency cannot be treated whereas tetrahydrobiopterin deficiency responds to daily doses of biopterin (Smith & Brenton, 1995).

Phenylketonuria can best be regarded as phenylalanine hydroxylase deficiency, the metabolic block occurring at the point in the pathway where phe is converted to tyr (see Figure 2). The severity of the disease depends on the degree of PAH activity. In 'classical' PKU, there is little or no PAH activity

and phe accumulates rapidly in the bloodstream and tissues. Classical PKU is operationally defined as phe levels in blood consistently in excess of 1200  $\mu\text{mol/l}$  (Bickel *et al.*, 1973). In the West, neonates with blood levels of this magnitude or greater are always treated.

Mild hyperphenylalaninaemia (hyperphe) is indicative of a reduction of PAH activity rather than its absence or near-absence. Children with very mild hyperphe whose levels do not consistently exceed 400  $\mu\text{mol/l}$  are not usually treated (Medical Research Council, 1993a). The condition is regarded as benign, and intellectual and neurological impairment have not been demonstrated in later life (Weglage *et al.*, 1996).

If classical PKU is untreated, high concentrations of circulating phe result. In the absence of PAH, phe is metabolized to phenylpyruvic acid, phenyllactate and phenylacetylglutamine which are excreted in urine in approximately equal amounts. Phe is a) decarboxylated to phenylethylamine, 90 per cent of which is oxidized to phenylacetic acid and 10 per cent to mandelic acid; and b) transaminated to phenylpyruvic acid, about half of which is converted into phenyllactate and hydroxyphenylacetic acid (Fulton *et al.*, 1980).

Phenylpyruvic acid permits the green ferric chloride diagnostic test.

Phenylpyruvic acid and phenylacetic acid are ketones and are probably sweated as well as excreted, giving rise to the typical aroma.

Because of the metabolic block, blood tyrosine levels remain very low.

Normally, some tyrosine is catabolized to the brown pigment, melanin, by the enzyme tyrosinase. Melanin is found in skin, hair and eyes; sunlight increases production, resulting in tanning. However, in addition to failures of

hormone production, tyrosine deficiency affects melanin synthesis in untreated PKU. Furthermore, the increased phe pool inhibits tyrosinase. Depigmentation or partial albinism are characteristic, producing the fair-haired and blue-eyed appearance found when the disease has been left to run its natural course (Salway, 1994).

## Genetics

Phenylketonuria is an autosomal recessive disorder whose transmission follows Mendelian principles (Jervis, 1939). Almost all (99%) of hyperphenylalaninaemia mutations map on to the phenylalanine hydroxylase gene on chromosome 12. Lidsky *et al.* (1985) first localized the mutations for PKU on the long arm of chromosome 12. Woo and his team in Texas found the gene that coded for PAH and identified 12q24.1 as the locus for the great majority of mutations (Eisensmith & Woo, 1991). Analysis of 3986 mutant chromosomes by 81 investigators in 26 countries has, to date, identified 243 different mutations in 788 different associations (Scriver *et al.*, 1996). By the end of the 1980s, some 18 mutations only had been defined (Cotton, 1990). Despite the enormous genetic heterogeneity of the condition, a few PKU alleles occur at high relative frequencies and often on only one haplotype. About two-thirds of PKU genes in Europeans carry just four PAH mutations. Thus, within a population there are a few prevalent mutations and a number of rare ones. These concentrations suggest positive selection and recurrent mutation in the past. Around 75 per cent of European PKU genotypes are compounds which may first have arisen 10,000 to 20,000 years ago when major demic immigration from the Middle East into Northern Europe is thought

to have occurred. The high frequency of PKU - in terms of abnormal genetic traits - is probably not due to a high mutation rate of the human PAH gene but rather the expansion of a handful of mutant alleles in the population (Woo, 1989).

The PAH enzyme is a protein consisting of 452 amino acids arranged in a genetically-determined order. Each PAH gene has 13 coding sequences (exons). Mutations occur when the units in the coding sequence become exchanged. This results in a wrong amino acid in the enzyme string. Different mutations exert different influences on PAH activity. Variation in the severity of PKU - classical, moderate or mild - is dictated by the combination of mutant genes contributed by each parent carrier. Thus, PKU is a spectrum of PAH deficiency disorders. Guttler (1996) has proposed that genotyping of newborns will, in the near future, provide a method for precise diagnosis of the severity of PKU, as there is a high correlation between mutation genotype and blood phe level.

## **Neuropathology**

The neuropathogenic mechanism by which the growing CNS is damaged in untreated PKU has yet to be described. A mouse model of PKU has been created by McDonald (1994) who produced PAH deficiency in the normal mouse by administration of the germline mutagen ethylnitrosourea.

Laboratory animals so treated possess the characteristics of human PKU such as inter-individual variation in hepatic PAH activity and hypopigmentation that abates when treated by a low-phe diet. Animal research has the potential to

circumvent some of the practical and ethical difficulties surrounding controlled experimentation with human subjects, not the least of which is the rarity of the condition. But the promise of further understanding of the pathophysiology of PKU by means of animal studies has yet to be fulfilled.

Controversy even exists about whether phe itself is the neurotoxic agent. Following the proposal by Kaufmann (1977) that phe is the culprit, most authorities subscribe to this view, a conclusion consistent with animal studies in which high doses of phe injected in infancy have been shown to result in brain damage in the mature organism (Hommes *et al.*, 1982). However, Fulton and his co-workers theorised that phenylacetic acid is the neurotoxic agent, its action interfering with synaptic development and ultimately producing mental handicap. These authors were able to demonstrate that the other metabolites of phe (phenylpyruvic acid, phenyllactate and mandelic acid) are harmless, but were initially unable to separate the effects of phenylacetic acid from those of phe in an empirical study (Fulton *et al.*, 1980). In a later study, the same group injected experimental animals with phenylacetic acid and found selective impairment of cholinergic and GABAergic synaptic terminals but not of noradrenergic, serotonergic and glutamatergic terminals (Loo *et al.*, 1984). Perturbation in the untreated CNS appears much more widespread than the narrow range of synaptic effects associated with excess phenylacetic acid reported by Fulton and his colleagues, and the phenylacetic acid theory has not received subsequent support.

The candidate most favoured by present-day theorists as the target for phe is the neurone, the formation of myelin sheath having received particular

attention (Gordon, 1976). Early brain biopsy studies of patients with untreated PKU consistently revealed disturbances in myelin development. Alvord *et al.* (1950) were among the first to show neuropathological alterations indicating arrested myelin growth. Corsellis (1953) and others confirmed the presence of disrupted myelination. In autopsies of eight severely mentally retarded patients who died between the ages of 9 and 40 years of age, Malamud (1966) found no evidence for neuronal degeneration, but extensive white matter changes. Frank demyelination was only sporadically observed in severely handicapped adult cases, but sponginess and arrested myelination were ubiquitous in the brains of both adults and children. He concluded that a dysmyelination pattern was characteristic in younger patients whereas active demyelination seemed confined to older age groups.

Shah *et al.* (1972) found myelin deposition in untreated PKU reduced to 62 per cent of normal, but concluded that the general pattern was more suggestive of dysmyelination rather than demyelination as pathological changes were indicative of an increase in the rate of myelin turnover rather than decomposition.

On the basis of microscopic histological analyses of the brains of three adults with untreated PKU, Bauman & Kemper (1982) concurred with Shah *et al.*'s interpretation, finding no evidence for demyelination or decomposition. Instead they found widespread abnormalities in myelin deposition. In addition, they found poor dendritic arborization. What was most notable in Bauman and Kemper's data was the selective sparing of structures like the oculomotor nucleus, which matures pre-natally, compared with disruption of cells in areas like association cortex which have a long post-natal cycle of

myelination. This apparently time-structured pattern led them to propose that the metabolic abnormality in PKU interferes with or disrupts normal growth processes rather than damaging existing structures.

Hommes (1991) argued that elevated concentrations of phe produce adverse neuronal effects at all stages of development by means of a dysmyelination mechanism. It was clear from animal models of PKU that phe increases myelin turnover in the immature brain (Vorhees *et al.*, 1981). In an earlier animal study, Hommes subjected 25 day-old rats to a high phe supplement. At this age the rat brain is considered mature, yet he found phe-related disturbance of myelin biosynthesis (Hommes *et al.*, 1992). In human subjects, the small amount of evidence available is consistent with the view that the deleterious effects of phe depend on the maturational stage of the nervous system. For instance, Bick *et al.* (1991) compared magnetic resonance imaging (MRI) scans of children who were early and late treated. MRI is highly sensitive to myelin change. These authors found the severest white matter abnormalities in those who were either treated very late, very poorly or both. Only in this group of patients were there signs of subcortical involvement. In early treated patients, these areas were unaffected, the only minor anomalies being found in the posterior periventricular areas. Similar findings were reported by Thompson *et al.* (1990) whose MRI scans of the brains of early-treated, off-diet patients in adulthood showed highly localized rather than diffuse white matter anomalies, periventricular locations again being those most commonly found.

Bick and his colleagues suggested that white matter disturbance in PKU needs to be understood within the context of the normal sequence of

myelination. From an autopsy survey of 162 infantile deaths, Kinney *et al.* (1988) formulated the following rules about early myelination.

- 1) Proximal pathways precede distal.
- 2) Sensory pathways precede motor.
- 3) Projection areas precede association.
- 4) Central sites precede cortical.
- 5) Occipital poles precede frontal.

It follows from this scheme that sites that myelinate slowly during infancy may be particularly vulnerable to prolonged phe exposure. From these premises, Bick *et al.* (1991) speculated that in untreated PKU, the brain stem area may be spared as it myelinates before birth, but diffuse cortical and pyramidal damage may occur. In early-treated PKU, dysmyelination should be minimal. Nevertheless, myelination proceeds hierarchically from areas subserving the lowest (vegetative) functions to the highest (symbolic) functions (Yakovlev and Lecours, 1967) and loss of phe control, even after early initiation of the restricted diet, could carry the risk of subtle disorders of frontal or prefrontal function. This model remains to be validated empirically, but a fair generalization is that infancy is a critical time for myelin formation over wide areas of the CNS and that the system is especially vulnerable to phe during this period.

Even though myelination in humans continues till 20 years of age, major deposition is complete by two (Davison, 1973). Results from post-mortem histological analyses of brain tissue in untreated PKU led Davison to conclude that, in PKU, infancy was the stage of maximal vulnerability and in non-PKU,

the pre-natal stage. Thus, nutritional hyperphenylalaninaemia should affect CNS growth most adversely in early life, producing lesser disruption as maturation proceeds and perhaps reaching a point where none occurs that is of any functional significance. In defence of this sensitive period hypothesis, Davidson drew attention to the devastating effects on the foetal CNS of sustained high concentrations of phe in the PKU mother even though the foetal liver is not deficient in PAH. Vorhees *et al.* (1981) warned that the myelin malformation theory in human PKU is based on correlative, rather than dynamic, evidence and this remains the position today.

A major unanswered question is, once the CNS in PKU has successfully myelinated as a result of phe restriction during the vulnerable early period and beyond, can the mature system withstand subsequent phe exposure without suffering structural damage or loss of function; and, if so, at what age does this protection emerge?

Myelin defects impair *intra*-neuronal impulse conduction. However, an allied neurochemical question centres on *inter*-neuronal communication and the functioning of the synapse. Doubt has been raised about the integrity of the synapse in PKU because of the neurotransmitter shortage that results from the metabolic block. Moreover, poor dendritic arborization, as seen in untreated PKU, may stem from synaptic dysfunction (Davison, 1973). Specifically, the issue is whether dietary discontinuation in the mature human organism and its attendant risk of neurotransmitter deficiency manifestly affects mental and motor functioning. Knowledge of neurotransmitter biosynthesis suggests it would, but experimental evidence on the matter is only now beginning to emerge as subjects become available who have been screened, treated from

an early age, stopped diet and been exposed to high phe levels.

Guttler and Lou (1986) articulated the theoretical basis for concern about neurotransmission problems in post-treatment PKU. They contrasted what they termed the juvenile form of PKU with the adolescent form. In the untreated, juvenile form of the disease, severe mental handicap results. This, they proposed, is probably due principally to myelin malformation. In the treated form of PKU, stopping the low-phe diet in adolescence produces hyperphenylalaninaemia, but in a more mature organism. Thus, the juvenile and adolescent form of PKU are both chemical insults to the brain but at different stages of maturity. What is unknown is whether the biochemical imbalance associated with untreated PKU is harmful later in life. Put otherwise, the question is whether the mature brain can function normally despite an abnormal biochemical environment. Guttler and Lou suggested that maturity brings protection against the devastating effects of phe, but that subtle dysfunction might arise due to impaired biosynthesis of neurotransmitters such as dopamine and noradrenaline.

If myelination and associated neurohistological growth processes such as synaptic and dendritic formation are the key to understanding the mechanism of progressive brain deterioration in untreated PKU, questions arise about how best to control hyperphenylalaninaemia. For instance, is a maturational point ever reached when elevated phe levels cease to have a deleterious effect on the structural integrity or functional capacity of the CNS? Issues about the effect of phe on neurological growth and psychological development are at the core of treatment practice and policy in PKU and are outlined in the next chapter.

## **Chapter Two**

### **TREATMENT ISSUES**

There are three core issues relating to outcome in dietary treatment of PKU. The purpose of treatment outcome research is to further theoretical knowledge of the neurochemistry and neuropsychology of the disease and ultimately guide treatment policy towards the formulation of an optimum regimen. In the form of questions, the issues are:

- 1) at what age should dietary treatment be started in order to avoid early damage to the CNS?*
- 2) within what range is circulating phe innocuous during treatment?*
- 3) at what age, if any, is it safe to discontinue treatment?*

These matters have been approached almost exclusively from the standpoint of research into psychological outcome, although some additional evidence has been derived from autopsies and, within the past decade or so, brain scans of living patients.

### **Psychological outcome research**

#### *Sample definition*

In the reviews that follow and in the empirical studies described in later

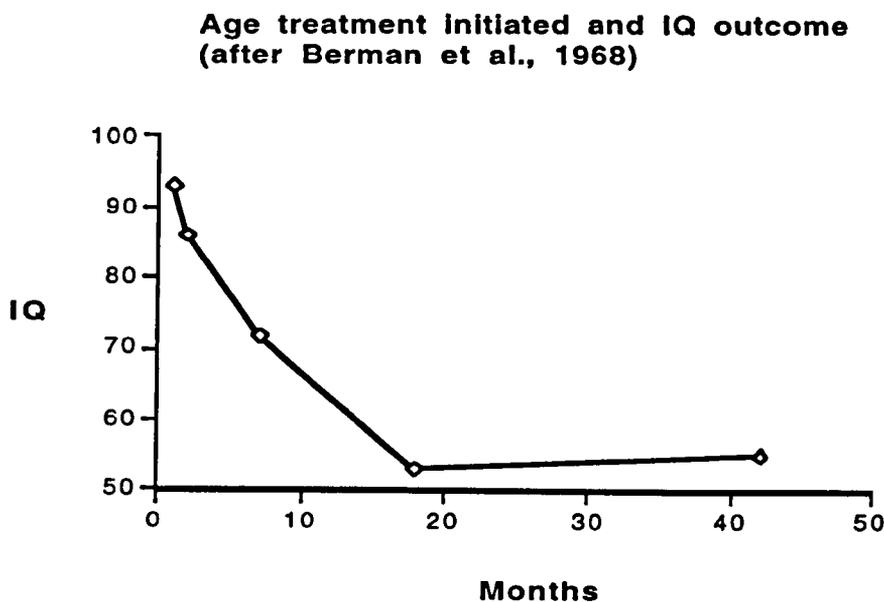
chapters, it should be noted that data are derived from subjects representing the classical PKU population unless otherwise stated. 'Classical PKU' refers to extreme enzyme deficiency in which there is little or no PAH activity. The classical subgroup is by far the largest among the hyperphenylalaninaemias, constituting over 90 per cent of the population. Even so, classical PKU is not caused by the entire PAH gene being deleted (Woo, 1989). It is operationally defined as blood phe levels persistently in excess of 1200  $\mu\text{mol/l}$  and is usually contrasted with the 'variant' subgroup, whose levels range from 600 to 1200, and the 'benign' subgroup whose levels fall below 600 (Koch & Wenz, 1987). These divisions are arbitrary. In reality, the degree of hyperphenylalaninaemia in PKU is a continuum, reflecting the broad heterogeneity of the underlying genotype. But, for convenience sake, the tripartite scheme of classical, variant and benign remains in operation. The figure of 1200  $\mu\text{mol/l}$  employed to define the threshold for classical PKU is ten times the level of blood phe found in the normal population (Walter, 1995).

#### *Age of treatment commencement*

The question about what age treatment should be initiated in classical PKU was one of the first to be addressed following Bickel's demonstration of successful dietary prophylaxis. There is now no doubt that the earlier treatment is begun the better the prognosis for sparing of intellect. One of the first studies of the relationship between treatment age and IQ concluded that treatment should be immediate (Berman *et al.*, 1968). This principle has nowadays become almost axiomatic, though in practice treatment is not effected until some days after birth as elevated phe levels, as measured by

Guthrie tests, take time to stabilize.

In a meta-analysis of patients treated at progressively later ages, Berman and her colleagues found an inverse relationship between age and IQ outcome. The figure below is adapted from their data and shows the dramatic and systematic decline in IQ to a severely mentally handicapped level following a treatment delay of one to two years. Datum points are mean IQs from samples of about 10 classical PKU subjects at each age.



**Figure 3.** Decline in IQ as a function of treatment delay.

In some cases, outcome IQ was measured over four years after initiation of treatment and no attempt was made to control for dietary compliance. Thus, starting age and compliance were probably confounded factors.

Nevertheless, assuming compliance was random, this series of patients provided strong evidence for an inverse relationship between the age treatment commenced and later IQ.

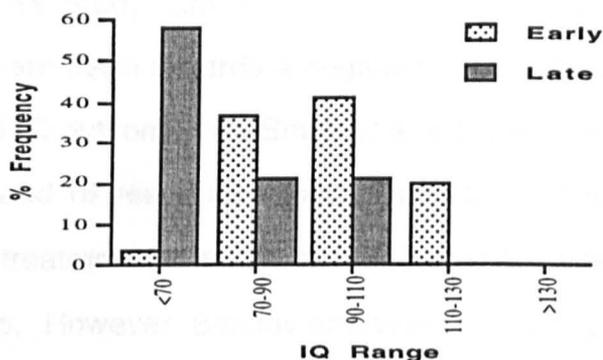
As a historical aside, some of the early-treated patients entered into Berman's analysis were among the first to be picked up by newly-established screening programs. Others, mainly those who were late-treated, were siblings of this group who were born before screening was in place.

In the late 1960s, Kaplan (1969) reviewed the available evidence for the effect of treatment onset age on IQ and concluded that if treatment was delayed until 6 to 18 months, figures in the mildly handicapped range (IQ 50 to 70) were to be expected; whereas a delay of beyond 24 months almost invariably led to severe mental handicap (IQ < 50).

Hanley *et al.* (1971), in a study of Canadian children, corroborated this view, demonstrating a steady decline in later IQ as a function of treatment delay. These authors reported IQ averages of 94, 72, and 55 for delays of 2, 6 and 12 months respectively.

In a well-constructed, early study of 58 children all selected for classical PKU, Kang *et al.* (1970) demonstrated the slippage in IQ that occurred when diet was implemented later than 8 months of age compared with implementation before 6 weeks. The percentage frequency distributions below are adapted from their data and show the abnormal intellectual levels associated with late commencement.

**IQ of Early- and Late-Treated Patients  
Proportional Comparison (after Kang  
*et al.* (1970))**



**Figure 4.** IQ distributions of early- and late-treated patients.

Bearing in mind that the population norm for the IQ distribution is 100 and the standard deviation 15, the slight positive skew in the above data for the early-treated group suggests that even 6 weeks might be sub-optimal for treatment commencement. Analyzing IQs from children in this group in terms of the 27 who began prior to 3 weeks and the 12 who began between 3 and 6 weeks, Kang *et al.* found a mean difference of 10 points, the average IQ for the pre-3 week group being 99 and that for the 3-to-6 week group being 89. Compared with their siblings, the pre-3 week group's IQs were non-significant, but the 3-to-6 week group's were significantly lower ( $P < 0.001$ ).

Apart from confirming the suspected dynamic relationship between age of treatment commencement and subsequent IQ, Kang and her colleagues' findings pointed to the need for starting treatment as early as possible to avoid loss of intellect.

Over the past quarter-century the picture has remained fundamentally unchanged. The direct link between early exposure to phe and later IQ was weakly confirmed by some of the first analyses to emerge from the UK Collaborative Study (Smith *et al.*, 1973), these researchers finding a nonsignificant trend towards a negative correlation between the age treatment began and IQ outcome. As Smith *et al.*'s patients were tested between the ages of 6 and 16 years, compliance and age of treatment cessation were additional treatment variables, but their possible effects were not analysed by the authors. However, Barclay and Walton (1988) provided controls for these factors by testing a well-controlled, on-diet group and found a significant inverse relationship between age of treatment onset and IQ.

Bickel's own results were reviewed by Schmid-Ruter and Grubel-Kaiser in 1977 (see Bickel, 1980). This survey was limited to patients whose treatment was started in the first year of life. Normal intellectual development was achieved only in those who began treatment in the first two months, IQ declining progressively thereafter.

Data from the larger sample of 134 patients available to the US Collaborative Study showed a dramatic fall-off in IQ following early treatment delay. The mean for patients who started in the first month of life was 95 whereas that for patients who started in the second was only 85 (Koch & Wenz, 1987). Stepwise regression analysis of a subset of the US Collaborative Study sample established that age of treatment onset was a significant predictor of later IQ, though age of treatment cessation and parental IQ were stronger (Legido *et al.*, 1993).

There is no evidence that the damaging effects of early exposure to phe can be reversed by later treatment. Neither late-treated children nor adults show any signs of intellectual normalization. It should be remembered that Bickel's first therapeutic trial was on a two-year-old who would nowadays be regarded as late-treated. While her demeanour, attention and social behaviour improved, her intellectual status remained severely mentally retarded. He wrote that older children and adults would probably not benefit from treatment because of pre-established brain damage and presumed that 'the best results of dietetic treatment of phenylketonuria will be obtained if treatment is started in infancy and particularly in the neonatal period' (Bickel *et al.*, 1954). While this view has been vindicated by subsequent research, today there is mounting interest in the use of the low-phe diet in previously untreated patients, not because of any prospect of fundamentally changing intellect, but because mood and behaviour may be improved. Studies of this issue published so far have been equivocal. The majority are frankly contradictory (see Harvey & Kirk, 1995) and interpretations of apparent improvements have tended to ignore non-phe factors such as Hawthorne effects.

Treatment policy as proposed by Bickel in the fifties has never been fundamentally challenged, subsequent treatment onset findings serving largely to substantiate his original presumption. Research in this area has been formally translated into treatment policy in the UK and Continental Europe, though curiously no treatment guidelines have been published for the USA (Schuett, 1996). The UK recommendations are for commencement of treatment before 20 days if phe levels are persistently above 600  $\mu\text{mol/l}$  (Medical Research Council, 1993a). In Europe, treatment is recommended before the third week of life when phe levels are in excess of 400  $\mu\text{mol/l}$

(Schuett, 1996). In practice, Guthrie tests are usually taken close to the sixth day of life and treatment often begun the week after (Smith & Brenton, 1995).

### *Dietary control and optimal therapeutic range*

The setting of a lower limit to the therapeutic range for phe in PKU is dictated primarily by dietary needs and thus recommendations have been largely based on theoretical considerations. By contrast, recommendations for the value of an upper limit are, at present, entirely empirically based and have shown some fluidity over the period since the inception of screening and treatment programmes.

The objective in setting lower and upper limits for phe during treatment is to minimize adverse long-term neuropsychological effects without compromising nutrition. In the early days of treatment, some children were severely deprived of phe in infancy. Their intake was only around 25 mg per kg of bodyweight daily for many months and they stopped growing, became dystrophic and anorexic, and developed bone lesions (Feinberg & Fisch, 1962). This was before it was realised that young infants with PKU require around 50 mg of phe per kg daily to prevent protein catabolism and aminoaciduria due to essential amino acid deficiency. Warnings about overtreatment were issued (e.g. Woolf, 1967) and consensus was reached by clinical practitioners on 120  $\mu\text{mol/l}$  as a minimum level for blood phe control during treatment. This is the figure formally recommended in current UK and continental European guidelines as the lower target limit for treatment irrespective of age (Medical Research Council, 1993a; Schuett, 1996). However, Bickel (1980) took the

view that 60  $\mu\text{mol/l}$  was tolerable and treated his patients in Heidelberg accordingly without apparent harm.

In dietary practice, the small amount of phe necessary for growth and normal protein metabolism is obtained from natural food. A daily dose of phe is calculated and carefully measured amounts of protein having known phe concentrations are permitted. These are usually low-protein foods such as cereals, potatoes or milk - including breast milk and baby milk in infancy. The dose prescribed rises slightly with age and is expressed in units of equivalence known as 'exchanges'. Each exchange contains 15 mg of phe. Thus, one exchange is equal to 15 g of rice cereal, 60 g of boiled potatoes, or 100 g of a proprietary apple yoghurt dessert (National Society for Phenylketonuria, 1996).

The exact upper limit of phe to be aimed at during treatment remains controversial and varies between clinics and countries (Editorial, 1991). The decision about what constitutes an optimum level has been informed by research studies, usually entailing outcome comparisons between naturally arising groups of well- and poorly-controlled patients or correlations between dietary control and intellectual outcome. In relation to the magnitude of the issue, such studies are surprisingly few in number.

Reporting findings from the USA Collaborative Study, Dobson *et al.*, (1977) were among the first to demonstrate the influence of dietary control during treatment on subsequent IQ. These authors divided a sample of 111 four-year-olds with classical PKU into nearly equal groups, one representing good control over the period and the other poor. Degree of control was indexed by

calculating individual children's median phe levels for the entire period and group allocation determined by whether this figure fell into the upper or lower half of the distribution. Children with indices above 600  $\mu\text{mol/l}$  were excluded from the analysis. The cut-off point (median of medians) that emerged was 350  $\mu\text{mol/l}$ . The mean Stanford-Binet IQ for the well-controlled group (medians <350  $\mu\text{mol/l}$ ) was 97, that for the poorly-controlled group (medians 350-600  $\mu\text{mol/l}$ ) was 89 ( $P < 0.01$ ). Treatment cessation was not a factor as all subjects were on-diet but age of treatment onset was not partialled out. Assuming this was random, the IQ difference suggested that quality of dietary control in the first four years of life was a factor determining intellectual outcome and that average phe levels in the range 350 to 600  $\mu\text{mol/l}$  might be harmful.

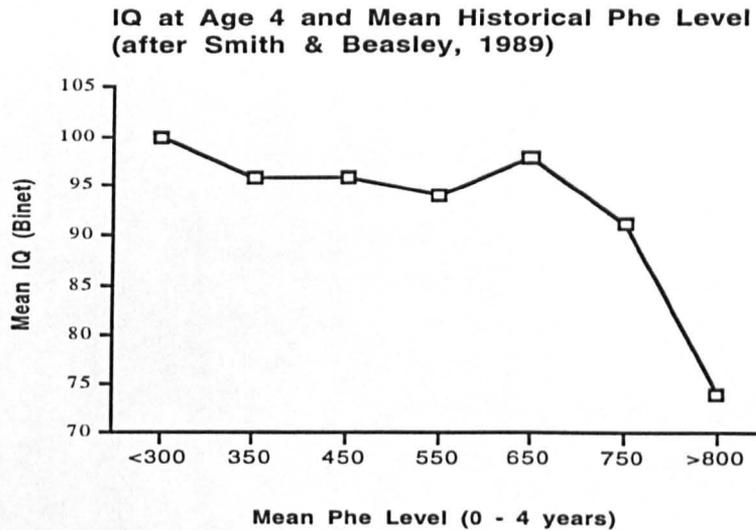
In a multiple regression analysis of an enlarged sample ( $n = 132$ ) from the USA Collaborative Study, Williamson *et al.*, (1981) ascertained the relative weightings of maternal IQ, age of treatment onset and dietary control in the determination of IQ at age 6. They found, not surprisingly, that maternal IQ contributed the highest, accounting for 25% of the variance; whereas age of treatment onset accounted for 7% and dietary control only 4%. These authors nonetheless concluded that blood phe levels during the early treatment period affected outcome and should be maintained between 120 and 600  $\mu\text{mol/l}$ , despite the earlier finding by others within the same research team that the latter figure may be too high.

Berry *et al.* (1979) provided indirect evidence for the hypothesis that intellectual deficit is related - at least in part - to degree of dietary control, obtaining a correlation of -0.68 ( $P < 0.01$ ) between concurrent phe level and

IQ (Binet) at age 4. No analysis based on historical phe data was performed, the assumption here being that concurrent phe reflects (*i.e.* correlates with) earlier phe levels.

In a study of groups from the Heidelberg clinic classified as having low or high phe levels during the first 10 years of treatment, Schmidt *et al.* (1987) found a mean difference of 18 IQ points in favour of the low-phe group, the average phe-levels being 276 and 546  $\mu\text{mol/l}$  for the low and high groups respectively. Schmidt and her colleagues interpreted the finding as supporting the phe-control hypothesis, but failed to test the data statistically. The indication is that when control is governed by an upper limit of, say, 300  $\mu\text{mol/l}$ , a higher level of intellectual functioning results. However, the possibility must be acknowledged that dietary control and parental IQ may be confounded. If parents with lower IQ were those whose treatment control was worse, the lower IQ of the children in the poorly controlled group could simply have been a reflection of that of their parents. It could be that parental IQ and dietary compliance are associated.

PKU research is riddled with methodological and interpretive difficulties, many arising out of lack of statistical analysis or experimental control. For instance, in reviewing the data sent to the UK PKU Register in the seventies and eighties, Smith & Beasley (1989) found IQ at age 4 to be fairly stable irrespective of phe level up to 650  $\mu\text{mol/l}$  but to decline when this level was exceeded. The graph below depicts mean IQs for groups with increasing mean phe levels for the four year period.



**Figure 5.** IQ and historical phe level.

This finding contradicts that of Dobson *et al.* (1977) in that no deterioration was evident between mean phe levels of 300 to 600  $\mu\text{mol/l}$ , suggesting that an upper limit of 600 may be acceptable as a dietary control target in the pre-school years. However, as with Schmidt *et al.*'s (1987) data, Smith and Beasley provided no statistical test of the mean differences, thereby undermining confidence in the IQ pattern being reliable.

Because of the paucity of information about what might constitute an optimal therapeutic range - especially in the seventies after the introduction of mass screening - national policies have varied. Some countries have exercised more clinical caution than others. In the USA, the range widely adopted was 180 to 600  $\mu\text{mol/l}$  throughout the entire treatment period (Schuett & Brown, 1984), whereas in Germany control within the range 60 to 360 was sought from birth to 10 years of age, and 120 to 720 for 10 to 15-year-olds (Burgard,

1991). In Denmark, upper limits of 420 were thought desirable from birth to 7, 600 from 7 to 10 and 720 from 10 to 12 (Andresen, 1991).

A consensus remains to be reached. In the UK, the range in common use was 180 to 600 for pre-school children. This was superseded by a clear recommendation in the early nineties that clinicians aim to keep phe levels below 360  $\mu\text{mol/l}$  in the pre-school period, below 480 in the primary school period and below 700 in adolescence (Medical Research Council, 1993a). Germany retains the tradition, established by Bickel, of endeavouring to achieve even lower levels: less than 360  $\mu\text{mol/l}$  before age 10 and less than 600 between 10 and 15 years (Burgard, 1994).

Some of these goals may be more idealistic than realistic. Even in the best controlled children, levels inexorably rise with age. For entire populations, even when vigorously treated, loss of control occurs progressively. For instance, despite the recommendations of the German Paediatric Society, Weglage *et al.* (1993) reported mean phe levels rising from 461 to 648  $\mu\text{mol/l}$  between the ages of 5 and 10 and progressing to 812 by age 15 in their sample of 34 Munster patients.

### *Age of treatment cessation*

No single issue in PKU research has generated more controversy than the age at which dietary treatment should be stopped. As with dietary control criteria, national differences have emerged. For example, in France cessation of treatment as early as age 5 was commonplace in the seventies and early

eighties (Saudubray *et al.*, 1987). Over this period in the USA, many children discontinued at age 6 (Schuett *et al.*, 1980). In the UK, at least 50 per cent of subjects on the Register were on relaxed diets (500 to 1000 mg of phe per day) by age 8, (Smith *et al.*, 1990a) though the West of Scotland Centre treated strictly till age 10. In Germany, strict control was attempted at least for the first 7 years of life (Burgard, 1994).

Eight years after Bickel's first report, concern was expressed about the lack of knowledge surrounding relaxation of dietary restriction (Moncrieff & Wilkinson, 1961). Bickel's own approach to the question of dietary cessation was theoretical. Jervis and others had shown that untreated patients did not deteriorate in adolescence; Gruter had suggested this might be due to brain myelination being complete by puberty; and Crome had demonstrated that phe could interfere with myelination: from this information, Bickel concluded that dietary termination in adolescence might be safe (Bickel, 1980).

Empirical data at that time were only available from a handful of case studies. For example, Murphy (1963) described outcomes in two children who were treated before 2 months, discontinued at 4 years of age and followed up at 7. In the first, the on-diet IQ was 85 and the off-diet 88; in the second, the on-diet IQ was 90 and the off-diet 85. Such equivocal findings clearly emphasized the need for larger-scale research, but it was not until the late seventies when information from the various collaborative studies began to emerge.

The very earliest studies revealed more about the need for careful methodology than about the influence of treatment termination age on subsequent intellect. For instance, Brown and Warner (1976) reported an on-

diet group's IQ being eight points higher than an off-diet group's but, as the groups were unmatched for age of commencement and control, and even age of cessation was not given, no clear conclusion could be drawn about the relationship between the two variables.

It was Cabalska *et al.* (1977) who perhaps did most to raise concern about diet discontinuation in the pre-school period. Comparing IQ during dietary restriction with IQ after four years of free diet in two samples of classical PKU children drawn from the Polish Collaborative Study who, respectively, stopped at 2 years 6 months and 4 years 8 months, these authors found both overall between-group differences and within-group decline in scores at follow-up. The table below shows the pattern of mean IQs.

	Mean IQ	
	On-diet	Four years later
Group One (n=22, cessation: 4 yrs. 8 mths)	101	90
Group Two (n=10, cessation: 2 yrs. 6 mths)	91	77

**Table 2.** IQ following dietary cessation (after Cabalska *et al.*, 1977)

Superficially, the data strongly suggested that 4 or 5 is too young an age to terminate diet without risk of compromising intellectual development and this was the authors' conclusion. Unfortunately, the study contained several methodological flaws which weakened the argument. First, Group Two, as well as stopping early, were also late-treated: at 40 days as opposed to 25 in

Group One. Second, Group One contained three children whose dietary control was poor. And third, no statistical analysis of the main factors and, even more importantly, the interaction term was undertaken. Thus, besides sampling and between-group comparability problems, the reliability of the data obtained was not tested.

Smith *et al.* (1978) published the first group study on data from the UK Register pertaining to the treatment cessation issue. Only a subgroup of 21 in their overall sample was homogeneous for age of commencement and cessation, nonetheless these 21 uniformly started under 6 weeks and stopped at 8 years. On-diet, their average IQ was 96, but this fell on follow-up after two years of free diet to 91 ( $P < 0.05$ ), suggesting that exposure to excess phe during the later primary school years might still carry a risk, and that even 8 might be too young an age to discontinue.

These findings were undermined by a report from Boston by Koff *et al.* (1979). This well-controlled, longitudinal study revealed no significance difference in pre-termination and off-diet IQ in a group of 30 classical PKUs who had stopped between 4 and 6.5 years of age and who were followed up, on average, 3.5 years later. These authors saw the result as a vindication of the, then, usual US practice of discontinuing diet between 4 and 6. However, this may be too simplistic an interpretation as no attempt was made to relate IQ to the period of post-dietary exposure to phe. The case would have been weakened, had a significant negative correlation emerged from such an analysis.

By the early eighties it was becoming clear that the criteria for evaluating

methods used in treatment discontinuation research were good control of non-treatment factors such as severity of the disease and of treatment factors such as age of commencement and compliance with therapeutic limits. Also, post-cessation management can be a factor if children are placed on relaxed rather than free diets as is customary in Germany. Furthermore, length of follow-up period is an important factor as the extent of brain damage caused by a given level of phe over a given time seems to diminish with age and short-term follow-ups may yield false negatives.

Apart from this last criterion, the others were met by the initial study on diet discontinuation from the US Collaborative Study (Koch *et al.*, 1982). These authors adopted a quasi-experimental approach to the question, randomly allocating 115 early-treated children at their sixth birthday to continuing and discontinuing groups - matched for size, age, sex and severity - and retesting IQ at age 8. They found no reliable change in IQ for either group despite the discontinuers' phe level having doubled, leading them to conclude that if diet is stopped at 6 no major intellectual deterioration occurs in the short-term. Nevertheless, the authors added the caveat that the high phe group may fare worse in the long term. In a follow-up study of the groups conducted after a further two years of high phe exposure for the discontinuers, Fishler *et al.* (1987) then found a significant IQ difference in favour of the continuing group. While the discontinuers' IQs remained near-constant, the continuers' rose. Combining mean data from the two studies gave the following table. Figures at 6 years are baselines when both groups were on diet. The interaction between group and age was significant ( $P < 0.05$ ).

	6 yrs.	8 yrs.	10 yrs.
Discontinuers	97	97	96
Continuers	101	102	105

**Table 3.** Treatment discontinuation and serial IQ (after Fishler *et al.*, 1987).

Thus, the suspicion was raised that six years of age may be too young to discontinue treatment. Though there appeared to be no dramatic decline in IQ in the slightly longer term following treatment termination at this age, compared with the continuing group those who stopped seemed to fare worse in the late primary school period.

In a report published almost simultaneously with that of Fishler *et al.* (1987), considerable weight to the argument that treatment cessation at 6 is premature was added by an analysis of another subset of the US Collaborative Study's cohort (Koch *et al.*, 1987). In this work, IQs were obtained at the later age of 12 for two groups, one of which stopped diet at 6 and one that stopped at 8. The outcome here was strikingly different to that found by Fishler *et al.* in that the mean IQ for the earlier-stopping group was well below that of the later-stopping group: 87 compared to 101. The significance of this apparently large difference is unknown as no statistical test of the distributions was provided.

Earlier findings reported by Seashore *et al.* (1985) from a small-sample study (n=14) closely mirrored the above 14-point discrepancy and cast more doubt on the wisdom of terminating treatment as early as 6 years old. These authors found mean IQs fell from 104 at age 6 when the subjects came off diet to 90 at age 13, this being a significant difference ( $P < 0.05$ ).

By the mid-1980s, the tentative conclusion from the US Collaborative Study was that treatment should continue at least until age 8 and possibly beyond (Holtzman *et al.*, 1986). As children detected in the early days of mass newborn screening grew older and increased in number, empirical research improved due to greater sample selectivity and investigation of longer dietary - and, indeed, phe exposure periods - becoming possible. Hence, by the 1990s workers from the US Collaborative Study were able to compile a sample of 95 treated subjects with classical PKU, all of whom commenced early (average = 17 days), were reasonably well-controlled and lost dietary control at different ages but with some continuing beyond 10. Because of these characteristics, Azen *et al.* (1991) were able to compare IQ results from three modestly sized subgroups that varied on the termination parameter, one stopping before 6 years of age (n=23), one before 10 (n=47) and one carrying on beyond 10 (n=25). She and her colleagues found a progressive and significant deterioration in mean WISC-R IQs at age 12 as a function of age of dietary cessation, the mean for the early group being 87, that for the middle group 96 and that for the late group 101 ( $P < 0.01$ ). These researchers concluded that the PKU diet should be maintained until at least 12 years of age, however they failed to take into account the fact that quality of control and age of cessation were confounded. Thus, even before age 4 when all the subjects were nominally on diet, the early cessation group had the worst

control and the late cessation group the best, mean phe levels for the period from diagnosis being respectively 576 and 354  $\mu\text{mol/l}$ . The middle group had an intermediate level of 462.

Because of the often retrospective nature of treatment cessation research, other treatment factors cannot be controlled experimentally. An alternative approach to comparing IQ outcome in on- and off-diet groups, is multiple regression analysis. The theoretical basis for this methodology is the now accepted view that outcome in treated PKU does not depend exclusively on the single treatment factor of age of cessation but on an amalgam of constitutional and treatment variables related to phe. The causal model can be expressed as an equation as follows.

**Severity + age begun + quality of control + age stopped = mental level**

The constitutional variable of *severity* refers to genotypic expression and the level of phenotypic PAH activity associated with different genetic mutations. The other three variables concern how treatment is managed clinically and domestically. While this simple predictive model is generally accepted by practitioners and scientists, what genetic and biochemical research has endeavoured to elucidate is the weightings of each variable and how they interact. This quest is ongoing and the multiple regression technique lends itself to the holistic approach required once a model such as the above, with its multiplicity of interactive parameters, is acknowledged. However, few studies to date have attempted to incorporate data into a multivariate scheme.

One example is a follow-up study by Chang *et al.* (1983) of 49 patients, 33 of whom were off diet when tested. The aim of their study was to identify and weight factors affecting IQ outcome by means of multiple linear regression. Data on both phe and non-phe factors were fed into the analysis. They found that parental education, commencement of treatment, on-diet phe level and length of treatment all predicted IQ. Leaving aside parental education, which is a non-PKU-specific predictor, higher IQ was most associated with early treatment and, to a lesser extent, low on-diet phe levels and prolonged treatment. However, the design and sample characteristics of the study precluded these authors from making a statement about the best age for treatment cessation.

In another study based on a multiple regression analysis, Michals *et al.* (1988) corroborated Chang and his colleagues' findings, obtaining IQ results from a sample stratified by age of cessation. These workers found that parental IQ and age of treatment commencement were, respectively, directly and inversely proportional to later IQ and, furthermore, that blood phe level after treatment discontinuation was a predictor. Curiously, they recommended that diet be continued beyond 10 years. However, this conclusion was not necessarily warranted by their findings. None in their sample was older than 10, thus no off-diet data after this age were available.

### **Treatment control and discontinuation policy**

Research has largely settled the issue of treatment commencement, overwhelming evidence supporting an early intervention policy, ideally with

treatment beginning as soon as is practicable. In age terms this is the second week of life as diagnosis before then is unreliable due to transient hyperphenylalanaemia. During treatment, a phe-level ceiling as high as 600  $\mu\text{mol/l}$ , such as used to be commonplace in the US, may be too high and nowadays in the UK and Germany a formal policy has been adopted of restricting phe, at least in the pre-school years to within a 360  $\mu\text{mol/l}$  limit.

Treatment discontinuation policy remains to be better informed by research. Some practitioners have taken the extreme view that treatment should be lifelong (*e.g.* Naughten *et al.*, 1987) and this position has been more widely advocated in recent years. However a 'diet-for-life' treatment regimen may be personally and socially disruptive for the patient and some concern has been already raised that the policy may be precipitate, overstated and unjustified by empirical findings (Walter, 1995). The logic of extrapolating from research that suggests that diet should be maintained until, say, age 10 to a lifelong treatment regime does not stand scrutiny, yet this line of reasoning has been followed by several authors in the nineties.

For instance, Smith *et al.* (1990a) analyzed longitudinal IQ data from two large cohorts within the UK PKU Register - the first, children born between 1964 and 1971 ( $n=224$ ) and the second, children born between 1972 and 1978 ( $n=375$ ). To control for possible shifts in IQ population norms, she and her co-workers converted obtained IQs into Z-scores based on the estimated true mean for the population at the individual child's fourth birthday (when the first IQ measure was taken) and the standard deviation of the test. These transformed IQs were then regressed against historical phe concentrations. A significant negative correlation was found between IQ at age 8 and phe levels

in the previous 4 years. IQ at age 10, however, showed only a weak inverse association with control in the preceding two years in one cohort and none in the other, and IQ at ages 12 and 14 in relation to phe levels after age 10 was random.

Smith *et al.* concluded that by middle childhood 'maturation of the nervous system has advanced to a stage where the parameters of intelligence have stabilised and are therefore less vulnerable to the biochemical effects of phenylketonuria'. Notwithstanding this clear statement about the CNS eventually reaching a stage when it achieves protection against the malign influence of phe, and despite the lack of empirical evidence for a harmful effect of phe on IQ after age 10, Smith *et al.* went on to recommend that patients with PKU will 'have to be advised to continue a low phenylalanine diet indefinitely'. Their reasons for this apparent *non sequitur* stemmed from sporadic case reports of adults who were treated early showing abnormal MRI scans (*e.g.* Villasana *et al.*, 1989; Thompson *et al.*, 1990) but without knowing whether such findings relate to post-treatment phe exposure in the second decade.

Similarly, at the end of a review of 12 treatment discontinuation studies, Potocnik and Widhalm (1994) concluded that the evidence pointed to treatment continuing until 10 years but then, paradoxically, went on to recommend a diet for life. This has already become the stated policy of clinics throughout Eire (Naughten *et al.*, 1987) and to a lesser extent in Holland (Slijper *et al.*, 1988).

## **Aim of present research**

Treatment research and policy are inseparable in PKU. A more sober assessment might suggest that generalizations about discontinuation policy be withheld until further knowledge is obtained about long-term psychological functioning in adolescence and adulthood following treatment. PKU research continues to evolve and studies of older patients who were well treated in childhood is the next natural stage in the process. Hitherto, such investigations were not possible. Because national mass screening programmes only began in the late sixties and early treatment criteria were undecided, until now there were insufficient older individuals available for constructing samples in which severity and treatment variables were satisfactorily controlled.

Compilation of such samples is now a reality and the first aim of the research described in this volume is to shed light on the issue of whether neuropsychological functioning is impaired by excess circulating phe occasioned by dietary cessation following strict phe control up to the tenth year of life. This was treatment policy in Glasgow from the inception of mass screening and one that enabled a group to be collated that had good homogeneity for treatment variables. Evidence of post-diet dysfunction in such a group would imply ongoing CNS susceptibility to the neurotoxic effects of phe and would indeed provide support for treatment continuation at least into adolescence - though not necessarily for life. The phe-restricted diet is a harsh regime whose psychosocial cost both for child and caretaker is considerable. Research must demonstrate that the benefits of remaining, perhaps indefinitely, on a phe-restricted food intake outweigh the demands

and privations of the treatment.

## **Chapter Three**

# **INFLUENCE OF POST-TREATMENT HYPERPHENYLALANINAEMIA ON NEUROPSYCHOLOGICAL FUNCTIONING**

## **Background**

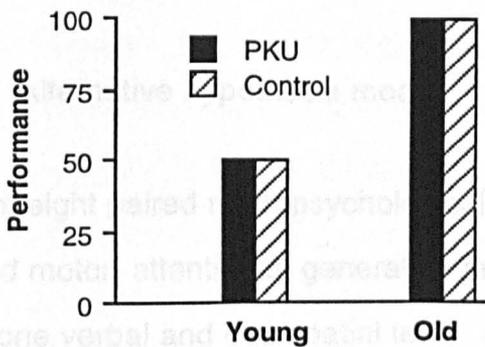
The study reported in this chapter arose out of an opportunity to examine neuropsychological functioning in a group of adolescents and young adults with PKU, all of whom had been diagnosed and treated early, had remained on a well-controlled diet until age 10, had ceased treatment abruptly at this age and had subsequently been hyperphenylalaninaemic for three years or more. These sample characteristics were determined by the policy in the West of Scotland of keeping patients, diagnosed at neonatal screening as having classical PKU, on a strict diet throughout childhood, then discontinuing treatment on their 10th birthday. Compared with, say, the early French and American approaches of terminating treatment at 6 or even earlier (Schuett & Brown, 1984; Saudubray *et al.*, 1987), the policy in Scotland was a conservative one founded on a combination of theory and caution.

In the 1970s, insufficient empirical data were available to guide clinical practice. Theoretical considerations of when might be safe to stop treatment centred on the knowledge that myelination of the human CNS is largely complete by 10 years of age. It was therefore assumed that, thereafter, the CNS would be sufficiently protected to withstand the high levels of circulating

phe that would result from dietary liberalization. The prediction stemming from this theoretical position was that post-treatment neuropsychological functioning would be unaffected by long-term exposure to phe after the age of 10. This was the hypothesis tested in the study.

A cross-sectional design was necessary to investigate outcome in patients falling into this category. A longitudinal study would have been preferable, as each subject could have acted as their own control, but this was precluded by an absence of on-diet test results. A control group procedure was chosen instead, with results from a carefully selected group of ten older off-diet PKUs being compared with those from a group of matched non-PKU controls. Further control for phe-effects was achieved by incorporating into the design a group of younger on-diet PKUs matched with a second non-PKU group. All four groups were tested using the same neuropsychological battery, the subtests of which were selected so as to be suitable for a wide age range.

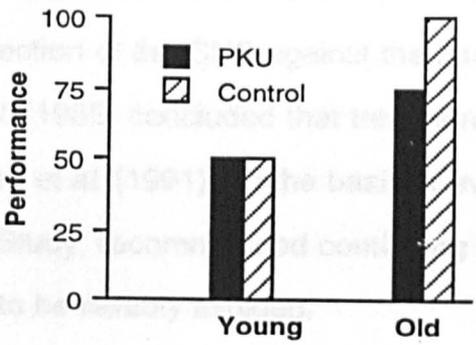
The model for the null hypothesis was as follows.



**Figure 6.** Null hypothesis model.

The figures on the ordinate for test performance are notional. The age factor

was expected to show a difference in favour of better performance by both older groups - PKU and control - on developmental grounds. The PKU factor was expected to show either no difference or perhaps a slight loss of performance for both PKU groups, the latter being a common finding even in well-treated PKU (Beasley *et al.*, 1994). The interaction was the critical term in the analysis. If post-dietary phe had had no specific effect, then the interaction between treatment status and age would be non-significant. If, however, phe had had a significant deleterious effect after an exposure period of three or more years, then the interaction term would reveal this, the score pattern for the alternative hypothesis being as follows.



**Figure 7.** Alternative hypothesis model.

A battery of eight paired neuropsychological tests was devised. The four pairs represented motor, attentional, generative and memory functions. Each pair contained one verbal and one spatial test. No between-test differences were hypothesized *ad hoc*, the principal aim of the study being to explore whether a specific phe effect would emerge that could be attributable to treatment cessation.

## Introduction

Global mental handicap in classical phenylketonuria (PKU) can be prevented by early implementation of a phenylalanine-restricted diet. However, research over the past decade or so has suggested that subtle psychological deficiencies might arise in patients with phenylketonuria after discontinuation of prophylactic treatment, despite early commencement. Problems reported include mild diminution of IQ (Koch *et al.*, 1982), scholastic underachievement (Seashore *et al.*, 1985), spatial dyspraxia (Pennington *et al.*, 1985), visuo-motor incoordination (Faust *et al.*, 1986) and atypical sensory evoked potentials (Pueschel *et al.*, 1983). In all these studies, however, patients' dietary control was lost at around 5 to 6 years of age, which may be too early to ensure protection of the CNS against the harmful effects of phenylalanine. Holtzman *et al.* (1986) concluded that treatment should continue until at least age 8 and Azen *et al.* (1991), on the basis of evidence from the USA Collaborative Study, recommended continuing dietary restriction until age 10 if loss of IQ is to be reliably avoided.

The optimum age for dietary termination in PKU remains controversial. No world-wide age standard has been established for stopping treatment, past figures varying from 4 to 14 years (Smith *et al.*, 1978; Schuett *et al.*, 1980). Since the introduction of national screening in the 1960s, the age adopted in the West of Scotland for treatment discontinuation in PKU has been 10 years. During the first 5 year period of treatment the objective has been to maintain phenylalanine levels below 400  $\mu\text{mol/l}$  and during the second 5 years below 600  $\mu\text{mol/l}$ . Because of this policy, the rare opportunity arose of investigating the psychological status of a selected sample of PKU patients who were well

treated until age 10 but who then discontinued diet and subsequently experienced several years of high phenylalanine intake as adolescents or adults. The present study explored the question of whether 10 years of treatment during the most vulnerable neuromaturational period is sufficient to provide long-term immunity against the mentally disabling effects of phenylalanine. What distinguishes it from previous research is not only the prolonged duration of treatment, but also the subjects' length of exposure to high levels of phenylalanine thereafter: three years in all cases and as much as 11 in some, these latter being among the first babies detected following the introduction of neonatal screening in the UK (Wolff, 1981).

Psychological investigation focussed on motor, perceptual and cognitive functions. Measures derived from neuropsychology were used (Griffiths, 1989), the rationale being that information-processing, multi-trial learning, sustained attention and decision-making tasks consistently show greater sensitivity to non-focal organic impairment than IQ tests with their high proportion of long-term memory items (Barbizet, 1970; Miller, 1977).

Test performance of off-diet PKU subjects was compared with that of unrelated, developmentally normal adolescents and adults. As a control for dietary effects, a group of younger, on-diet PKU subjects was also tested and their results compared with normals. We hypothesized that, if exposure to phenylalanine is psychologically harmful even after 10 years of treatment, the older, off-diet group would show adverse effects (*i.e.* diminished performance) on neuropsychological tests; by contrast, the younger, on-diet group who had not been exposed to chronically high levels of blood phenylalanine would not show the same effects. On the basis of the well-established finding that minor

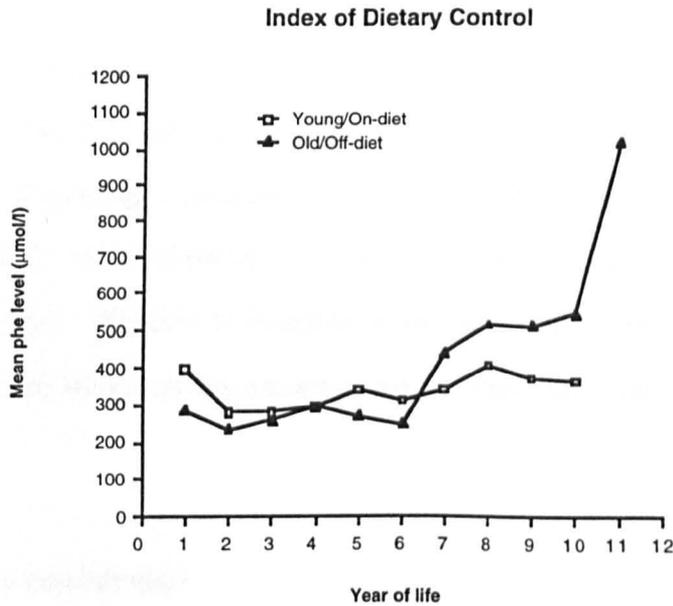
cognitive deficiencies are present even in treated PKU patients who remain on diet (Berry *et al.*, 1979; Brunner *et al.*, 1983; Beasley *et al.*, 1994), we expected some diminution in performance relative to normal by the PKU sample overall, but that the decrement would be greater in the off-diet group if post-treatment hyperphenylalaninaemia were a deleterious factor.

## Methods

### *Subjects*

Forty subjects in total were tested, 20 PKUs and 20 controls. The PKU sample consisted of 10 younger children (seven boys, three girls) who were still on diet and 10 older adolescents and adults (same sex ratio) who had stopped. All 20 had been diagnosed as having classical PKU at birth, the mean neonatal blood phenylalanine level before treatment for the younger group being 1977  $\mu\text{mol/l}$  (s.d. 500) and for the older group 1912  $\mu\text{mol/l}$  (s.d. 1077). The groups did not differ statistically in severity ( $t = 0.17$ ,  $P = \text{N.S.}$ ). All were diagnosed and treated within three months of birth, the median age at which treatment commenced for the younger group being 15 days (range 12-20) and for the older group 31 days (range 8-75). Again, this difference was not significant ( $U = 25$ ,  $P = \text{N.S.}$ ). On average, both PKU groups maintained their blood phenylalanine levels below 600  $\mu\text{mol/l}$  throughout treatment and at least for the first six years of life stayed within what is nowadays recommended as the therapeutic range of 120 to 360  $\mu\text{mol/l}$  (Medical Research Council, 1993a). The older group's level rose steeply on discontinuation of the restricted diet at age 10. Figure 8 shows changes over

time for the two groups' indices of dietary control.



**Figure 8** Index of dietary control: average phenylalanine blood levels for both PKU groups during each year of life. Means are derived from annual median values for each subject. The restricted diet was formally discontinued in the older group at age 10. The figure shows a sharp rise in the average level for the subsequent year.

At testing, the median age of the younger PKU group was 7.53 years (range 5.58-9.75), that of the older group 20.59 years (range 13.58-28.42). The median age for dietary discontinuation in the older patients was 10.17 years (range 9.50-11.42) and all had been exposed to high levels of phenylalanine for at least three years, the median period being 10.50 years (range 3.58-17.00). The mean phenylalanine level of the younger group when tested was 348 µmol/l (s.d. 167), that of the older group 1014 µmol/l (s.d. 216). This difference was highly significant ( $t = 7.72, P < 0.001$ ). Overall, the PKU

sample had a working-class bias.

Each PKU subject was matched with a non-clinical control on the basis of age, sex and approximate socio-economic status. The average ages of the younger and older control groups did not differ significantly from those of the PKU samples (younger controls: median 7.59 years, range 5.66-9.75,  $U = 49.00$ ,  $P = \text{N.S.}$ ; older controls: median 20.54 years, range 13.58-27.92,  $U = 47.50$ ,  $P = \text{N.S.}$ ). All control subjects were free from major developmental disorders as far as could be ascertained by parental- and self-report.

### *Tests and administration*

Eight tests were selected on the basis of Pennington's (1991) developmental neuropsychological model, each providing respectively a measure of motor speed, sensory-motor coordination, visual perception, visual attention, verbal fluency, visual fluency, list learning and position learning. In addition to their sensitivity to organicity, a major factor in the choice was their applicability to a wide age range.

Test details are as follows. 1) *Simple Reaction Time* Button-pressing in response to single light stimulus. Preferred hand. Ten trials summated. 2) *Peg Transfer* Purdue Pegboard (Lafayette Instruments) without collars and washers. Total time for transferring 25 pegs from left to right column of holes and back again. Preferred hand. 3) *Matching Figures* Visual matching-to-sample of complex drawings. First ten items from MFFT20 (Cairns and Cammock, 1978). One-minute time-limit per figure. 4) *Letter Cancellation*

Visual letter-search (Neisser, 1967). Twenty target stimuli, 1180 distractors. Four minute time limit. 5) *Verbal Fluency* Thurstone's cued lexical retrieval task. Number of words beginning with letters 'F', 'A' and 'S' generated in one minute respectively (Spreen & Benton, 1969). 6) *Design Fluency* Spatial counterpart of Verbal Fluency. Number of non-repeated designs made from two straight and two curved pieces of flat plastic in three minutes (Griffiths, 1991). 7) *Rey Verbal Learning* Recall of orally-presented 15-word list, repeated five times (Rey, 1941). 8) *Rey-Davis Manual Labyrinth* Spatial counterpart of Verbal Learning. Systematic search and recall, repeated five times, of single peg positions in each of four stacked wooden boards, each board containing a further eight distractor pegs (Zangwill, 1946).

### *General procedure and design*

Subjects were tested individually and given standardised instructions. To minimise order effects, the eight tests were randomised and the first-to-last sequence counterbalanced. Each measure yielded a single, objective performance score. The study conformed to a cross-sectional mixed design with matched pairs, independent samples and non-repeated measures.

## **Results**

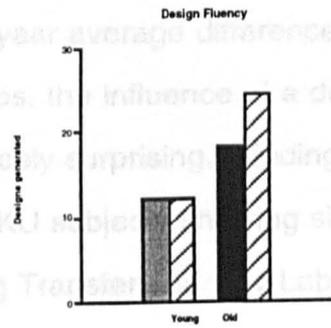
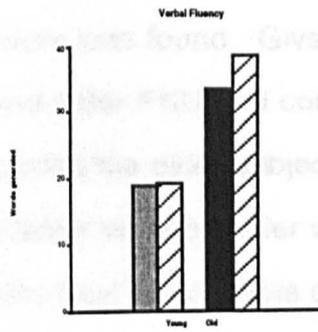
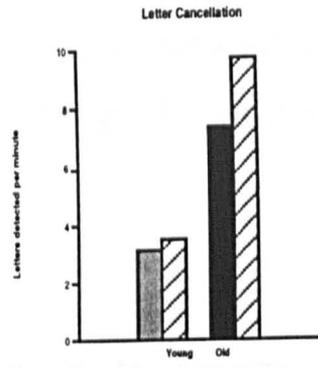
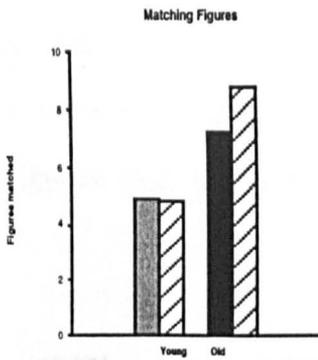
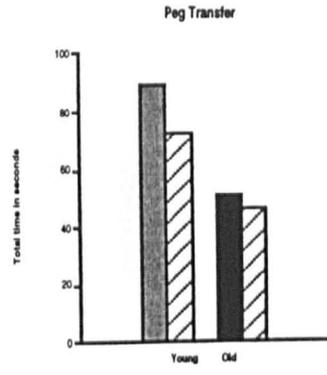
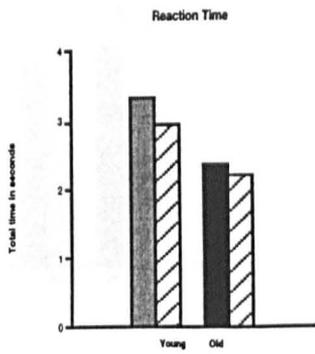
For each subject, scores across items or trials on each neuropsychological test were summated. Table 4 shows the mean scores and standard deviations on the eight tests obtained by the four groups. Figure 9 depicts mean data only. Reaction Time and Peg Transfer values are times in seconds: the lower

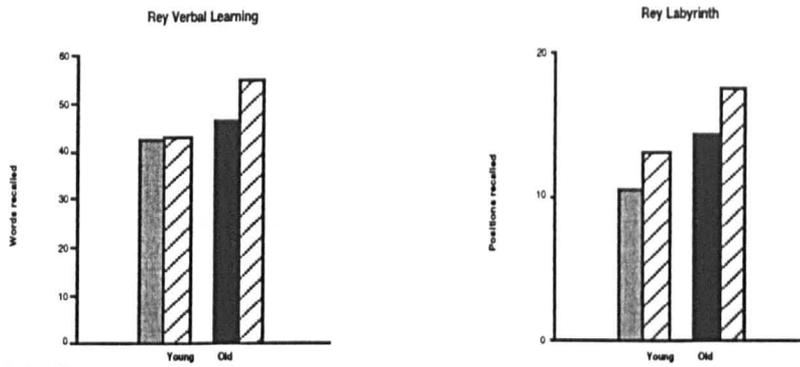
the value the better was the performance. The remaining values are number-correct scores. No basal or ceiling problems were encountered on any test.

Tests	Young				Old			
	PKU (ondiet)		Control		PKU (off-diet)		Control	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Reaction Time (secs)	3.31	0.89	2.95	0.52	2.37	0.77	2.20	0.47
Peg Transfer (secs)	88.50	19.56	72.70	12.80	51.50	5.42	46.90	4.75
Matching Figures	4.90	1.52	4.80	1.55	7.20	1.40	8.80	1.23
Letter Cancellation	3.33	0.77	3.51	1.46	7.36	3.52	9.76	5.06
Verbal Fluency	18.80	8.95	19.00	9.63	33.80	9.43	38.90	8.63
Design Fluency	11.90	2.56	12.20	2.35	18.20	3.43	24.50	3.41
Rey Verbal Learning	42.40	10.09	43.00	7.65	46.70	8.08	55.10	7.40
Rey Labyrinth	10.50	4.74	13.10	4.98	14.50	3.95	17.60	1.08

**Table 4.** Means and standard deviations for neuropsychological tests. Data are performance scores expressed as total correct items except for Reaction Time and Peg Transfer tasks which are scored in seconds. See also Figure 9.

Data from the test battery were analysed by ANOVA. The focal point was the interaction term as it provided an indication of differential performance between the on- and off-diet groups and therefore whether post-treatment hyperphenylalaninaemia impaired psychological function.





**Figure 9.** Group mean scores on neuropsychological tests. Bars: (grey) younger, on-diet PKUs; (black) older, off-diet PKUs; (hatched) controls. All interactions were non-significant except for Design Fluency ( $P < 0.01$ ) and Matching Figures (n.s. trend,  $P = 0.07$ ).

Design Fluency

Rey Verbal Learning

Rey Labyrinth

Separate ANOVAs were conducted on each set of test results. F-ratios and significance levels are displayed in Table 5. A highly significant age effect on each measure was found. Given the 13-year average difference between the younger and older PKU and control groups, the influence of a developmental factor favouring the older subjects is scarcely surprising. Findings for the diagnosis factor were patchier with the PKU subjects showing significantly poorer ability than the controls on the Peg Transfer and Rey Labyrinth tests irrespective of age or treatment status. Design Fluency also showed a PKU-effect. Furthermore, Design Fluency was the only test to reveal a significant interaction between diagnostic and treatment status, the older, off-diet PKU group doing differentially much worse than the younger on-diet group (Scheffe,  $P < 0.001$ ). Apart from Design Fluency, none of the other interaction terms was significant at the 5 per cent level, although that for Matching Figures showed a trend towards significance at the 7 per cent level, again with the off-

diet group specifically performing worse than the controls. There was no significant PKU-effect on this variable.

Tests	Factor					
	Age/diet		Diagnosis		Interaction	
	F	P	F	P	F	P
Reaction Time	15.20	0.001	1.50	N.S.	0.19	N.S.
Peg Transfer	63.35	0.001	7.22	0.05	2.28	N.S.
Matching Figures	48.47	0.001	2.75	N.S.	3.53	N.S.T.
Letter Cancellation	26.50	0.001	1.79	N.S.	1.05	N.S.
Verbal Fluency	36.23	0.001	0.84	N.S.	0.71	N.S.
Design Fluency	97.73	0.001	12.31	0.01	10.17	0.01
Rey Verbal Learning	9.59	0.01	2.89	N.S.	2.17	N.S.
Rey Labyrinth	11.28	0.01	5.07	0.05	0.04	N.S.

**Table 5.** F-ratios from two-factor ANOVA and significance levels at  $P \leq 0.05$ . 'Age/Diet' factor refers to younger, on-diet PKU children and controls versus older, off-diet PKU adolescents/adults and controls. 'Diagnosis' factor refers to all PKU subjects, *i.e* both on- and off-diet, versus all controls. Degrees of freedom 1/36 throughout. Only Design Fluency interaction term achieved significance.

Matching Figures interaction showed a non-significant trend at  $P=0.07$ . Peg Transfer, Design Fluency and Rey Labyrinth showed an overall PKU effect (see also Figure 9).

Correlating raw test scores derived from the older PKU group ( $n = 10$ ) with the

lengths of time these subjects had been off diet yielded no significant statistics. Six of the eight coefficients were negative, but the strengths of the associations were insufficient to indicate any relationship between time off diet and poor test performance.

As a final analysis, test scores for the off-diet group were standardised by converting each into a Z-score based on the mean and standard deviation derived from the older control group's distributions on each of the tests. With signs reversed for Reaction Time and Peg Transfer, individual PKU subjects' Z-scores were summated across the eight tests to provide a composite score for the battery as a whole. Composite scores were then correlated with time off diet (*i.e* period of hyperphenylalaninaemia), but the coefficient was not reliable ( $r = -0.52$ , d.f. 8,  $P = \text{N.S.}$ ). While the negative sign implies that overall test performance declined with time off diet, the association was weak and failed to represent even a non-significant trend at  $P < 0.10$ . No correlation was found between composite scores and post-treatment phenylalanine level ( $r = -0.04$ , d.f. 8,  $P = \text{N.S.}$ ), nor with age of treatment cessation ( $r = 0.03$ , d.f. 8,  $P = \text{N.S.}$ ), although the range of the latter variable was very narrow. None of the treatment-related variables correlated reliably with any of the tests singly.

## **Discussion**

Overall, the results from the present study failed to provide convincing evidence that prolonged exposure to unrestricted phenylalanine during adolescence and early adulthood is harmful to cognitive and motor functioning in well-treated PKU when dietary termination is delayed until at

least 10 years of age. Methodological factors such as diagnostic and treatment variables were tightly controlled to produce homogeneous experimental samples. Nevertheless, as group sizes were small and functions tested limited, inferences about the wider PKU population should be made with caution. The findings from the present study, however, appear to vindicate the West of Scotland practice of continuing dietary treatment until age 10, but do not necessarily support a diet-for-life policy as advocated by some authors (*e.g.* Beasley *et al.*, 1994). Further research, ideally multi-centre, with larger numbers and more varied tasks is required to settle this issue.

The finding that both the on- and off-diet PKU groups under-performed the controls on the Peg Transfer, Design Fluency and Rey Labyrinth tasks was broadly in keeping with previous work. Although marginally lowered IQ in early-treated PKU, despite good dietary compliance, has been reported by several authors (*e.g.* Berry *et al.*, 1979; Brunner *et al.*, 1983; Smith & Beasley, 1989), it may be that the discrete functions tapped by these neuropsychological measures (rapid eye-hand coordination, visual creativity and learning of spatial sequences) are especially vulnerable to prolonged exposure to the mildly elevated phenylalanine blood levels found even in well-treated patients.

Welsh *et al.* (1990) found specific impairments on neuropsychological tests tapping executive motor and cognitive functions in pre- and early-school treated PKU children. She and her colleagues hypothesized that even well-treated PKU children are susceptible to mildly elevated levels of phenylalanine early on in life and, in particular, suffer from selective deficits in executive

functions such as attentional control, maintenance of mental set and motor planning. These authors proposed a prefrontal dopamine depletion mechanism to explain the specificity of the functional data. Diamond *et al.* (unpublished) have pursued this line of enquiry, arguing that the conventional therapeutic range of 100 to 600  $\mu\text{mol/l}$  is too liberal, given that the upper phenylalanine limit is five times the norm. On the basis of comparative and neuropsychological evidence, she and her colleagues have concluded that elevated circulating phenylalanine as customarily found in well-treated PKU has a specific negative effect on cognitive-executive functions mediated by the prefrontal cortex of the brain even as early as infancy. In view of this, it is noteworthy that the three tests performed badly by both the on- and off-diet PKU groups fall broadly into the category of prefrontally-mediated executive functions, being characterised respectively by behavioural regulation, ideational productivity and spatial planning. While these findings may lend support to the idea of a specific executive-function defect in treated PKU which is identifiable early in childhood and persists into adulthood, the results nevertheless contained exceptions to this scheme. Scores on Verbal Fluency and Letter Cancellation were inconsistent, for example, even though these tasks are generally considered to represent executive activities dependent on prefrontal integrity (Beaumont, 1983).

What is perhaps of greater concern is that deficits of any kind were found in the PKU patients overall. Both the on- and off-diet groups' average indices of dietary control in the early treatment period were within or close to the lower, revised therapeutic range as currently recommended of 120 to 360  $\mu\text{mol/l}$  (see Figure 8) and not the more liberal range of 180 to 600  $\mu\text{mol/l}$  that was actually in operation at the time (Medical Research Council, 1993a). Despite

families' above-standard compliance with treatment early in the life of the children, three of the eight neuropsychological tests showed significant abnormalities of function later, two of these being unrelated to subsequent exposure to unrestricted phenylalanine.

The Design Fluency task showed not just a general PKU effect but an off-diet effect as well. Impaired design fluency has been found in patients with focal frontal lesions (Jones-Gotman & Milner, 1977), adding weight to Welsh *et al.*'s view that the prefrontal cerebral cortex and its functions may be particularly vulnerable to the harmful effects of phenylalanine. Furthermore, this finding extends the prefrontal theory by implicating the adolescent and early adult period - when most patients have either terminated or relaxed the diet - as a time of possible additional risk. Whether deficient skill in the invention of visual designs after treatment for PKU is a temporary toxicological phenomenon and therefore perhaps reversible by dietary reinstatement remains to be investigated.

It should be emphasized that the Design Fluency result was an isolated one, necessitating replication, and may represent at best a chance phenomenon or at worst a minor dysfunction. Further research is required before any firm conclusions can be drawn about prolongation of dietary treatment into adolescence and beyond.

## **Commentary**

Neuropsychological measures were chosen for testing phe effects on the

grounds that these are more likely to be sensitive to subtle cognitive dysfunction in PKU than IQ tests (Burgard, 1996). Nonetheless, the failure to incorporate IQ measures in the study could be regarded as a shortcoming as the presence of such data, in addition to shedding light on the concurrent validity of the experimental measures, would have made the results more readily comparable with those in the bulk of the literature.

The non-significant interactions found on the Rey Verbal and Rey Labyrinth tests suggested no derangement of verbal or spatial memory as a consequence of post-treatment exposure to phe. However, this result conflicts with one recently obtained by Demellweek (1996) who found a significant discrepancy ( $P < 0.05$ ) between the mean Full-Scale IQ on the WAIS-R (101) and the mean General Memory Index on the WMS-R (82) in six adolescents and adults with treated PKU. Though fewer, the individuals in his sample were all early-treated, but had been off diet for slightly longer: 5 to 13 years. The degree of comparability between the two samples as regards severity and phe-control is unclear as concurrent and historical phe levels were not reported. Demellweek's finding is striking but, to date, an isolated one that remains to be substantiated.

Non-homogeneity of the treatment commencement variable was a minor methodological flaw. Statistically, the older off-diet did not differ from the younger on-diet group, but the age range at which treatment began in the former patients was greater - as much as 75 days in one case. This slight confound did not cause interpretive difficulties as the analyses, except one, were uniformly negative. Nevertheless, it highlights the need for strict inclusion-exclusion criteria in sample construction. Many earlier studies of

treatment outcome have been lax in their sample specifications (see Chapter Two). Now that screening and treatment services are better integrated, a reasonable aim for future research might be to avoid including patients treated later than, say, 28 days after birth.

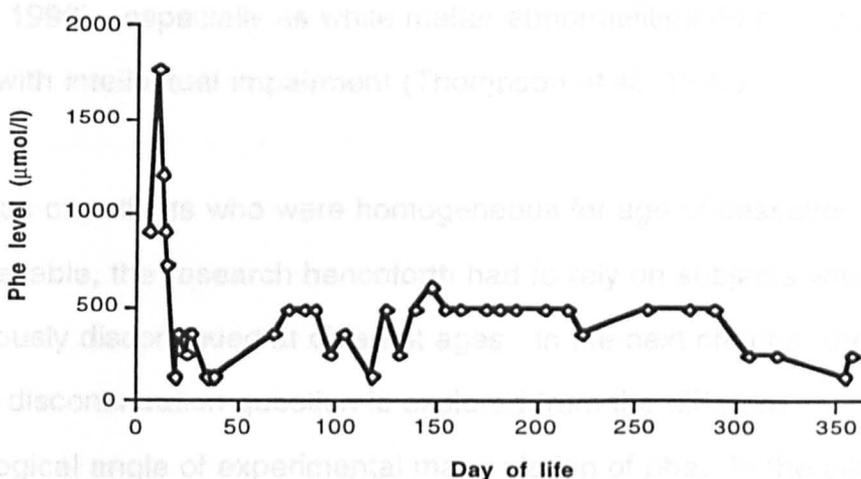
With a specific phe effect only being shown on Design Fluency and non-specific phe-effects being revealed by Peg Transfer and the Rey Labyrinth, the weight of evidence was in favour of accepting the null hypothesis. However, the risk of making a type 2 (Beta) error must be acknowledged because of the small sample sizes involved. Clearly, collaborative multi-centre research with larger, assiduously compiled samples is desirable. To date, few well-controlled studies addressing the question of treatment cessation in the older patient have emerged from the national PKU databases.

The significant interaction and PKU effect on Design Fluency are consistent with the hypothesis that executive functions are specifically impaired in treated PKU. However, these findings were not corroborated by Verbal Fluency or Letter Cancellation, tests which are regarded by some as executive measures (Kelly *et al.*, 1996). Further studies on this question are reported below in Chapters Eight and Nine.

The index of dietary control (IDC) chosen was the median capillary blood phe reading per annum. This method of quantifying dietary control is in agreement with Rupp and Burgard (1995), though some authors have used the median for six months only. Throughout the studies reported here, the IDC has been standardized on the annual median as shorter time-spans in some subjects would have reduced the number of phe-samples to single figures. Diurnal

and day-to-day phe-levels vary considerably (Rylance *et al.*, 1996). The more readings included in a unit measure, the more the standard error diminishes.

The following is an illustration of the raw data from which yearly IDCs were calculated in a single patient, the abscissa representing the date of the assay and the ordinate the blood phe level in  $\mu\text{mol/l}$ . Readings are from the first year of life, the first being taken on day 6. In the case of this child (M.McD.) the median annual level for year 1 was  $360 \mu\text{mol/l}$ . The graph below clearly shows neonatal levels well above the  $1200 \mu\text{mol/l}$  threshold for the diagnosis of classical PKU.



**Figure 10.** Case illustration of neonatal phe control.

The IDC curves in Figure 8 are characteristic of well-controlled PKU. The first-year IDC is often high as good dietary control takes a while to be accomplished. The pre-school period is usually a time of very strict control, but after school-age levels start to creep up as children become more self-determining in their eating habits, fall victim to peer pressure or actively defy

the rule of parents and dieticians (Weglage *et al.*, 1993).

At the time of writing, the possibility of replicating this study or expanding the sample no longer exists. Shortly after it was conducted, the long-standing policy in Glasgow of dietary discontinuation on the 10th birthday was reversed in favour of a permanent treatment regime. The reasons for this decision were principally that MRI abnormalities were beginning to be reported in scans of PKU patients despite early and continuous treatment (Villasana *et al.*, 1989). It was thought that dietary prolongation might furnish some protection against white matter disturbance and so it was considered prudent to advise patients to keep phe levels down indefinitely. This issue remains controversial (Brenton, 1996) - especially as white matter abnormalities do not appear to coincide with intellectual impairment (Thompson *et al.*, 1993).

As samples of patients who were homogeneous for age of cessation were no longer available, the research henceforth had to rely on subjects who had spontaneously discontinued at different ages. In the next chapter, the treatment discontinuation question is explored from the different methodological angle of experimental manipulation of phe. In the pilot study reported in Chapter Five, the matched groups design was replaced by a simple repeated measures design, and cognition and motor function studied in treated adults who returned to diet after several years of *ad libitum* phe intake.

## ***Chapter Four***

### **EXPERIMENTAL MANIPULATION OF PHENYLALANINE: THEORETICAL AND METHODOLOGICAL BACKGROUND**

In the 1980s, due to the widespread policy, especially in the USA, of stopping treatment at or around the end of the pre-school period (Schuett *et al.*, 1980), research into treatment factors mainly concentrated on intellectual outcome after dietary cessation (*e.g.* Koch *et al.*, 1982; Seashore *et al.*, 1985). The commonest methodology was the retrospective follow-up. Nevertheless, having a large sector of the PKU population off diet permitted a more experimental approach, namely, of deliberately manipulating the independent variable by reinstating treatment in previously treated but currently off-diet subjects. Exploration of possible phe-related changes in cognitive and behavioural functioning by this method could shed light on whether the neurotoxicological potency of phe extended into later childhood, adolescence or even adulthood. Moreover, the experimental approach could provide insights into the reversibility of phe-effects, if found.

In addition to these important clinical issues, controlled experimentation held the promise of expanding theoretical knowledge about the role of CNS maturation as a protective mechanism against the damaging effects of phe and the related question of how to conceptualise damage in the more mature organism. This latter issue concerns whether high levels of phe have predominantly structural or synaptic effects on neural tissue and whether the magnitude of these effects alters dynamically as a function of CNS maturation.

As discussed in Chapter One, neuropathological theories implicate the neurone as the prime target for phe. In particular, myelination was singled out as a likely candidate for disruption, though synapse formation and dendritic arborization occur simultaneously and are impaired by phe. Mental capacity is related to all three.

### **Naturalistic versus experimental methodologies in the study of phenylalanine effects**

Throughout the history of phenylketonuria research, the influence of excess, circulating phenylalanine on neural structure and psychological functioning has been investigated both by naturalistic and experimental methodologies. For reasons of treatment ethics, naturalistic studies constitute the majority of studies. These may be small-scale, as in single-centre projects where sample sizes are customarily restricted by the availability of patients; or large-scale, as in the National Collaborative Studies of the UK (Smith & Beasley, 1989) and the USA (Koch *et al.*, 1987). Irrespective of scale, naturalistic studies tend to be characterised by retrospective designs, a frequently adopted paradigm being to relate anatomical or functional data to dietary history (*e.g.* Behbehani *et al.*, 1986). Such designs necessitate careful subject selection to avoid ambiguous effects from confounding or extraneous variables, and sufficient and consistent sampling of phe levels to enable comparability between subjects. Where samples are heterogeneous or where control has been weak, problems of causal attribution can arise.

Control problems are inherent in the naturalistic approach. The independent

factor (phe level) typically varies according to forces over which the researcher may have little or no direct control, such as age of treatment onset or dietary compliance. In the experimental approach, on the other hand, the independent variable is deliberately manipulated by the researcher. Dependent variables studied are usually common to both approaches and may be neurological - for example, clinical or brain scan parameters; or psychological - for example, measures of skill, learning or adjustment.

Experimentally raising phe during dietary restriction or lowering phe following dietary discontinuation are alternative means by which treatment issues have been addressed. If, as might be inferred from Davison's position (Davison, 1973), there is an age in PKU beyond which the CNS is unaffected by hyperphenylalaninaemia, then raising phe from low levels or lowering phe from high levels, either artificially or naturally, should fail to potentiate changes in structure or function in individuals who fall into that age category. This logic has been the driving force behind the experimental approach in PKU, the origins of which date back to the 1970s.

### **Experimental phenylalanine manipulation by challenge**

By nature, the experimental approach is prospective and fosters planned comparisons. Six years or so after mass neonatal screening for PKU was introduced in Europe and the USA, the first study emerged in which phenylalanine was experimentally manipulated in an endeavour to explore functional outcome (Frankenberg *et al.*, 1973). In this prototype study, the authors selected six non-retarded children with treated classical PKU aged

four to five years and subjected them to a 20-day period of raised phe by means of a high phe supplement that simulated the average daily phe intake of a normal four-year-old. The group's mean blood phe-level rose seven-fold from 342 to 2400  $\mu\text{mol/l}$ . A double-blind, repeated measures protocol was adopted. Each day, mothers rated dimensions such as withdrawal, hyperactivity and emotionality on an *ad hoc* instrument but the results revealed no measurable behavioural consequences of the manipulation, at least in the brief term of the experiment. Thus, the study failed to support the view that temporary hyperphenylalaninaemia produces psychiatric symptomatology in treated PKU. However, caution about generalizing the finding is required, given the small sample and short exposure period.

Frankenburg *et al.*'s study illustrates the use of a challenge paradigm in experimental PKU research. In this approach, subjects who are being treated and therefore maintaining a low phe diet are temporarily exposed to high phe and concomitant measures of neurological or psychological variables taken. The integrity of the CNS is thus challenged by transitory hyperphenylalaninaemia and possible neurotoxic effects explored by tests of structural or functional change. This approach has also been used as an investigative tool with non-PKU subjects. For example in a study of the food additive aspartame which contains phe in large proportion, Zametkin *et al.* (1987) found that hyperactive children did not show behavioural toxicity effects after two weeks during which they were given phe at a level equivalent to the 99th percentile for adult consumption.

Three other studies that fall into the challenge category are those by Krause *et al.* (1985), Hogan *et al.* (1986) and Clarke *et al.* (1987), the last two

stemming from the same Canadian laboratory.

Krause *et al.* improved upon Frankenburg *et al.*'s methodology by adopting a triple-blind, double-crossover design. These authors adjusted phe level in a similar fashion by adding phe to the dietary supplement. However, their sample was very heterogeneous for age (6 to 24 years), experimental treatment periods (low vs high phe) were only a week long and reporting of data was highly selective. Of eight neuropsychological tests administered (rapid automatised naming, trail-making A and B, grooved pegboard, choice reaction time, digit span, Benton visual retention and digit symbol), Krause and her colleagues only documented results derived from choice reaction time. They reported having found an inverse relationship between transitory phe level and psychological performance, but choice RT was the sole basis for the putative effect and the claim was not substantiated by statistical analysis.

Hogan *et al.* (1986) studied seven subjects aged 13 to 17, six of whom had terminated treatment and been hyperphenylalaninaemic for four to 14 years. The age at which diet was stopped in this group was very young: between two and eight years. All had low IQs and some were mildly mentally handicapped, the overall IQ range being 60 to 90 with an average of 79. The combination of early dietary cessation and low IQ strongly suggested phe-related neuropathology. These authors first returned the children to a restricted diet with a standard phe-depleted formula (Lofenalac or Phenyl-Free), then imposed a challenge by means of a high-phe supplement and finally reinstated the low-phe supplement. During challenge, phe rose to an average of 1327  $\mu\text{mol/l}$  from an average of 713  $\mu\text{mol/l}$  during normal

treatment. The two intervals during which phe was experimentally altered were each five weeks long. Serial neuropsychological testing was conducted, the examiners being blind to condition. The measures included simple and choice reaction times, pegboard and trail-making tests, and others not specified in the report. The results showed no systematic effect of phe level on neuropsychological functioning, thus weakening the case for an element of phe neurotoxicity extending beyond childhood into adolescence. The negative findings did not provide evidence justifying the prolongation of the phe-restricted diet into late adolescence, but the borderline retarded sample studied was atypical and not representative of the well-treated PKU population.

The same Canadian group published findings a year later (Clarke *et al.*, 1987) that partly contradicted Hogan *et al.*'s (1986). Again based on a very small sample (n=9), it is not clear whether the subjects were largely those who participated in Hogan *et al.*'s study, but tested more extensively. This is the suspicion, however, as many of the features of the two samples were identical. Clarke *et al.* manipulated circulating phe by administering oral capsules containing high or low phe instead of a modified supplement. They employed a triple-blind, single cross-over design and collected data from the WISC, the WRAT and an eight-subtest neuropsychological battery, including choice reaction time, from a group of adolescents with classical PKU. Apart from the small sample size, subjects were non-homogeneous for treatment variables, some having terminated diet in infancy and showing IQs in the mildly retarded range. Clarke *et al.* found no reliable hyperphenylalaninaemia effects on any test from the WISC, WRAT or neuropsychological battery except for choice RT, thus providing some support for Krause *et al.*'s contention that choice RT

might be a phe-sensitive measure.

The strength of the challenge model lies in the tight control the experimenter can exert over the independent variable. Equivalent control over treatment variables to avoid heterogeneous samples is just as desirable but, as none of the studies so far reported was conducted on patients whose treatment conformed to modern criteria, generalization to the PKU population as treated by present-day standards is difficult. The available literature from challenge studies can inform about the effects of transitory changes in phe on conduct and psychological test performance - albeit in patients many of whom discontinued diet earlier than would now be considered safe and whose exposure periods were very varied. The results to date, however, are equivocal and suggest that neither widespread nor severe intellectual deterioration attend experimentally induced hyperphenylalaninaemia irrespective of age of treatment discontinuation. If indeed any cognitive dysfunction does occur as a response to temporarily elevated phe, it may be in the area of rapid decision-making, but research data derived from representative samples and subjected to appropriate statistical analysis is currently lacking.

### **Experimental phenylalanine manipulation by resumption of diet**

In experimental studies of the possible dynamic relationship between phe and CNS function in PKU, the dietary resumption paradigm has been more commonly employed than the challenge model. In the dietary resumption procedure, the independent variable, circulating phe, is manipulated by

returning treated patients to the restricted diet who have terminated treatment and therefore experienced a period of hyperphenylalaninaemia, perhaps several years long. Both the resumption and challenge paradigms are reversal techniques but resumption is the mirror image of challenge. Viewing each longitudinally, the challenge method follows a low-high-low sequence, while the resumption method follows a high-low-high sequence. In terms of time-series treatment designs where A represents baseline and B treatment (Vasta, 1979), challenge can be represented as BAB and resumption as ABA (bearing in mind that the pre-treatment baseline in PKU is characterised by hyperphenylalaninaemia). Ethical principles dictate that, in challenge studies, the period during which the CNS is exposed to high phe is strictly limited, subjects always being returned to diet thereafter. However, in resumption studies, some subjects choose to remain on diet indefinitely.

Apart from the difference in the sequence of conditions between the challenge and resumption approaches, the resumption method relies on manipulating phe by food and supplement variation rather than by supplement alone. In challenge studies, dietary restriction remains constant across all phases, only the supplement being altered so as to be phe-loaded or phe-limited. This readily enables blinding of subjects, parents and researchers to the condition imposed and enables tight control of the phe variable so long as subjects' dieting behaviour remains consistent. As a research procedure, the resumption method is the more erratic and difficult to control because it depends on behavioural change, and therefore motivation, on the part of subjects. Thus, unlike the challenge method, where phe intake can be altered imperceptibly, the resumption method cannot avoid awareness by participants of the transition between the high and low phases.

Achieving patient compliance with low-phe dietary treatment is perhaps the hardest task for the clinician. For the researcher, the goal may be doubly difficult to attain a) because resumption occurs after a period of unrestricted eating, and b) because it is usually imposed in adolescence or adulthood when parental mediation of dietary intake has been replaced by self-determination. Furthermore, representative samples of patients may not be achievable. Because of the self-discipline and motivation required to return to diet, those prepared to do so will inevitably constitute a self-selected subset of the treated PKU population.

Despite the intrinsic methodological shortcomings of the resumption procedure for experimentally manipulating phe levels, it has enjoyed a certain popularity in recent years amongst researchers investigating possible transitory neurotoxic effects of phe in treated PKU. Returning patients to diet in later life can provide information about adverse phe effects and their possible reversibility. Thus, if function is found to improve following dietary resumption in the older patient, this might be interpreted as evidence for reversal of phe-related toxicity occasioned by a free diet.

Although the bulk of such work has concentrated on patient groups in their teens and twenties, Michals *et al.* (1985) explored dietary resumption in a group of 12 primary school children with PKU from the USA Collaborative Study. These authors followed up a cohort of 28 well-treated children with classical PKU who started treatment early and discontinued around their sixth birthday, as was the norm in the US throughout the seventies. Of this number, 11 showed falls in IQ greater than seven points when tested after three years

on a normal diet and 24 had scholastic learning problems as evidenced by the need for repeating classes or for special tutoring. These findings principally called into question the wisdom of terminating treatment as early as six, but gave rise to the ancillary question of whether loss of intellectual function could be regained by resumption of treatment. Michals and her colleagues addressed the reversibility issue by identifying twelve patients with an average age of 11 years who, as a group, had shown a mean fall in IQ of eight points: from 96 when treatment was stopped to 88 some three years later. For a further three years, they returned this group to a phe-restricted diet, retesting them at the end of the period. The researchers were unable to report that retreatment had been effective. They observed no change in IQ and scholastic attainments showed a downward trend, suggesting that reintroducing diet later in the developmental period, even for several years, cannot rectify intellectual decline.

This finding carries important implications for the theory of brain damage in PKU. While it fails to illuminate the precise neuropathological mechanism, it points to the six to nine year period as one during which the CNS remains vulnerable. As the adverse effects of phe appeared to be irreversible, it is a reasonable conclusion that the nature of neural disruption consequent on high phe in middle childhood is likely to be structural rather than biochemical - neuronal and histological rather than synaptic and neurohumoural.

In a companion study to Michals *et al.* (1985), Schuett *et al.* (1985) surveyed a larger and more heterogeneous group from within the USA Collaborative Study who had resumed dietary treatment after a period off diet. Seventy-two patients with PKU were identified who had restarted because of clinical

problems such as hyperactivity, moodiness and skin disease; and scholastic problems such as poor classroom performance. As the average age for treatment cessation in this group was six, there were probably many who stopped even earlier. Several subjects had IQs in the mentally handicapped range, suggesting CNS damage due to poor phe control. Thus, they could not be regarded as having been well-treated by contemporary criteria. The age for resuming diet varied between six and 19 years, about 30 per cent relapsed and few were able to maintain blood phe below 600  $\mu\text{mol/l}$ . Off- versus on-diet changes were measured by a questionnaire completed by parents, teachers and clinic staff. After dietary resumption, improvements were recorded in scholastic learning, mood, activity level and maturity in 42 patients. Eleven worsened on these dimensions and the remainder stayed unchanged. No comparison was made between improvement and age of dietary cessation, however a subgroup analysis by IQ band revealed that the greatest changes for the better were reported in the mildly and severely mentally retarded patients. These were mainly in mood and maturity.

Though alterations in behaviour coincided with a return to diet, the case for attributing behavioural change to phe reduction was weakened by a lack of a significant correlation between off- versus on-diet blood phe differences and the questionnaire variables. No attempt to control extraneous factors was made and the authors did not acknowledge expectancy effects or the possibility of a Hawthorne effect accounting for improved maturity. Nor did they consider that the intrinsically healthier PKU diet - with its high vegetarian content and low proportion of refined or manufactured foods - might have accounted for a lessening of hyperactivity (Egger *et al.*, 1985).

The equivocal results of Schuett *et al.* (1985) illustrate the weaknesses of the resumption method. Apart from motivational factors and the difficulty of diminishing phe consistently across subjects, naivety on the part of subjects, caretakers and researchers as to dietary condition is impossible to achieve. This control limitation may result in positive changes in outcome variables being falsely attributed to phe reduction when effects are really due a third factor such as increased social attention or a generally better dietary regimen.

Pietz *et al.* (1993) in Germany chose performance measures rather than rating scales in a study of phe effects on attention in a group of classical PKU adults with a history of good phe control. Though only five in number, compared with Schuett *et al.*'s (1985) sample the group was homogeneous for treatment, all subjects having been treated early and continuously until 12 years of age. Furthermore, dietary compliance was good, the average index of dietary control being 418  $\mu\text{mol/l}$  during childhood. IQ scores in the 105 to 126 range at age 20 provided evidence that early damage from elevated phe during the vulnerable infancy and early childhood period had been avoided. Phe was manipulated by dietary restriction and relaxation, the study conforming to a counterbalanced rather than a reversal design. This meant, however, that subjects and investigators were aware of the conditions imposed and that subgroups were only two or three strong. A dot-pattern task was used to measure sustained attention, subjects being required to press a key with the left finger in response to a three or five dot stimulus and with the right to a four dot stimulus. The authors claimed that sustained attention worsened under the high- compared with low-phe condition though numbers were too small to permit inferential statistical testing.

Though minimal, Pietz *et al.*'s (1993) evidence hinted that functioning in treated PKU might be compromised following dietary relaxation even after prolonged and well-controlled dieting. As a sequel to this study, members from same Heidelberg group (German Collaborative Study of PKU) conducted a more thorough investigation with a larger sample (n=14) and employed a reversal design incorporating a return to baseline, though phe was again manipulated by natural diet intake rather than by supplement (Schmidt *et al.*, 1994). All subjects were measured on de Sonneville's dot-pattern task while off diet, on diet, then off again at intervals of five weeks. Average phe-levels during the three phases were 1320, 630 and 1410  $\mu\text{mol/l}$ . Data for accuracy of detection (hits) were not reported. Analyses of errors and response latencies revealed covariance between phe and performance with the middle phase (low phe) the best. The authors interpreted these findings as indicative of a temporary adverse phe effect on attention. Such a conclusion may not be so straightforward, however, as phe and practice were confounded in the first and second phases. A significant improvement in attentional performance between the second and third phases would have vindicated the conclusion, but this was not found on *post hoc* testing. Thus, the main effect was a decline in errors and an increase in speed on the dot-pattern task between the first off-diet phase (baseline) and the middle on-diet phase (treatment), but this effect was not reliable for the return-to-baseline transition, leaving open the possibility that practice or familiarity with a novel task could have been the active factor.

The combination of a possible confound and selective data reporting with hit-rates not presented, calls into doubt the credibility of the conclusion drawn by the authors that, in adulthood following good and prolonged treatment,

transient phe produces a toxic but reversible state in PKU. Further anecdotal evidence undermining the plausibility of the phe-toxicity theory comes from the authors themselves who found that subjects did not complain of relative cognitive impairment or improvement when questioned about everyday functioning during the high-low-high phe conditions. Some were university students, but individuals in this subgroup did not describe cognitive changes in relation to studying, intellectual activity and academic performance.

### **MRI data and dietary resumption**

No MRI studies published to date have followed the challenge model. However, MRI scanning in the context of dietary resumption was undertaken by McCombe *et al.* (1992) in a case study of a 19 year-old male with classical PKU. Their patient had been treated from 16 days *post partum* until 12 years when diet was relaxed. Control during childhood was described as fair rather than good, perhaps accounting for an IQ at 16 years of 88. Diet was abandoned completely at 18 and phe levels rose to around 1900  $\mu\text{mol/l}$ . Eight months later the patient developed hypertonicity in all four limbs. MRI revealed abnormalities in deep cerebral white matter. Although the authors do not say so explicitly, the suspicion must be of pyramidal system involvement. The patient was returned to diet, reducing phe levels to around 900  $\mu\text{mol/l}$ , and the motor signs gradually improved. A repeat MRI, however, was unchanged.

The amelioration of the patient's motor disorder following dietary resumption implies reversibility of the factors underlying the condition. McCombe *et al.*

considered post-diet elevated phe concentrations to be the active agent and sporadic reports of motor dysfunction, such spasticity and ataxia, associated with hyperphenylalaninaemia following treatment do exist in the literature (e.g. Rylance, 1989; Smith, 1985; Thompson *et al.*, 1990). However, the apparently reversible toxicological effect may be related to vitamin B12 rather than phe. Vitamin B12 is strongly suspected of being implicated in the aetiology of slowly progressive spastic paraparesis in PKU (Hanley *et al.*, 1993). Vitamin B12 forms part of the amino acid supplement used in treating phe and some patients discontinuing diet may run the risk of vitamin B12 deficiency if they fail to change their eating habits to include natural foodstuffs containing this vitamin (Walter, 1995).

### **Treatment policy and experimental research**

The worth of experimental approaches to treatment questions in PKU lies in their potential to relate phe and function directly, minimize extraneous effects and thereby inform policy-making. Of the two methods of varying phe intake - supplement manipulation and dietary resumption - the former is the more scientific, the latter more ethically acceptable. While foreknowledge of the hypothesis cannot be withheld in either procedure, supplement manipulation can be conducted blindly and phe more rigidly controlled as subjects are essentially passive during the periods when it is altered. A disadvantage of this method is that there are ethical constraints on the length of time dieting patients can reasonably be exposed to high levels of phe. Dietary resumption entails practical and methodological difficulties. As volunteers are required who are prepared to recommence diet, a self-selected sample will result that,

as a subgroup, may not be representative of the treated PKU population. More problematic is the degree of phe reduction and control that might be achievable in a group that previously had eaten normally. Whereas the supplement manipulation method requires no motivation to change eating behaviour, dietary resumption necessitates a fundamental realignment of motivation and eating habits, both on the part of patients and caretakers, that may be difficult to achieve.

## *Chapter Five*

### **EXPERIMENTAL MANIPULATION OF PHENYLANINE: PILOT STUDY OF DIETARY RESUMPTION IN TREATED OFF-DIET ADULTS**

With the considerations outlined in Chapter Four in mind, we faced the decision of which approach to adopt in investigating the effects of varying phe after treatment until at least 10 years of age. As screening in the West of Scotland began in the late 1960s and policy at that time was to discontinue treatment on the tenth birthday, a small group of young adults with classical PKU was available for research who had terminated treatment at age 10 and who had been on normal diet for over a decade thereafter. Individuals were approached and asked if they would be prepared to return to diet for a period of six months. Four agreed to be retreated and thus a small-scale pilot study was possible using a resumption approach. However, obtaining a potentially larger, though younger, group for a supplement approach was also a possibility as treatment policy in the early 1990s had switched from termination at 10 to indefinite continuation. These patients, who were mainly in their teens, offered the possibility of a challenge paradigm as all were on diet. The quandary was whether they could be persuaded to be temporarily exposed to high levels of phe and if so for how long.

We decided to proceed with an exploratory study of the adult group as a first step. In addition to measuring psychological concomitants of dietary resumption, we were able to investigate neurological correlates by means of

MRI scanning, clinical neurological examination and transcranial magnetic stimulation (TMS) - a technique for quantifying nerve conduction in the brain and spinal cord (Devinsky, 1993; Belmaker and Fleischmann, 1995). Funding was obtained from the Scottish Society for PKU to cover the costs of dietary preparations and scans. Sufficient funds were available to follow up the dietary resumption study with a challenge study if recruitment of subjects was successful.

### **Methodological and theoretical aims of the resumption study**

The methodological aims of the dietary resumption study were two-fold: 1) to appraise the soundness of the procedure, in particular the practicalities of returning adults to diet after many years off, and 2) to try to identify specific neuropsychological tests that would be maximally sensitive to functional change. The theoretical aim was principally to determine whether, in the mature CNS, any empirical evidence could be found either for improvement in psychological function or normalization of neurological structure (or both) that could be attributed to medium-term reduction of circulating phe.

A multiple case-study AB design was adopted. Baseline measures were taken when subjects were off-diet and compared with repeated measures taken six months after they had returned to a phe-restricted diet.

## **Hypothesis**

It was hypothesised that, if long-term exposure of the CNS to phe in adults with treated PKU has a functional neurotoxic effect, reducing the level of phe circulating in the bloodstream for a six-month period would lead to improved performance on psychological and neurological measures, and possibly result in reversal or partial reversal of white matter abnormalities.

## **Method**

### *Subjects*

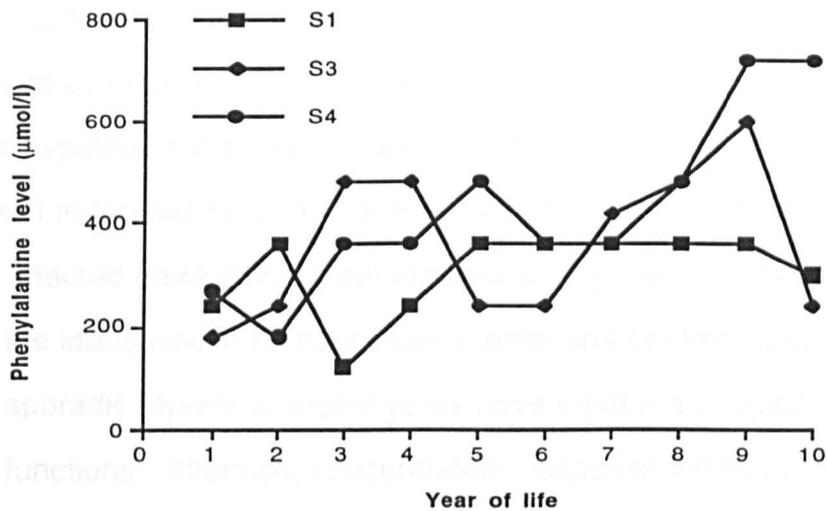
Of the four young adults who agreed to resume diet, two were males and two females. All were aged between 20 and 25 years on entry to the study and had classical PKU. Subjects 1, 3 and 4 were detected and treated early. As infants, they were among the first to be identified by the mass newborn-screening programme introduced in the West of Scotland in 1967. Only patchy information on Subject 2 was available from hospital records. This girl was diagnosed and treated elsewhere, her family moving to Glasgow when she was in her teens, at which time she was inducted into the maternal PKU programme at RHSC. In each case, including Subject 2, treatment was discontinued on or around the 10th birthday. Thereafter, eating habits were unrestricted. Table 6 shows demographic data, neonatal phe levels, age treatment begun, chronological age on entry to the pilot study and IQ for the four subjects.

No. (S)	Sex (M/F)	DoB	Neonatal phe ( $\mu\text{mol/l}$ )	Age treated (days)	CA (yrs)	IQ (FS)
1	M	02/03/67	>1250	32	25.17	101
2	F	10/03/72	>1200	?	20.17	92
3	M	26/11/68	>1250	22	23.58	122
4	F	01/02/67	>1250	12	25.33	111

**Table 6.** Patient characteristics.

The late 1960s predated modern biochemical assay techniques such as fluorimetry and high pressure liquid chromatography and measures above 1250  $\mu\text{mol/l}$  were not reported precisely. However, all subjects' neonatal phe levels exceeded the 1200  $\mu\text{mol/l}$  threshold for classical PKU.

Figure 11 shows historical phe data as indices of dietary control (median phe level per annum) for each of the first 10 years of life for subjects 1, 3 and 4. Lifetime phe indices for subject 2 could not be calculated as biochemical records were very incomplete. Despite this, sufficient data exist to ascertain that by the 10th month of life her levels were below 120  $\mu\text{mol/l}$ , though by year seven one at least had risen to around 720.



**Figure 11.** IDCs for first decade of life.

At the time, the local therapeutic target range for blood phe levels was 120 to 480  $\mu\text{mol/l}$  for the first 10 years of life. Throughout this period Subjects 1 and 3 maintained yearly averages within this band. Subject 4's control was similar until the ninth and tenth years when the upper limit was exceeded.

At the time of psychological testing, the occupations of Subjects 1 to 4 were respectively: joiner, VDU operator, university student and nurse.

*Psychological tests: measures, rationale and conceptual scheme*

Possible phe-related changes in psychological functioning were explored by a battery of eight neuropsychological tests. These were: *Matching Familiar*

*Figures, Letter Cancellation, Verbal Fluency, Design Fluency, Digits Forwards, Rey Verbal Learning, Purdue Pegboard and Stylus Tremor.*

The choice of tests making up the neuropsychological battery was dictated by their ability to cover a reasonably broad range of mental and motor functions and by the hypothesis that certain skills might be differentially phe-sensitive and deficient in treated PKU. Knowledge about whether particular cognitive areas are affected more than others remains poorly developed and insufficient to enable the identification of ideal experimental and control measures. However, sporadic reports in recent years have emerged suggesting that executive functions - attention, concentration, response inhibition and set maintenance - may be especially vulnerable (Welsh *et al.*, 1990; Diamond, 1994). For this reason, the MFFT and Letter Cancellation tests were incorporated as indices of selective and sustained attention; and Verbal and Design Fluency as indices of verbal and spatial inhibition and set maintenance. Memory impairment has been reported in older early-treated but off-diet patients with PKU (Demellweek, 1996), and so Digit Recall and the RVLTL were brought in as tests for possible short-term retentive and learning deficits respectively. Case reports of motor dysfunctions such as mild spasticity and Parkinsonian tremor in suboptimally treated patients (Rylance, 1989; Villasana *et al.*, 1989; Shaw *et al.*, 1990; McCombe *et al.*, 1992) led to the inclusion of the Purdue Pegboard and Stylus Tremor tests. In addition to documented findings, a further consideration underlying the choice of tests was that, in the PKU and maternal PKU Clinics at RHSC, off-diet patients had reported symptoms of poor concentration, memory and shakiness.

The tests divided into four pairs, each providing two measures of visual

attention/perception, ideational productivity (executive function), memory and motor coordination. Individual tests conform to the scheme in Table 7.

	<b>Domain</b>	<b>Construct</b>	<b>Test</b>
1)	Attention (Exec)	Selective attention Sustained attention	Matching figures Letter cancellation
2)	Set maintenance (Exec)	Verbal production Spatial production	Verbal fluency Design fluency
3)	Memory	Short-term Long-term	Digits forward Verbal learning
4)	Motor	Finger dexterity Motor steadiness	Pegboard Stylus tremor

**Table 7.** Test battery scheme.

*Details of the neuropsychological battery*

*Matching Familiar Figures.* Odd-numbered items from the MFFT20 (Cairns & Cammock, 1978). Ten in all. Visual matching to sample. One-minute time-limit per item. Total correct score. *Letter Cancellation.* Marking the letter 'Z' with a pencil in 50 rows of 24 random letter stimuli arranged in four columns of six. Twenty targets randomly located but never more than one in a block of six

or in adjacent rows. 1180 distractors. Upper case letters, 12-point size. Two-minute time limit. Total correct score. (Neisser, 1967). *Verbal Fluency*. Number of non-repeated words beginning with the letters 'F', 'A' and 'S' generated in one minute to each stimulus (Spreen & Benton, 1969). *Design Fluency*. Number of non-repeated designs created from two straight and two curved pieces of flat plastic in three minutes. (Griffiths, 1991). *Digit Recall*. Random digits forwards. Eight pairs. Two-long to nine-long strings. Total correct score. (Wechsler, 1992) *Verbal Learning*. Word list of 15 concrete nouns. Orally presented. Five repetitions. Recall after each presentation. Total correct score. (Rey, 1941). *Pegboard*. Modification of Purdue Pegboard. Matrix of 25 holes. Pegs inserted then collars and washers placed over pegs. Preferred hand. Time to completion. (Lafayette Instrument Company. Available from Campden Instruments Ltd., King Street, Sileby, Loughborough, LE12 7LZ, UK.) *Stylus Tremor*. Modification of Hole-Type Steadiness Tester. Insertion and withdrawal of stylus into seven holes of progressively smaller diameter without touching sides. Impulse counter measures contacts. Preferred hand. Untimed. Total contact score. (Lafayette Instrument Company. From Campden Instruments Ltd., address as above.)

### *Procedure and design*

Baseline psychological data were collected from the four subjects while they remained off-diet. The order in which the tests were given was: MFFT, Verbal Fluency, Digit Recall, Pegboard, Letter Cancellation, Design Fluency, RVLT and Stylus Tremor, *i.e.* the first attention, set maintenance, memory and motor task in the scheme shown in Table 7 followed by the second. A neurological

examination and MRI scan were also conducted. Dr R. McWilliam, Fraser of Allander Unit, RHSC, performed the neurological examination which consisted of ratings of briskness of tendon reflexes and severity of crossed adductor reflexes, maximum tandem (heel-to-toe) stance and transcranial magnetic stimulation (TMS). MRI scans were arranged by Dr D. Hadley, Institute of Neurological Sciences, Southern General Hospital, Glasgow. The reason for combining TMS with MRI was that the MR scanner available would image brain but not spinal cord. However, TMS can provide an index of spinal cord function. By measuring the evoked response in peripheral muscle following magnetic stimulation of the motor cortex, conduction time in central motor pathways can be assessed. The psychological testing and neurological examination were performed during the same out-patient appointment. MRI scanning took place within a week.

A repeated measures design was adopted with each subject acting as their own control, data first being gathered during the off-diet phase then at the end of the on-diet phase six or more months later. The independent variable, phe level, was measured concurrently along with other biochemical parameters such as tyrosine. The psychological tests were given in the same order on retest as baseline.

Informed consent and ethics approval were obtained and the subjects agreed to follow a six-month phe-restricted diet. Dietary control was supervised by Mrs B. Clark, Dietetics Department, Royal Hospital for Sick Children, Glasgow, who prepared menu lists and supplied the vitamin-enriched amino acid supplement, XP Maxamum (SHS International, 100 Wavertree Boulevard, Liverpool, L7 9PT, UK). The 'XP' prefix in the product description denotes

'without phenylalanine'.

## Results

### *Phenylalanine levels*

Table 8 shows median off-diet (baseline) and on-diet (treatment) phe levels in  $\mu\text{mol/l}$  for the four subjects with ranges in brackets. Phe readings were taken approximately fortnightly during the six-month period preceding the onset of dietary restriction and similarly during the six- or more month period of dietary restriction. With the alpha level set at 0.05, the results demonstrate that Subjects 2, 3 and 4 achieved statistically significant reductions in whole blood phe levels during the on-diet phase compared with off-diet (Mann-Whitney). Subject 1, however, did not lower his levels significantly, thus calling into doubt his compliance with the phe-restricted diet. The independent variable (phe) was therefore successfully manipulated in the desired direction in Subjects 2, 3 and 4, but not in Subject 1.

No.	Phe off-diet	Phe on-diet	U	P
1	900 (600-1038)	870 (617-1099)	103	N.S.
2	1200 (600-1363)	376 (170-1000)	6	<0.001
3	1200 (720-1200)	664 (255-1086)	17	<0.001
4	1200 (1200-1212)	253 (181-484)	0	<0.01

**Table 8.** Phe changes off and on diet. Median phe concentrations ( $\mu\text{mol/l}$ ) and ranges for the four subjects. Mann-Whitney comparisons and *P*-values. Table shows no significant reduction of phe in Subject 1.

#### *Neuropsychological test battery*

The next table contains the raw scores from the neuropsychological test battery for each of the four subjects when off and on diet. Abbreviations are as follows. *MFFT*: Matching Familiar Figures. *Cancel*: Letter Cancellation. *V/Flu*: Verbal Fluency. *D/Flu*: Design Fluency. *Digits*: Recall of forwards digits. *RVLT*: Rey Verbal Learning Test. *Pegs*: Purdue Pegboard. *Tremor*: Stylus Tremor Test. Data are all number-correct scores except those for the Purdue Pegboard, which are time-to-completion in seconds, and those for the Stylus Tremor Test, which are the number of contacts between the stylus and the holes. Thus, high scores on these last two tests indicate poor performance whereas high scores on the remainder denote good performance. The inter-test interval (ITI) in months is shown against the Subject number.

No.	ITI	MFFT		Cancel		V/Flu		D/Flu		Digits		RVLT		Pegs		Tremor	
		Off	On	Off	On	Off	On	Off	On	Off	On	Off	On	Off	On	Off	On
1	(6.87)	8	10	10	13	27	40	14	15	10	10	47	55	141	141	24	23
2	(6.63)	3	3	11	9	32	26	8	9	8	8	25	48	159	159	27	30
3	(7.57)	8	8	15	14	37	39	15	16	11	12	45	62	139	134	11	12
4	(9.30)	8	8	9	13	46	48	13	23	12	10	53	59	152	129	15	9

**Table 9.** Test battery scores.

The following table shows test score differences (d) between the dietary conditions for each subject. The polarity of the score difference indicates the direction of change in performance from the off-diet to the on-diet condition: improvement (+) or deterioration (-). The percentage (%) column shows the size of the change as a proportion of baseline level, multiplied by 100 with the sign retained.

No.	MFFT		Cancel		V/Flu		D/Flu		Digits		RVLT		Pegs		Tremor	
	d	%	d	%	d	%	d	%	d	%	d	%	d	%	d	%
1	+2	(+25)	+3	(+30)	+13	(+48)	+1	(+7)	0	(0)	+8	(+17)	0	(0)	+1	(+4)
2	0	(0)	-2	(-18)	-6	(-19)	+1	(+13)	0	(0)	+23	(+92)	0	(0)	-3	(-11)
3	0	(0)	-1	(-7)	+2	(+5)	+1	(+7)	+1	(+9)	+17	(+38)	+5	(+4)	-1	(-9)
4	0	(0)	+4	(+44)	+2	(+4)	+10	(+77)	-2	(-17)	+6	(+11)	+23	(+15)	+3	(+40)

**Table 10.** Change and percentage change scores.

When interpreting the results in the above table, it is important to bear in mind that Subject 1 failed to return to diet and thus any changes in his results cannot be attributed to diminution of phe. The other factor that could account for changes in test performance between conditions is practice as the repeated measures design meant that the subjects retook the same test at follow-up six to nine months later.

The results failed to reveal consistent gains in perceptual, cognitive and motor functioning between the off- and on-diet conditions. Subject 2 only improved on two of the eight measures; Subject 3 improved on five and Subject 4 on six, however. Subject 1 also improved on six of the measures but these increments in performance are paradoxical considering he did not show a reduction in phe level across the test-retest period and suggest another active factor. No change was recorded for Subject 1 on two measures, for Subject 2 on three, for Subject 3 on one and for Subject 4 also on one. Minor decrements in performance were recorded for Subject 2 on three measures, for Subject 3 on two and for Subject 4 on 1. Thus, the three subjects who achieved a significant reduction in phe produced six poorer scores between them on retest compared with none in Subject 1 who did not reduce his phe level.

Thus, in the three subjects who succeeded in reducing phe levels, none of the neuropsychological measures yielded systematic results indicative of improvement between the conditions with the exception of Design Fluency and the Rey Verbal Learning Test. On the Design Fluency Test, improvements were minimal for Subjects 2 and 3 and equal to the change found in Subject 1

whose phe level was unchanged. On the Rey Test, Subjects 2 and 3 made considerable gains on retest, though Subject 4 did not and did less well than Subject 1.

### *Neurological examination*

The data in the table below were collected and reported by Dr McWilliam. Columns refer to tendon reflexes (TR), crossed adductor reflexes (CAR), maximum tandem stance (MTS) and transcranial magnetic stimulation (TMS). The tendon reflexes variable is a four-point rating scale (0 = *absent*, 1 = *reduced*, 2 = *normal* and 3 = *increased*). The crossed adductor variable is also a rating scale (0 = *absent*, 1 = *present* and 2 = *pathological*). Ratings were made by the examiner. The maximum tandem stance score is the maximum time in seconds the heel-to-toe position could be held by the subject up to 25 seconds. The transcranial magnetic stimulation variable is average time in milliseconds for a single pulse over the skull above the left motor cortex to be conducted to the calf muscle of the right leg, the interval measured being that between the onset of the pulse and the onset of the EMG response.

Diet	TR (0-3)		CAR (0-2)		MTS (sec)		TMS (msec)	
	Off	On	Off	On	Off	On	Off	On
No.								
1	3	3	1	1	25	25	284	346
2	2	2	0	0	25	25	N/A	N/A
3	2	2	0	0	25	25	N/A	N/A
4	3	3	2	2	25	25	364	384

**Table 11.** Neurological examination data. Subjective ratings of tendon reflexes (TR) and crossed adductor reflexes (CAR): higher scores indicate possible pathology. Maximum tandem stance (MTS): time in seconds position held up to 25 sec. Transcranial magnetic stimulation (TMS): conduction time in msec. N/A (not available) indicates refusal to be tested by subjects 2 and 3.

All the indices chosen for inclusion in the examination were of motor functioning. Abnormal tendon and crossed adductor reflexes were found in Subjects 1 and 4. Otherwise the examination revealed no deviant neuromuscular responses. The ceiling scores on the maximum tandem stance test were contra-indicative of pathology. Administration of the transcranial magnetic stimulation test was incomplete, Subjects 2 and 3 withdrawing because of intolerance of rising energy levels in the magnetic field as the power was increased.

None of the indices showed any change between the off-diet and on-diet conditions in the subjects who demonstrably lessened their phe intake (2, 3 and 4) with the sole exception of a change in TMS for Subject 4. It is unlikely that this minor difference could have occurred as a function of the reduction in circulating phe as Subject 1 manifested an even greater shift in his TMS score

in the same direction as Subject 4 but in the absence of a concomitant phe-level change.

*Magnetic resonance imaging*

The information about the results of the MRI investigations was provided by Dr Hadley who assessed the MRI scans qualitatively. His appraisal of the images from the four subjects when off diet is summarized below.

---

<b>No.</b>	<b>Comments</b>
1	Foci of increased T2 signal
2	Periventricular white matter increase in T2 signal and foci of isolated increased T2 signal
3	Posterior periventricular white matter increase in T2 signal
4	Posterior periventricular white matter increase in T2 signal

---

**Table 12.** Summary of MRI findings.

When re-scanned at the end of the six- to nine-month period on diet, Dr Hadley reported no detectable differences in the scans of the four subjects. The minor periventricular white matter abnormalities persisted in the three subjects who had successfully restricted their phe intake, but the pattern had not changed following medium-term phe reduction.

## Discussion

Because of the AB design chosen for this pilot study, phe and practice were confounded factors. An ABA reversal design would have been preferable to separate out the effects of phe and familiarity with test demands and content. Taking return-to-baseline measures during a third experimental phase would have meant the subjects coming off diet again - a possibly unacceptable course for them - however, such a design was successfully implemented by Schmidt *et al.* (1994).

Having Subject 1 fail to return to diet as intended was a fortuitous circumstance. As Subject 1's test data were unrelated to any observed change in phe, they therefore provided a control for non-phe effects and could be compared with the other subjects' data. Apart from the Digits and Pegboard tests, all Subject 1's scores rose in varying amounts on retest. These gains in test performance strongly suggested a practice effect despite an inter-test interval of 6.87 months.

It is noteworthy that the instrument most likely to be influenced by long-term memory, the Rey Verbal Learning Test, was one of the two that showed consistent improvement across all four subjects. This may have been a savings effect with some of the items retained or at least recognised on retest. The other measure on which all subjects did better on retest was Design Fluency. This may have been a familiarity effect as the demand characteristics of the test (constructing stick patterns) are unusual.

When possible practice and familiarity effects are considered in the interpretation of the findings, there is little consistent evidence to suggest that a reduction in phe in the three adults who complied with the dietary manipulation led to improved perceptual, cognitive and motor functioning. Irrespective of the magnitude of change, 11 test scores were unchanged or worse compared with only 13 that were better and non-phe factors could have been responsible for the improvement. This unclear pattern emerged despite the three compliant subjects decreasing their average circulating phe levels to between a half and a fifth of their off-diet norm.

The principal value of this small-sample pilot study lay in its ability to highlight methodological weaknesses in the dietary resumption approach. The central problem in this method is adherence to the phe-restricted diet. The four subjects, as adults, had experienced a normal range of foodstuffs for over a decade. Dietary resumption necessitated voluntarily self-denial and a high degree of motivation to persist. Despite support from the Hospital service, one of the four subjects was unable to exercise the discipline needed to obtain a substantial reduction in blood phe.

Being unable to blind subjects and researchers to the conditions in a resumption study is another major drawback of the method. Collecting self-report data or subjective ratings has little worth because of contamination by expectancy effects. Because of this, no formal assessment of subjective state was made in the present study but, in principle, such a measure could be important. For instance, it might cut down the risk of false negatives if tests selected failed to detect real behavioural changes either through insensitivity or inappropriateness.

A further inherent difficulty with the resumption method is that, unless Guthrie cards are submitted very frequently, say daily, subjects can eat liberally but deliberately diet immediately prior to a blood test and so falsify the reading. Though it was impossible to control for wilfully fraudulent behaviour in the current study, there were no grounds for suspecting that subjects were being deliberately deceptive about their phe levels.

The independent variable in AB designs is open to confounding with practice when the treatment measure is a straightforward repeat of the baseline measure. Even an inter-test interval of over six months, as in the present study, may produce insufficient decay of memory traces or interference to prevent learning. Delaying repetition may enhance performance through consolidation, savings or a fresh problem-solving strategy.

Incorporating a control group in the design is one feasible solution to the problem of confounding phe and practice variables but, in small sample research, careful subject matching is required to avoid introducing a between-subjects confound. Building in a return-to-baseline phase and hypothesising that the negative effects of phe during this phase would disrupt the learning curve is second solution, but one that presents interpretative difficulties if test performance during the final off-diet phase is not worse than during the on-diet condition. A third solution is the use of parallel test forms; however, amongst the battery used here, only five out of the eight (Matching Familiar Figures, Letter Cancellation, Verbal Fluency, Digit Recall and Verbal Learning) lend themselves to this approach and, even for these, test-retest reliability coefficients are unknown.

## Conclusion

Clear evidence did not emerge from this small-sample pilot study of patients with treated PKU that reducing blood phe-level for six months or more after over a decade of a liberalized diet leads to improved psychological functioning, neurological functioning or normalization of minor neuroanatomical anomalies.

Methodological weaknesses undermined the interpretation of the findings. The treatment of three patients had adhered closely to the standards prevailing in the West of Scotland in the 1970s of dietary commencement in the first month of life, termination at 10 years of age and average phe levels being maintained within the 120 to 600  $\mu\text{mol/l}$  range. However, one subject's historical phe data were lost to investigation. The psychological measures were almost certainly contaminated with practice effects, a further subject's phe levels on retest showed a lack of compliance with dietary manipulation and the study was characterised by non-naivety about its aim on the part of all the subjects and the researchers.

Although the confounding variables in the psychological data could not be controlled statistically, scores from the single subject who did not diet provided points of comparison. This subject's performance either remained unchanged or improved on retest, sometimes with even greater magnitude than that of the three subjects who dieted successfully - a pattern suggestive of factors other than phe contributing to better test results.

An absence of change in the findings from the clinical neurological examination and MRI scans weakened support for the hypothesis that the CNS is sensitive to phe after at least 10 years of treatment. Medium-term attenuation of circulating phe after prolonged post-treatment exposure did not appear to provoke any detectable change in motor function or brain neuroanatomy.

In view of the methodological shortcomings of the resumption method for investigating phe effects in older PKU patients - non-naivety of subjects, difficulty of controlling the experimental variable and confounding factors - we sought an alternative investigative model. The pilot study described here amply demonstrated the severe methodological and interpretive problems intrinsic to the simple AB experimental design and we decided to abandon this model in favour of an ABA challenge paradigm with manipulation of the phe variable under the control of the experimenter. Though more exacting and ethically less defensible, we opted for a triple-blind crossover design to control for confounds and a study in which phe loading was varied by means of the amino acid supplement instead of by diet.

## **Chapter 6**

### **MANIPULATION OF PHENYLALANINE BY AMINO ACID SUPPLEMENT: COGNITIVE AND MOTOR OUTCOME**

#### **Background to study**

Aside from the author, the personnel in the research team at RHSC, Glasgow, who ran this study were Professor Cockburn (Director) and Dr Harvie of the Department of Child Health, Dr McWilliam of the Fraser of Allander Unit, Dr Logan and Mrs Cain of the Department of Biochemistry and Mrs Clark of the Department of Dietetics. Also involved were Dr Hadley from the Institute of Neurological Sciences at the Southern General Hospital, Glasgow; and Mr Nicholas Ward and Miss Caroline Smith from the Department of Psychology at the University of Stirling, both of whom were undergraduate students who assisted in the collection of some of the psychological data.

Blood phenylalanine levels in the subjects who volunteered were manipulated by means of special protein supplements prepared by Miss Katrina Roberts of the Research and Development Department of Scientific Hospital Supplies (SHS) International, Liverpool. Low and high phe supplements were coded by the suppliers who labelled each according to their order of administration for individual subjects. Each child was randomly assigned either to a low/high or high/low supplement condition and information about his or her specific protocol was kept secure by the project Director, Professor Cockburn. The Director made the group allocations and

thus knew the condition to which each child belonged, but was not involved in gathering data. The other members of the team, all of whom were involved in either biochemical, neurological or psychological testing, were blinded to the condition imposed, as were the children who participated and their parents.

The study followed a clinical drug-trial model (Stuart-Hamilton, 1995). In the psychological sector of the study, phe level was the independent variable and performance on various measures within a neuropsychological test battery the dependent variables. High phe could be thought of as the treatment or active drug, low phe as the placebo. Contrary to the usual prediction in drug trials, elevated phe level in this case was expected to be associated with deterioration of psychological performance rather than improvement. A randomised, single crossover design with repeated measures was adopted in order to separate phe effects from possible confounds such as practice (Yaremko *et al.*, 1982). The crossover design obviated the need for a control group as each subject acted as their own control (Vasta, 1979).

Subjects were selected from the West of Scotland PKU Register according to well-defined criteria in order to recruit a sample that was characterised by early and fairly strict treatment, and continuation of treatment until at least the end of the first decade of life. Following approval of the study by the Hospital and University ethics committees, an explanatory meeting was held for families identified as having suitable children at which the purpose and the method of the study were detailed. Parents and children had the opportunity to question and discuss what was intended with representatives of all the clinical disciplines on the team and were given several weeks during which to contact the staff for further information and come to a decision. From a total of

19, consent was given by 16. Reasons for refusal were not formally sought from the three families who declined to participate.

A three month period of phe-induced neurotoxicity was chosen as the effects of such lengthy exposure had not been explored before by studies employing a phe-supplement experimental paradigm. Krause *et al.* (1985) achieved only one week of hyperphenylalaninaemia in their subjects, Frankenburg *et al.* (1973) about three weeks, and Hogan *et al.* (1986) and Clarke *et al.* (1987) about five weeks. We assumed that more than doubling the phe-exposure period compared with Clarke *et al.*'s procedure would lead to increased confidence in the robustness of the neuropsychological outcome findings.

On the basis of Guttler and Lou's (1986) theory, we hypothesized that three months of elevated circulating phe levels would be sufficient to bring about changes in neurotransmitter biosynthesis in the CNS and that effects would be visible in lowered performance on neuropsychological tests. On the grounds that excessive phe decreases both tryptophan and tyrosine levels in the CNS and hence serotonin, dopamine, epinephrine and norepinephrine concentrations, Guttler and Lou argued that biochemical synaptic changes similar to those commonly found in untreated PKU (Behbehani & Langenbeck, 1982) also occur in the mature organism after treatment cessation. They concluded that, in treated older children and adolescents who discontinue diet, hyperphenylalanaemia gives rise to subtle neurological and neuropsychological deficiencies via the intervening mechanism of deficient neurotransmitter production.

The aim of the study was to test the hypothesis that transient elevation of phe

level by artificial means would result in neuropsychological dysfunction, the prediction being that phe intoxication would cause test performance to decline as a consequence of disrupted neurotransmission. If such a relationship were demonstrated, then it would have important implications for treatment policy. As all the subjects selected were over 10 years old, the age considered safe for treatment relaxation in Germany (Stellungnahme der APS, 1990), the finding of phe-related information processing deficits after this age would strengthen the case for continuation of dietary treatment or at least provoke debate as to whether the intellectual benefits of maintaining the diet outweigh the burdensome and socially handicapping nature of the treatment (Burgard *et al.*, 1997).

## **Introduction**

Little at present is known about life-span intellectual functioning in treated phenylketonuria (PKU, McKusick 26160). However, as individuals age who were detected by mass screening in the late 1960s and after (Medical Research Council, 1981), research opportunities become increasingly available for charting the long-term course of mental development in the treated disease and relating outcome to treatment history. It is now firmly established that implementation of the low phenylalanine diet is necessary as early as possible in the neonatal period to prevent subsequent mental handicap (Medical Research Council, 1993a). However, consensus on policy about optimal blood phenylalanine (phe) levels during treatment has yet to be achieved (Lancet, 1991) and the age at which prophylactic treatment may safely be discontinued remains controversial.

The principal method for illuminating the dietary termination issue has been the retrospective follow-up study in which IQ or neuropsychological test performance is compared either with phe levels during treatment or age of treatment cessation (Azen *et al.*, 1991; Smith and Beasley, 1989). Studies of early-treated subjects with good dietary control in childhood suggest that the risk of intellectual deterioration declines dramatically after the first decade (Behbehani *et al.*, 1986), the inference being that from 10 years of age onwards the nervous system may be sufficiently mature to withstand the neurotoxic influence of persistent hyperphenylalaninaemia. For example, widespread adverse effects on neuropsychological functioning were not found in a recent study of adolescents and young adults with classical PKU who were treated early (<20 days post-natally) and fairly strictly (average phe-level < 300  $\mu\text{mol/l}$  from birth to five years and <600  $\mu\text{mol/l}$  from five to 10), who discontinued the low phenylalanine diet abruptly at age 10 and who were hyperphenylalaninaemic (average >1000  $\mu\text{mol/l}$ ) for three or more years thereafter (Griffiths *et al.*, 1995). These findings corroborate IQ data from the United Kingdom National PKU Register (Smith *et al.*, 1990; Beasley *et al.*, 1994) and the German Collaborative Study on PKU (Schmidt *et al.*, 1996) which failed to demonstrate negative trends in IQ as a function of elevated phe levels after 10 years of age.

An alternative method to post-treatment follow-up in researching the dietary cessation question is experimental manipulation of phe by dietary supplement. This is a prospective approach that allows better control of the independent variable (phe level) than does retrospective food intake. The effects on neuropsychological performance of altering phe in a controlled

manner have been explored (Clarke *et al.*, 1987; Frankenburg *et al.*, 1973; Krause *et al.*, 1985), but in all these studies treatment factors were heterogeneous and none was based on samples characterized by early and continuous phe restriction.

To examine whether medium-term elevation of phe level has a deleterious outcome in treated PKU, we adopted a triple-blind, single-crossover design similar to that used by Krause *et al.*, (1985) in which we raised blood phe for three months in an on-diet sample of older children and adolescents by means of a phe complement. However, our subjects differed from those previously documented in that we included only those who had been continuously treated from early infancy until at least age 10.

We hypothesized that, if reducing phe by dietary treatment during the first decade of life confers protection on the immature nervous system, subsequent medium-term hyperphenylalaninaemia would not result in temporary neuropsychological impairment.

## **Method**

### *Subjects*

The inclusion criteria for entry into the study were classical PKU (neonatal phe above 1200  $\mu\text{mol/l}$ ), initiation of treatment within three weeks of birth, phe-control below 480  $\mu\text{mol/l}$  between birth and 5 years of age and below 900  $\mu\text{mol/l}$  between 5 and 10, continuation of dietary treatment and absence of

other diseases. Sixteen patients (10 boys, 6 girls) satisfied these criteria and, along with their parents, consented to participate following an explanation of the procedure. Full approval for the study was granted by the Ethics Committees of the Royal Hospital for Sick Children, Glasgow, and the Department of Psychology, University of Stirling. The children's mean phe levels were: pre-treatment 1893  $\mu\text{mol/l}$  (SD 523), 0-5 years 351  $\mu\text{mol/l}$  (SD 86) and 6-10 years 474  $\mu\text{mol/l}$  (SD 165). Mean age at entry was 12.60 years (SD 2.10) and mean IQ (WISC3-UK) was 100.94 (SD 10.71).

### *Dietary supplements*

Non-phe supplements were individual children's usual low phe amino acid mixture: either XP Maxamaid or PK Aid 3. High-phe supplements were these products with phe added to a value found in first-class protein. All products contained balanced mixtures of amino acids and were supplied by Scientific Hospital Supplies International Ltd., 100 Wavertree Boulevard, Liverpool, L7 9PT, UK. The high-phe supplements were specially developed so as to be facsimiles of the regular treatment products in visual appearance, smell, taste and texture.

### *Neuropsychological test battery*

The neuropsychological battery consisted of eight tests measuring between them selective attention, verbal memory, visuo-spatial memory and fine motor coordination. Individual tests were: Matching Familiar Figures (Cairns and Cammock, 1978), Rey Verbal Learning (Rey, 1941), Digits Forwards (Wechsler, 1992), Paired-Associate Learning (Wechsler, 1987), Corsi Block-

Tapping (Kolb and Whishaw, 1996), Rey-Davis Manual Labyrinth (Zangwill, 1946), Purdue Pegboard and hole-type Steadiness Tester (Campden Instruments Ltd, King Street, Sileby, Loughborough, LE12 7LZ, UK). The measures were not specifically selected for their sensitivity to executive function, a cognitive domain associated with prefrontal activity which has been hypothesized by Welsh (1996) and others as being susceptible to the supranormal phe concentrations found even in treated PKU. However, error scores on Matching Familiar Figures are regarded by some authors as indicative of executive function disorder (Pennington and Ozonoff, 1996) and, being a three-dimensional maze that requires a systematic search strategy, the Rey-Davis Manual Labyrinth satisfies the theoretical criteria for a test of executive ability (Kelly *et al.*, 1996). Including breaks, the complete battery took between 40 and 60 minutes to administer.

### *Group allocation*

Patients were randomly allocated to one of two groups. Group One consisted of ten patients, Group Two of six. The groups did not differ significantly in chronological age, sex ratio, IQ, pre-treatment phenylalanine level, age at diagnosis and start of treatment. Dietary control during treatment was assessed by calculating individual median annual phe levels over the first ten years of life (Rupp and Burgard, 1995). A group by year-of-life ANOVA of these data revealed no significant between-group effect, the mean phe levels over the period for Group One being 414  $\mu\text{mol/l}$  (SD 86) and for Group Two 410  $\mu\text{mol/l}$  (SD 89), and no significant interaction. A significant within-group effect was found ( $F = 5.24$ ,  $df 9/117$ ,  $P < 0.001$ ), indicating deterioration of dietary control over time for the sample overall. There was thus good

evidence that the groups' dietary compliance during the first ten years of treatment was comparable.

### *Procedure*

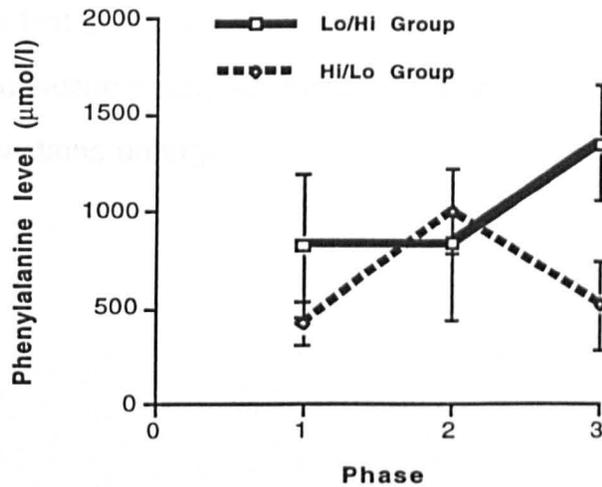
The study conformed to a triple-blind, single-crossover, repeated measures design with baseline, first experimental and second experimental phases. Circulating phenylalanine was manipulated by means of the protein-substitute supplement rather than by natural food intake. The experimental period lasted six months, switching of conditions occurring half-way. Both groups remained on their usual low-phe dietary regimen throughout. During the baseline phase, both groups took their customary non-phe protein supplement. For the three months of the first experimental phase, the ten patients in Group One continued with their non-phe supplement, but for the three months of the second experimental phase were given a supplement with added phe. They were designated the Lo/Hi Group. The sequence for the six patients in Group Two was the reverse. They were given a supplement with added phe for the first three month phase and their customary non-phe supplement for the second. They were designated the Hi/Lo Group. The perceptually identical non-phe and added-phe protein supplements were coded by the manufacturers for individual patients so that the children, their families and the research staff were unaware of the product being given during any particular phase of the study.

Phe levels were measured using high pressure liquid chromatography at baseline, at the end of the first three-month experimental phase and at the end of the second. The neuropsychological battery was repeated on the

same three occasions. Pre-test familiarization, test instructions and the test administration sequence were standardized. Biochemical and psychological testing during illness was avoided. All the children resumed their usual treatment on completion of the second experimental phase.

## **Results**

Blood phe levels recorded at baseline and at the ends of the first and second experimental periods were subjected to ANOVA (BMDP2V), which revealed significant group ( $F= 7.98, df 1/14, P < 0.05$ ), phase ( $F= 8.91, df 2/28, P < 0.01$ ) and interaction ( $F= 17.98, df 2/28, P < 0.001$ ) effects. The highly significant interaction term confirmed the intended experimental manipulation. The analysis also showed that the Hi/Lo Group's overall phenylalanine levels, irrespective of dietary condition, were significantly lower than the Lo/Hi Group's. Figure 12 shows mean phe levels with SDs for each group as a function of experimental phase.



**Figure 12.** Mean (SD) blood phenylalanine levels in  $\mu\text{mol/l}$  at baseline (1) and end of first (2) and second (3) manipulation periods for Group One (Lo/Hi) and Group Two (Hi/Lo). The group  $\times$  phase interaction was significant ( $P < 0.001$ ).

Performance on the neuropsychological tests was quantified by time in seconds to complete the Purdue Pegboard, the number of contacts between stylus and hole on the Steadiness test, and number-correct raw scores on the remainder. Results from each of the eight tests were analyzed separately by ANOVA, the interaction term providing an index of a possible phe-related crossover effect. Table 13 shows means and standard deviations for tests and conditions. None of the between-group or interaction terms was significant at  $P < 0.05$ . Similarly, within-group terms were non-significant except for Rey Verbal Learning ( $F = 14.06$ ,  $df\ 2/28$ ,  $P < 0.001$ ) and Paired Associate Learning ( $F = 4.69$ ,  $df\ 2/28$ ,  $P < 0.05$ ). The results from these two tests indicated gradual improvement over time in the performance of the group as a

whole irrespective of phe-level and suggested a practice effect with repeated testing. In view of the Hi/Lo Group's lower overall phenylalanine levels, test scores after the first and second supplement manipulations were subjected to ANCOVA with baseline phenylalanine as the covariate, but again no significant interactions emerged.

---

Supplement Manipulation

<u>Test</u>	<u>Group</u>	<u>Baseline</u>	<u>1st Phase</u>	<u>2nd Phase</u>
1) Matching Figures	Lo/Hi	6.30 (1.42)	7.00 (1.16)	6.80 (1.40)
	Hi/Lo	5.67 (1.03)	5.33 (1.51)	6.00 (1.41)
2) Rey Verbal	Lo/Hi	49.80 (3.94)	54.00 (3.09)	55.00 (4.74)
	Hi/Lo	43.67 (8.45)	49.50 (8.19)	50.33 (5.32)
3) Digits	Lo/Hi	10.10 (1.73)	10.30 (1.49)	10.40 (1.90)
	Hi/Lo	9.50 (2.66)	8.50 (2.95)	8.67 (1.75)
4) Paired Associates	Lo/Hi	27.30 (1.70)	25.90 (2.13)	27.40 (2.63)
	Hi/Lo	25.00 (3.10)	24.83 (2.86)	27.00 (1.27)
5) Block Tapping	Lo/Hi	7.90 (1.52)	7.20 (1.69)	8.30 (1.89)
	Hi/Lo	8.00 (1.79)	7.67 (2.50)	7.83 (1.47)
6) Rey Labyrinth	Lo/Hi	16.50 (3.92)	16.70 (2.95)	15.70 (3.09)
	Hi/Lo	15.00 (4.69)	14.33 (5.35)	15.50 (3.39)
7) Pegboard (s)	Lo/Hi	160.40 (8.45)	162.00(12.32)	155.90(13.15)
	Hi/Lo	154.33(11.98)	154.17(16.15)	152.83(16.50)
8) Steadiness	Lo/Hi	21.20 (9.81)	17.20(10.89)	15.20 (9.73)
	Hi/Lo	16.00 (6.87)	24.67(14.05)	13.50 (5.47)

---

**Table 13.** Results of neuropsychological tests in relation to group and experimental supplement manipulation. Means and SDs are number-correct

scores except for Pegboard which are seconds. High scores indicate good performance except for Pegboard and Steadiness where the reverse pertains. Group x phase interactions were uniformly non-significant.

## **Discussion**

Crossover effects on test performance as a result of experimentally manipulating phenylalanine would have been reflected in the interaction terms of the analyses. However, as none was significant, it appeared that medium-term changes in phenylalanine were not affecting either motor control or verbal and spatial information-processing as measured by the neuropsychological battery.

At the end of the two three-month periods of induced hyperphenylalaninaemia, the subgroups' average levels of 1355  $\mu\text{mol/l}$  and 1008  $\mu\text{mol/l}$  respectively were below those of many children with classical PKU who have discontinued diet and thus may not have provided a realistic simulation of normal levels after treatment cessation. For this reason, it would be incautious to conclude that the results unequivocally indicate age 10 as a safe time to stop dietary treatment as far as neuropsychological function is concerned. Nonetheless, Frankenburg *et al.*, (1973) raised the average phe level to 2400  $\mu\text{mol/l}$  in their subjects but found no adverse effects at the end of a three-week challenge period.

The results agree well with those of Halvorsen *et al.*, (1989) who allowed 23 PKU patients of similar age to those reported here natural protein up to the

point where blood phe levels reached 1200  $\mu\text{mol/l}$ . These authors reported no changes in their subjects' IQ or neuropsychological status after three years of the liberalized regime.

The neurotoxic mechanism by which excessive phe disrupts CNS development is poorly understood but, in untreated PKU, high phe levels appear to impair myelin, neurotransmitter and synapse formation (Hommes, 1991). Maturation of the CNS is largely complete by 10 years of age (Davison, 1973). Following treatment of PKU until this age, three months is probably too short a phe-load time-scale to potentiate changes in myelin turnover that would be reflected in functional indices such as neuropsychological measures. However, according to Guttler and Lou (1986), neuropsychological effects of the faster temporal dynamics of neurotransmitter biosynthesis might have been expected in the present sample within the manipulation period adopted. The fact they were not seen suggests that neuronal circuitry underlying verbal and spatial memory, fine motor coordination and selective attention was not affected by transient hyperphenylalaninaemia. However, the battery was not exhaustive and other psychological functions, not assessed, may be more vulnerable.

Following a suggestion by Lou *et al.*, (1985), Diamond (1994) proposed that excessive phe impairs dopamine synthesis in the prefrontal cerebral cortex which results in malfunction of executive programs governing attention, response inhibition and set maintenance. We found neither covariance between transient hyperphenylalaninaemia and attention or impulsivity as measured by the Matching Familiar Figures Test, nor deficient three-dimensional maze learning, another purported executive function, on the Rey-

Davis Manual Labyrinth. Nonetheless, short-term worsening of performance on executive tasks involving complex, speeded decision-making has been reported in older patients in whom phe was manipulated by supplement (Clarke *et al.*, 1987) and by diet (Pietz *et al.*, 1993). Though Griffiths *et al.*, (1995) found no effect of post-treatment hyperphenylalaninaemia of three or more years duration on the Rey-Davis task in subjects who had good dietary control until age 10, their data did show a non-significant trend ( $P < 0.07$ ) towards impairment on Matching Familiar Figures. In our current research programme, we are exploring the particular issue of whether executive functions are adversely affected by elevated phe levels in treated PKU.

Our findings on the Matching Familiar Figures Test are at variance with those of Davis *et al.*, (1986) though the two studies differ in design. Davis and her colleagues reported that Matching Familiar Figures discriminated averagely intelligent, treated PKU children from controls. However, they did not correlate phe levels with performance, thus failing to clarify whether the constructs of impulsivity and short attention span that the test purports to measure were associated with high historical or circulating phe. Our results suggest that temporary changes in circulating phe are not associated with these behaviours.

Neuromotor deficits have been reported in seven post-treatment PKU adults, a sample representing less than 0.5 per cent of the total United Kingdom National PKU Register (Thompson *et al.*, 1990). Only two of this number were treated early and remained well-controlled up to the same age as our patients. Three were developmentally retarded and one had epilepsy. As treatment factors and IQ were heterogeneous in this group, comparability with

ours was poor. In our early and continuously treated patients, we found no evidence of tremor or incoordination as a temporary, phe-related crossover effect.

Our results agree with those of Pietz *et al.*, (1995) who studied the immediate effects of a single oral dose of L-phe in a carefully selected group of eight adults with classical PKU who were treated early. These authors reported a steep rise in blood phe and a delayed rise in brain phe concentrations in their patients, but no impairment of attention or fine motor skill as measured by neuropsychological test performance during a 20-hour post-load period compared with pre-load baseline. A similar lack of attentional dysfunction following an acute phe load has been reported by Realmuto *et al.*, (1986), whose nine- to 20-year-old sample was closer in age to ours.

## **Conclusion**

Our findings suggest that elevation of blood-phe concentrations to between 1000 and 1300  $\mu\text{mol/l}$  on average for three months does not adversely affect verbal and spatial memory, motor coordination and selective attention in treated children with classical PKU whose intelligence is normal and who have been maintained on diet until and beyond the primary school period. These findings imply that the nervous system may be sufficiently protected against hyperphenylalaninaemia by the age of 10 or thereabouts and may not be susceptible to hypothesized neurotransmitter effects. However, the influence of longer-term exposure to excess phe on information processing remains to be researched, especially in relation to rapid decision-making and

sustained attention. It is important that patients with varying degrees of dietary control are followed up across the life span before a definitive statement is made about the optimal age for stopping treatment or whether dietary phe restriction should be a lifetime recommendation (Medical Research Council, 1993b).

## **Commentary**

A central purpose behind conducting neuropsychological studies of PKU is to determine whether subtle deficits revealed by neuropsychological tests in younger subjects might represent markers for later, perhaps more severe, impairments (Woolf, 1979). The oldest PKU patients who were detected by the mass newborn screening methods devised by Guthrie and subsequently treated are barely 30 years of age at the time of writing. Though older treated patients exist, they were generally identified and treated because of being younger siblings of children with PKU in whom the disease was diagnosed following behavioural signs of mental retardation. This post-30 group - apart from being very small in number and often late or poorly treated - are still less than 45 years old as it was 1953 before Bickel published the first demonstration of an effective dietary treatment. Thus we know nothing about the intellectual outcome of treated PKU in late middle-age onwards.

The value of conducting experimental challenge studies such as the one described above is in identifying phe-sensitive marker variables at an early stage that might have some predictive validity. In particular, psychological processes that are found to react adversely to temporary

hyperphenylalaninaemia in well-treated PKU might represent early risk indicators of later dysfunction should dietary treatment be relaxed or discontinued. After reviewing 21 published studies of neuropsychological measures in treated PKU, Waisbren *et al.* (1994) concluded that motor speed, speech, language, memory and basic logic were generally unaffected. By contrast, she and her co-authors found evidence for deficits in abstract reasoning, problem solving and executive functions. Executive functions are inherent in all purposeful behaviour and subsume planning and foresight (the guidance of actions by internal representation rather than by external stimulation), attentional deployment (selective and sustained) and self-restraint (response inhibition). The term is often used interchangeably with 'prefrontal lobe function' (Stuss, 1987).

The negative findings of the present study are consistent with Waisbren *et al.*'s view that non-executive functions are not subject to disruption by excessive phe. The results only partially agree with a recent report by Weglage *et al.* (1996). These authors found a similar lack of association between concurrent phe and digit recall in 11-year-old children, whereas performance on the Stroop and a test similar to Matching Familiar Figures was impaired. Their sample's test levels ranged between 85 and 1709  $\mu\text{mol/l}$  and thus were comparable, at least with the Lo/Hi Group's on challenge. Of the six cognitive tests in the battery, five were measures of memory and learning. Thus, a methodological weakness of the study was its failure to achieve a better balance between discriminant tasks such as memory measures and supposedly phe-sensitive tasks such as measures of executive function that may be differentially depressed by hyperphenylalaninaemia. Matching Familiar Figures and the Rey Labyrinth came closest to the criteria

for being classified as executive function tests. In failing to show a phe effect, they weakened the executive dysfunction theory, but further research is required to settle this question. A comprehensive and validated test battery suitable for assessing executive skills in children remains to be developed (Kelly *et al.*, 1996). The Stroop and Wisconsin Card Sorting tests, choice reaction time and mazes hold some promise of being able to reveal subtle executive disturbance in treated PKU but often studies on which such claims are based are characterized by small samples that are heterogeneous for age, treatment history and intellectual status (Waisbren *et al.*, 1994). The present study could not be criticized on these counts, but would have been improved by the inclusion of more tasks thought to be indicative of executive activity.

In a study very recently published, Burgard *et al.* (1997) compared two groups of early-treated adolescents on simple reaction time and de Sonneville's Dot Pattern Test (de Sonneville, 1993). The Dot Pattern Test is a speeded, decision-making task involving response suppression and purports to measure executive function. Burgard *et al.* found no difference in the level of sustained attention, as quantified by this task, between a French PKU group who had stopped treatment abruptly at 5 years and a German PKU group who had continued with the phe-restricted diet. No effect was found despite a large, though not statistically tested, difference in mean concurrent phe levels: 1272  $\mu\text{mol/l}$  for the French patients and 588  $\mu\text{mol/l}$  for the German patients. With mean ages of 13.6 and 13.1 years and mean IQs of 103 and 101 respectively, these groups were thus similar in age and IQ to the present sample. Furthermore, the British children were comparable with the French during the high-phe experimental phase and with the German during the low-

phe phase. While the children in the present sample were rendered hyperphenylalaninaemic for only three months, the French children had been subjected to high dietary phe levels for eight years on average but, notwithstanding this prolonged period of exposure, showed no neurotoxic effects on measures of executive function. Burgard *et al.*'s results therefore appear to support the conclusion of the present study that a point is reached in middle childhood at which the CNS can tolerate high circulating phe levels. However, a note of caution about Burgard *et al.*'s study needs to be sounded as their samples were neither homogeneous nor matched for severity of PKU. In the French sample, three of the nine children had mild (Type 2) rather than classical (Type 1) PKU (Guttler, 1980); in the German sample there was one Type 2 child. The relatively large proportion of non-classical children in the French sample may have diluted the effect of many years of hyperphenylalaninaemia and thus the non-significant test results may have been an artefact of sample factor differences.

An obvious anomaly in the design of the present study was the unequal subgroup numbers. This imbalance arose because of the method of group assignment adopted by the Project Director. A pure (heads/tails) randomization rather than an alternate allocation (random sequence, then odd/even) procedure was used to decide group membership. Partly because the rest of the research team were blind to the categories to which individuals had been assigned and partly because of a lack of communication between team members, this quirk did not emerge until after the final phase when the codes were broken. The analysis package (BMDP 2V) was able to handle unequal distribution totals but the large standard error inherent in the data from the smaller (n=6) Hi/Lo Group would have lowered the statistical (*i.e.*

inferential) power of the ANOVA (Gravetter & Wallnau, 1992).

A matched pairs design would have been preferable, especially in view of the significant difference between the groups in overall phenylalanine level. The lower mean phe levels shown in every condition by the Hi/Lo Group appeared to be the result of a combination of small subgroup size ( $n=6$ ) and three exceptionally well-controlled children being allocated to this group by the simple randomization procedure. These three subjects had low mean lifelong phe levels of 360, 362 and 327  $\mu\text{mol/l}$  respectively.

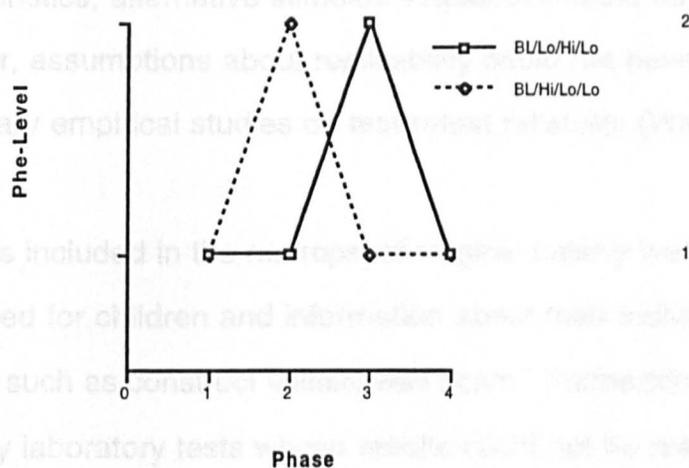
A small criticism that could be levelled at the biochemical aspect of the study is the absence of intermittent phe-level data during the experimental phases. Intra-individual diurnal fluctuation in phe concentrations is beginning to be recognized as a potentially troublesome phenomenon in PKU research. For instance, Rylance *et al.* (1996) reported a range of 130 to 280  $\mu\text{mol/l}$  in a group of 11 patients whose plasma phe concentrations were monitored for 24 hours. The group x condition interaction term was highly significant when subjected to ANOVA, thus vindicating the intended manipulation of the independent variable. However, the single phe readings taken at the end of each phase may not necessarily have been representative of average levels during the three-month experimental periods, only of levels pertaining during neuropsychological testing. It is conceivable that the significant between-group difference in phe could have arisen out of subjects in the Hi/Lo Group systematically and temporarily reducing dietary phe intake more than the Lo/Hi Group immediately prior to hospital attendance for examination. There is no evidence for this kind of deception by the subjects in the smaller group and individual differences in general dietary control or even genotype are

more likely explanations. Nevertheless, the simple expedient of gathering data from regular Guthrie tests throughout the duration of the experiment would have illuminated whether end-of-phase phe figures were representative of the phases in their entirety.

The use of the interaction term in the ANOVA should have permitted phe effects to be identified irrespective of differences in group baseline phe levels. A minor weakness in the argument for relying on the interaction term to reveal a dynamic relationship between phe and neuropsychological function, however, is whether absolute phe levels are indeed factors in the equation that can be dismissed as irrelevant. The possibility of a threshold effect has to be acknowledged where an adverse influence of elevating phe level only occurs when the baseline level surpasses a certain point. This is conjecture and there is no evidence for such a mechanism in treated PKU. In untreated PKU, when very high phe levels similar to those achieved in the present study by the Lo/Hi Group (*i.e.* >1200  $\mu\text{mol/l}$ ), are maintained from infancy until adulthood, a qualitative change appears to occur in the extent and location of white matter abnormalities in the brain (Malamud, 1966; Pearsen *et al.*, 1990). But the treated form of the disease in which experimental hyperphenylalaninaemia is imposed after the great majority of CNS maturation is complete and the untreated form cannot reasonably be compared, given the fundamental differences in onset and duration of high phe levels that delineate the two.

A further shortcoming of the experiment was its failure to incorporate a return-to-baseline condition. Such an addition would have been eminently practicable as, unlike the adults in the dietary resumption pilot study

described in Chapter Five, all the children continued with their low-phe diet and were returned to their normal phenylalanine-free amino acid supplement on conclusion of the second manipulation period. A final administration of the test battery after a further three months of normal treatment might have provided extra data on the reversibility of phe effects or, in the case of negative results, additional support for the null hypothesis. The sequence of conditions would continue to represent a single crossover model as in Figure 13. While this is a less effective demonstration of reversal effects than a double crossover model, it is ethically more acceptable and would shed a little more light on the neurotransmitter theory of treated PKU, the theory positing dynamic yet temporary and reversible alterations in rapid information processing as a function of phe level fluctuation in late childhood and adolescence (Lou *et al.*, 1987).



**Figure 13.** Phe manipulation model with return to baseline phase incorporated.

Practice effects bedevil repeated measures designs and their control necessitated the adoption of a crossover design (Stuart-Hamilton, 1996). Alternative control procedures such as preliminary asymptoting of performance would have been feasible on the two motor tasks, Pegboard and Steadiness, but not on the remainder as this approach would probably have resulted in stimulus material being transferred to long-term memory and thus being arguably less vulnerable to transitory neurochemical effects (Claridge, 1970). The employment of parallel test forms was not an option unless detailed preparatory work was undertaken. Though structurally equivalent forms for Matching Familiar Figures, Rey Verbal Learning, Digits Forwards, Paired Associates could have been constructed, equality of content and facility level could not have been achieved without extensive psychometric standardization. The Corsi Block Tapping and Rey Labyrinth tasks tap recall of novel information (spatial positions). While the apparatus for these tasks could not have been altered without changing their intrinsic demand characteristics, alternative stimulus sequences could have been devised. However, assumptions about replicability could not have been made without preliminary empirical studies on test-retest reliability (Walsh, 1989).

The tests included in the neuropsychological battery were not norm-referenced for children and information about their individual psychometric features such as construct validity was scant. Furthermore, they were uniformly laboratory tests whose results could not be readily translated into everyday psychological functioning. Finally, they were exclusively either cognitive or motor in nature and thus unable to throw light on any personality changes that might have occurred in response to raised phe concentrations. These shortcomings were not recognized at the planning stage of the study

but became evident once it was under way. In order to make good possible deficiencies in their psychometric soundness, ecological validity and coverage of personality variables, a further set of tests was introduced, the subjects being given these at the end of the first and second experimental phases. This additional battery was not prepared in time to be included when the initial baseline measures were made using the neuropsychological tests. The chapter which follows describes a subsidiary study in which results on IQ, everyday memory and problem behaviour tests were compared with manipulation of the phe variable.

## **Chapter 7**

### **MANIPULATION OF PHENYLALANINE BY AMINO ACID SUPPLEMENT: IQ, EVERYDAY MEMORY AND PERSONALITY OUTCOME**

#### **Background to study**

In the previous study, the effects of the independent variable, phe concentration, were investigated under three conditions: baseline, first experimental manipulation and second experimental manipulation. In the present study there were two only, the baseline phase being omitted. The subjects, group allocation and procedure were the same, but the following tests were given at the end of the first and second three-month experimental phases in addition to the neuropsychological battery. As before, the same measures were repeated on retest.

The aim of the study was to explore whether transitory elevation of phe level in older treated PKU children would affect cognitive functioning as evidenced by changes in IQ, everyday memory as measured by the Rivermead Behavioural Memory Test (Wilson *et al.*, 1991) and behaviour disorder as rated by the children's parents on a modified version of the Rutter Scales (Rutter *et al.*, 1970). A perennial difficulty with the interpretation of neuropsychological test results, such as those derived from the study reported in Chapter Six, surrounds their ecological validity. Positive or negative findings on laboratory-style tests may not translate readily into obvious

practical consequences or identifiable shortcomings in everyday commerce with the world (Taylor, 1988; Griffiths, 1996) and may fail to provide a firm foundation on which to build a convincing argument to put before patients and their families for dietary discontinuation or prolongation. Thus, we selected the Rivermead measure whose ecological validity is purportedly good - or at least which was constructed in the spirit of relating item content to real-life tasks. Furthermore, we realized that the experimental format of the study with both parents and children being blinded to the supplement condition presented us with a valuable opportunity to collect data on the subjective judgements of the patients and their carers about which supplement - high or low phe - was being administered. To our knowledge, no previous study had asked such a basic question, despite the history of experimental manipulation of phe in PKU research extending back over 20 years (Frankenburg *et al.*, 1973).

A short-form of the Wechsler Intelligence Scale for Children, 3rd UK Edition (WISC3-UK) was given. This comprised two subtests from the Verbal Scale - Similarities and Vocabulary - and two subtests from the Performance Scale - Block Design and Object Assembly. This tetrad of subtests was chosen for being representative of verbal and spatial cognitive functions and for its high correlation with Full Scale IQ, at least in terms of subtests from the WISC-R standardization. Applying McNemar's formula, the part-whole correlation coefficient is +0.943 (Sattler, 1974).

The Rivermead Behavioural Memory Test (RBMT) is a fairly new measure for children that is little known outside the UK. The philosophy underlying its creation is that the preponderance of existing psychometric instruments

devised to quantify memory are clinical or laboratory tasks that use artificial stimuli and have little bearing on successful day-to-day adjustment to the real-life world. Hence, the authors believe, there is a need for information about whether patients are able to meet everyday demands on memory, such as remembering appointments and people's names and faces. The strength of the RBMT, its authors claim, lies in its *ecological validity*, a concept alluded to by Cole and Bruner (1971), but articulated in detail by Brooks and Baumeister (1977) who, taking the field of mental handicap as an example, called for a move towards more naturalistic studies and, in general, applied psychological research with more environmentally and socially relevant objectives than those of the traditional laboratory. Stuart-Hamilton (1995) defines *ecological validity* as a 'term describing a study which is a realistic simulation of a real life event or which tests skills used in real life'. The RBMT consists of seven main items: recalling names, recognizing faces, locating a hidden object, remembering a story, recognizing pictures, learning a route and recalling a message. There are also temporal and spatial orientation items. Immediate and delayed recall of route learning and message recall is tested. For children over 10, the overall score is norm-referenced. No parallel forms exist.

To the author's knowledge, this is the first time an attempt has been made to investigate memory for everyday stimuli in treated PKU by means of the RBMT.

Several reports have suggested that treated PKU may be associated with poor psychosocial adjustment. Smith *et al.* (1988) gave teachers the Rutter scales to complete on 544 eight-year-old children from the UK National Study

cohort. Overall, the PKU children were rated significantly higher for neurotic and hyperactive behaviours than controls, though not for anti-social and aggressive behaviours. In an uncontrolled study of patients on the US Collaborative Study's register, Holtzman *et al.* (1986) found parental ratings of problem behaviours to be highest for children with early dietary cessation, but so was loss of IQ. Above average psychiatric symptomatology, including hyperactivity, was found in a strictly treated group from the German Collaborative Study at the age of 13, but disturbed behaviour was unrelated to historical phe levels (Burgard *et al.*, 1994). Distractability, passivity and impersistence emerged as PKU-related personality variables from a questionnaire study by Schor (1983) but, again, these traits did not correlate with either historical or concurrent phe levels.

While there is evidence that, as a diagnostic category, treated PKU carries with it an elevated risk of mild maladjustment compared to age standards, a key research issue is the underlying reason. Given the onerous nature of dietary treatment and the pressures compliance exerts on both patient and carer, it is simplistic to assume a direct link between phe level and personality development. In the untreated condition, uncontrolled circulating phe appears to potentiate a deviant personality that is typified by hyperkinesis, temper tantrums and self-punitive acts (Wood *et al.*, 1967), but these immature behavioural characteristics are common in numerous mentally handicapping conditions (Heaton-Ward & Wiley, 1984) and may not be PKU-specific. Much of the behavioural deviance reported in treated PKU may be iatrogenic, the chronic and burdensome treatment itself carrying with it the potential for peer and family conflict, feelings of alienation in the patient and, of course, the ever-present fear in the parents of damaging the child's

intellect.

In an early study of children treated until at least five years of age, Chang and Fish (1976) found the opposite to what would be predicted by the hyperphenylalaninaemia hypothesis. Paradoxically, children in their sample who were treated the longest showed the highest frequency of problem behaviours, this result providing some support for the view that dietary prolongation may be accompanied by the negative side-effect of accumulating stress. In a later study (Chang *et al.*, 1983), these authors argued that unstable behaviour associated with treated PKU may be due to non-phe factors such as parental education and implied that, when treatment is continued after the age of seven, the risk of intellectual deterioration declines to near-zero but the risk of emotional disturbance rises considerably.

The value of observing problem behaviour parameters during phe manipulation in older treated children lies in the potential power of such a model to provoke change by challenge and thereby establish whether a dynamic relationship exists between phe and behavioural deterioration when treatment administration factors are held constant. It is worth remembering that the contemporary treatment of PKU by dietary restriction is inherently a behavioural methodology that requires, for its success, the formation of atypical feeding and eating habits on the part of the parent and the patient. Raising phe by supplement allows these habits to remain unmodified and thus the design falls into the univariate category where any systematic change in personality can reasonably be attributed to phe and not extraneous variables.

The theoretical rationale for considering that hyperphenylalaninaemia in

treated PKU may influence personality functioning is the assumption that certain central aspects of individuality such as adherence to moral codes, self control and concentration are dependent on the integrity of the prefrontal cortex. The bulk of evidence for this view derives from clinical research on patients with damage to the anterior area of the brain who subsequently exhibit the typical frontal lobe syndrome of disinhibition, poor foresight and lack of moral restraint (Walsh, 1978). There is no suggestion of such a dramatic personality change being occasioned by hyperphenylalaninaemia in treated PKU subjects, but Welsh (1996) has recently proposed that more subtle - and also temporary and reversible - signs of the syndrome could arise, the underlying mechanism being disruption by phe of the dopaminergic system in the prefrontal cortex.

## **Introduction**

Classical phenylketonuria (PKU) is operationally defined as blood phenylalanine (phe) levels consistently in excess of 1200  $\mu\text{mol/l}$  after the first week of post-uterine life (Koch & Wenz, 1987). Sustained hyperphenylalaninaemia during the early years of life is neurotoxic and almost invariably results in severe intellectual impairment (Berman *et al.*, 1968). The effect is irreversible (Marholin *et al.*, 1978). The pathogenesis of mental retardation in untreated PKU is unclear. Poor dendritic arborization and reduced synaptic spine density are characteristic (Bauman & Kemper, 1982). The neuropathological mechanism is assumed to be a disturbance of myelination (Bick *et al.*, 1991), elevated levels of circulating phe during the vulnerable period of cerebral maturation appearing to compromise neuronal

conduction by reducing myelin turnover (Hommes, 1991). Neurotransmission may also be affected. The metabolic block in PKU provides a theoretical basis for a neurotransmitter hypothesis as dopamine, serotonin and norepinephrine deficiencies are consequential on the failure or partial failure of phenylalanine hydroxylation to tyrosine (Vorhees *et al.*, 1981).

The devastating effects of hyperphenylalaninaemia in infancy are well-established, however its effects on the mature nervous system in treated PKU following dietary cessation require ongoing research. While myelination is largely complete by early adolescence, elevated phe levels may adversely affect neurotransmitter activity in the short-term and such changes could be manifested in subtly impaired cognitive and behavioural functioning (Guttler & Lou, 1986). Current treatment recommendations are for dietary restriction to be implemented within the first month of life and maintained until at least 10 years of age with control well within the 120 to 600  $\mu\text{mol/l}$  range (Legido *et al.*, 1993). Briefly challenging the CNS with phe following a treatment regime that meets these criteria and measuring concomitant psychological function is a methodology that can illuminate treatment and theoretical issues in PKU such as age for dietary discontinuation and, indirectly, neurotransmitter depletion. Temporary deterioration of psychological function in response to transient hyperphenylalaninaemia might suggest premature cessation of treatment and imply short-term, but probably reversible, neurotransmitter deficiency.

In a prototype study, Frankenburg *et al.* (1973) experimentally manipulated phe intake in a small group of six non-retarded, treated PKU children aged four to five years, measuring adverse behaviours such as withdrawal, short

attention span and emotional lability in relation to two high-phe challenges. These authors noted a short-lived deterioration in behaviour coincident with onset of challenge but no sustained effect over three weeks of exposure to phe. Clarke *et al.* (1987) conducted a similar study with nine adolescents with IQs below 90 who had discontinued treatment between two and eleven years of age, returning them to a protein-restricted diet before imposing a high-phe challenge for four to five weeks. A neuropsychological battery failed to detect psychological changes following alterations in phe except for slight increases in choice reaction-time. Together, these studies suggest that widespread behavioural and cognitive deficits are not consequent on temporary elevations in phe. Also, the lack of circumstantial evidence for short-term neurotoxic effects implies that a neurotransmitter mechanism may not be involved. However, while both studies adopted a triple-blind approach, they suffered from methodological shortcomings. For instance, Clarke *et al.* (1987) did not incorporate a cross-over design to control for practice effects, neither sample conformed to current US and UK treatment recommendations of maintaining diet until at least 10 years of age (Azen *et al.*, 1991; Medical Research Council, 1993a) and Frankenburg *et al.*'s (1973) findings may have been contaminated by confounding variables.

To establish whether temporary hyperphenylalanaemia in PKU results in diminished psychological functioning when the disorder is treated according to contemporary standards, we simulated a period of temporary dietary cessation in an early and continuously treated sample of 10 to 16 year-old, non-retarded children with classical PKU. We experimentally manipulated phe intake over two successive three-month blocks, inducing hyperphenylalaninaemia in one of these, and measured cognitive and

behavioural outcome. A triple-blind crossover design was adopted to control for subject, parent, experimenter and practice variables. Psychological measures included short-form IQ, everyday memory and problem behaviours. We also asked parents and children to judge which condition (low or high phe) had been imposed.

## **Method**

### *Subjects*

A recruitment meeting attended by families with children who met the inclusion criteria of early and continuously treated classical PKU was held at which a presentation of the study was made by a senior clinician. The purpose and methodology of the study were fully explained and parents and children discussed details and concerns with the research team. Thereafter, individual families were given a few days in which to consider their decision, prior to formal consent being sought.

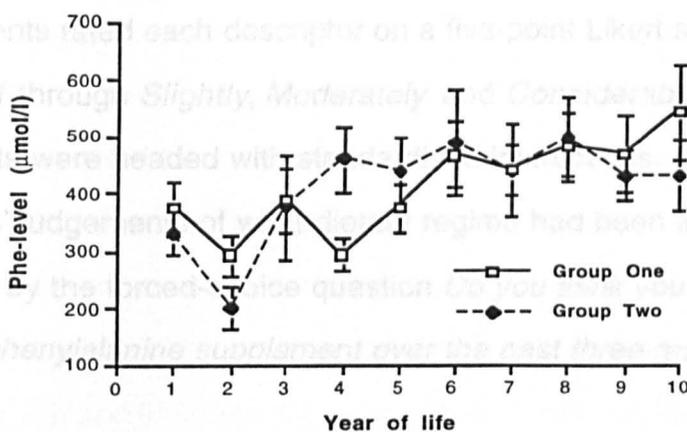
Three families declined to participate, leaving a final sample of 16 children, ten boys and six girls. All had classical PKU (*i.e.* phe > 1200  $\mu\text{mol/l}$  following neonatal screening) and had been continuously treated from early infancy. The target therapeutic range for the first ten years of life was 120 to 600  $\mu\text{mol/l}$ , some degree of dietary relaxation occurring thereafter. None had neurological or psychiatric disorder. On entry to the study, the median chronological age of the sample was 11.48 years (range 10.83 to 16.42), the median phe level 507  $\mu\text{mol/l}$  (range 209 to 1236) and the median Wechsler IQ 101 (range 89 to 124). Concurrent phe and IQ were not significantly

correlated ( $r = -0.21, p = \text{n.s.}$ ).

Each child was randomly allocated to one of two groups. Group One consisted of ten patients with a median chronological age of 12.17 years on admission to the study (range 10.83 to 16.42). Group Two consisted of six patients with a median age of 11.29 years (range 10.83 to 14.67). The groups' age distributions did not differ statistically ( $U = 20, p = \text{n.s.}$ ). There were six males and four females in Group One (60% male), four males and two females in Group Two (67% male). The mean IQ for Group One at the outset of the study was 104.90 ( $SD$  10.09), that for Group Two 99.67 ( $SD$  8.71). This difference was non-significant ( $t = 1.05, p = \text{n.s.}$ ). The mean pre-treatment phenylalanine level for Group One was 2070  $\mu\text{mol/l}$  ( $SD$  449) and for Group Two 1597  $\mu\text{mol/l}$  ( $SD$  540). This difference was also non-significant ( $t = 1.90, p = \text{n.s.}$ ). The median age at diagnosis was 16 days for both groups (range: Group One, 14 to 25; Group Two, 10 to 40) which was non-significant ( $U = 29, p = \text{n.s.}$ ). The median age for initiation of treatment was also 16 days for both groups (range: Group One, 15 to 25; Group Two, 11 to 40) which was similarly non-significant ( $U = 26, p = \text{n.s.}$ ). The groups were thus matched for chronological age, sex, IQ, pre-treatment phenylalanine level (severity), age at diagnosis and age at start of treatment.

Following the recommendations of Rupp and Burgard (1995), median annual phe levels were taken as indices of dietary control (IDCs). Each child's IDCs over the first ten years of life, *i.e.* up to the age of the youngest child in the sample, were averaged and the distributions subjected to ANOVA, although one patient from Group One had to be omitted because of incomplete historical data. No significant between-group effect was found ( $F = 0.01, df$

1/13,  $p = n.s.$ ), the mean IDC over the ten years for Group One being 414  $\mu\text{mol/l}$  ( $SD$  86), that for Group Two 410  $\mu\text{mol/l}$  ( $SD$  89). A significant within-group effect was found ( $F = 5.24$ ,  $df$  9/117,  $p < 0.001$ ), indicating deterioration of dietary control over time for the sample overall (see Figure 14), but no interaction ( $F = 1.54$ ,  $df$  9/117,  $p = n.s.$ ). The groups' dietary control during the first ten years of treatment was therefore comparable. The sample as a whole could be regarded as well-controlled, average IDCs falling within the recommended 120 to 600  $\mu\text{mol/l}$  therapeutic range in operation at the time.



**Figure 14.** Mean indices of dietary control (IDCs) with standard errors. Groups One ( $n=10$ ) and Two ( $n=6$ ) over first ten years of life. See text for analysis.

The protocol for the study was based on a triple-blind, crossover design with repeated measures, each subject acting as their own control. In each group

**Apparatus**

phenylalanine intake was experimentally manipulated over a six-month period, switching of conditions occurring at the half-way point. Both groups

1) Intellectual functioning was quantified by a short-form version of the Wechsler Intelligence Scale for Children (WISC-3 UK) comprising the Similarities, Vocabulary, Block Design and Object Assembly subtests (Sattler,

1982). 2) Everyday memory was assessed by the Rivermead Behavioural Memory Test. The adult version was chosen as no differences in level of performance between a sample comparable in age to ours and the normal adult population have been reported by its authors (Wilson *et al.*, 1991). The Rivermead contains tasks that measure memory for names, pictures, places, times, faces, routes, commands and stories. It purports to possess greater ecological validity than laboratory-style measures of memory. 3) Ratings of parents' perceptions of problem behaviours were obtained from a 20-item adjective checklist based of the Rutter Scales (Rutter *et al.*, 1970). Each item in the Problem Behaviour Checklist was a word or phrase such as *anxious, destructive, forgetful, unable to concentrate, lacking in energy, clumsy* and so on. Parents rated each descriptor on a five-point Likert scale ranging from *Not at all* through *Slightly, Moderately* and *Considerably* to *Extremely*. Checklists were headed with standardised instructions. 4) Parents' and childrens' judgements of what dietary regime had been imposed were obtained by the forced-choice question *Do you think you were given the low or high phenylalanine supplement over the past three months?*

### *Procedure*

The protocol for the study was based on a triple-blind, crossover design with repeated measures, each subject acting as their own control. In each group phenylalanine intake was experimentally manipulated over a six-month period, switching of conditions occurring at the half-way point. Both groups adhered to their normal restricted-protein dietary regime throughout the six month study period. For the first three months, however, the children in Group

One received their customary low-phenylalanine supplement but for the other three months a high-phenylalanine supplement. This group was designated the Lo/Hi Group. The children in Group Two received the reverse: a high-phenylalanine supplement for the first three months and their customary low-phenylalanine supplement for the remainder. This group was designated the Hi/Lo Group.

High phenylalanine intake was achieved by replacing the patients' usual low phenylalanine protein supplement with a formula identical in bulk, composition, appearance, taste and texture but with half the tyrosine content replaced by phenylalanine. Thus, the experimental product was the children's customary amino acid, carbohydrate, mineral and vitamin supplement but with phenylalanine added and tyrosine reduced to give an identical volume. Dietary restriction was otherwise invariant before and throughout the study and individual children remained on diet and took their daily supplement as usual, the only difference being that, during the first three-month phase for the Hi/Lo Group and during the second three-month phase for the Lo/Hi Group, the supplement contained phenylalanine equivalent to the average daily intake of someone without PKU. Both the children's usual formula and the experimental product were prepared and coded by Scientific Hospital Supplies (UK) Limited, 100 Wavertree Boulevard, Liverpool, L7 9PT. Neither the children, their parents, nor the researchers knew which product was being administered until testing was completed. The key to the code was kept by a senior member of the hospital medical staff.

WISC and Rivermead data, parents' ratings on the Problem Behaviour Checklist and parents' and childrens' guesses about the supplement were

gathered during the week before crossover at the end of the first three-month period. The testing was repeated during the week before the end of the second three-month period. The same parent who completed the first Checklist and made the first guess also responded the second time. On the second occasion, parents and children who wished to change their first guess were allowed so to do (their second guess then becoming the opposite). Parents and children made their guesses independently of each other. The Problem Behaviour Checklist was headed with standardised instructions for parents who were asked to rate their child's behaviour as they saw it over the immediately preceding week. The testing sequence was standardised to control for order effects.

Blood samples were collected at the end of each three-month period on the morning of the day of testing. Phe levels were established by high pressure liquid chromatography. All subjects returned to their restricted diet and normal protein supplement immediately on conclusion of the second three-month experimental period.

## **Results**

Phe levels were measured in  $\mu\text{mol/l}$ . Analyses of the WISC subtests were based on scaled scores and analysis of the Rivermead test on standard score totals, high scores on both representing good intellectual and memory skills respectively. Individual item-ratings on the Problem Behaviour Checklist were coded on a 0 to 4 scale (zero for absence of the problem) and these figures totalled to give a score out of 80, high scores representing maladjustment.

Table 14 shows both the independent variable (phe level) and dependent variables (psychological test and rating scores). All were analysed by ANOVA (BMDP 2V), the between-group factor being group allocation (Lo/Hi or Hi/Lo), the within-group factor phase of dietary manipulation (1st or 2nd) and the interaction term the cross-over effect. Significant interactions would have supported the hypothesis that systematic changes in phe affect intellect, memory and behaviour. According to the theoretical model, WISC and Rivermead scores should decline with rising phe, and Problem Behaviour scores increase.

		Dietary manipulation				ANOVA		
Test	Group	1st: Mean (SD)		2nd: Mean (SD)		Term	F	P
Phe:	One	836.90	(396.06)	1354.50	(291.87)	Between	7.49	0.02
	Two	1007.67	(217.54)	515.00	(236.14)	Within	0.01	n.s.
						Interaction	23.79	0.00
Sim:	One	11.70	(2.00)	11.90	(3.11)	Between	0.96	n.s.
	Two	10.17	(2.04)	11.17	(2.48)	Within	1.12	n.s.
						Interaction	0.50	n.s.
Voc:	One	9.80	(1.32)	9.80	(2.70)	Between	3.26	n.s.
	Two	7.50	(2.26)	8.50	(2.43)	Within	0.82	n.s.
						Interaction	0.82	n.s.
BD:	One	10.50	(1.72)	9.10	(2.89)	Between	0.39	n.s.
	Two	9.50	(2.59)	8.67	(2.16)	Within	6.06	0.03
						Interaction	0.39	n.s.
OA:	One	10.90	(2.13)	13.00	(2.94)	Between	2.63	n.s.
	Two	9.83	(3.31)	10.67	(1.50)	Within	3.23	n.s.
						Interaction	0.60	n.s.
RBMT:	One	22.30	(2.06)	22.50	(1.51)	Between	2.07	n.s.
	Two	20.67	(3.93)	20.50	(3.94)	Within	0.00	n.s.
						Interaction	0.08	n.s.
PBC:	One	7.20	(7.98)	15.80	(18.90)	Between	0.40	n.s.
	Two	10.67	(10.01)	5.67	(2.42)	Within	0.24	n.s.
						Interaction	3.49	n.s.

**Table 14.** Cross-over effects on independent variable (phenylalanine level) and dependent variables (psychological tests). Phe levels, and WISC (Similarities, Vocabulary, Block Design and Object Assembly), Rivermead (RBMT) and Problem Behaviour Checklist (PBC) results as a function of dietary manipulation and group. *df* 1/14 throughout. Alpha-level 0.05.

A highly significant interaction between dietary phase and group was found for blood phe levels (Table 14), indicating that phe differed according to the supplement taken and that manipulation of the independent variable was as intended. For the sample as a whole, overall phe levels did not differ over time. However, the significant between-group term indicated that, during the course of the experiment Group One's overall phe level was higher than Group Two's (mean, Group One: 1095.70  $\mu\text{mol/l}$ ; mean, Group Two: 761.33).

Except for the Problem Behaviour Checklist which showed a non-significant trend ( $P = 0.08$ ), none of the interaction terms for the psychological variables was significant, suggesting that intellectual and memory test scores and behavioural ratings were not changing consistently in a similar manner to phe levels. None of the major terms in the analyses of the psychological tests was significant, except for the within-group *F*-ratio on Block Design. This statistic revealed a reliable difference in performance on retest for the entire sample, irrespective of dietary condition. The second mean (8.94) was less than the first (10.13): a counter-theoretical effect in terms of practice as familiarity would have resulted in the opposite. A secondary analysis of Rivermead raw (unweighted) scores also yielded non-significant *F*-ratios (Between 3.46, *df* 1/14,  $p = \text{n.s.}$ ; Within 0.01, *df* 1/14,  $p = \text{n.s.}$ ; Interaction 0.07, *df* 1/14,  $p = \text{n.s.}$ ).

The proportion of parents and children whose guesses about the dietary regime agreed with actuality and with each other are shown as percentages in Table 15, chance level being 50 per cent. On a binomial test, in Group One (n=10), nine or more parents and children (90 per cent) guessing correctly would have been required for statistical significance at  $p < 0.05$ ; in Group Two (n=6) all six (100 per cent) would have had to have been correct; and for the sample as a whole (n=16) the number would have had to have been 12 (75 per cent). The data confirmed that neither the parents, nor the children in Groups One and Two, nor the sample as a whole reliably guessed which phe supplement had been administered. Likewise, inter-observer agreement, as measured by the proportion of concurring parents' and childrens' judgements, was random.

Group	Parent	Child	I/O
One	40%	60%	40%
Two	67%	33%	50%
All	50%	50%	44%

**Table 15** Guesses about dietary manipulation. Per cent correct judgements of parents and children and inter-observer (I/O: parent-child) agreement for Groups One and Two and complete sample. Chance level 50 per cent.

## Discussion

The results showed that changes in IQ, everyday memory and behaviour ratings did not parallel significant elevation and reduction of blood phe levels over three-monthly periods of experimental dietary manipulation. These negative results corroborate those of Frankenburg *et al.* (1973) and Clarke *et al.* (1987). Though restricted in size by availability of patients, the sample in the present study was nonetheless homogeneous for severity and treatment factors and, collectively, the three studies provide evidence that temporarily raising phe in treated, classical PKU does not produce a measurable neurotoxic effect, at least as reflected in tests of psychological performance.

Of the three studies, the exposure period of the present study was the longest. However, despite being subjected to three-months of elevated phe levels, the groups did not show adverse test scores at the end of these periods. A limitation of the research design was that psychological measures were not taken immediately following onset of the high phe phase. Frankenburg *et al.* (1973) found behaviour ratings worsened briefly at this time, suggesting rapid reaction of the CNS to the change in phe before stabilising 10 days later at baseline levels. Our findings cannot attest to the validity of this phenomenon, but no parent reported sudden deterioration in their child's functioning on administration of the high phe supplement either at the time or in retrospect. Furthermore, the brief behavioural disorder noted by Frankenburg and his colleagues may have been due to illness rather than phe as the two factors were confounded in four of their six subjects.

The two experimental groups were matched for demographic, treatment and IQ variables. However, they differed in overall phe level during the course of

the study, Group One being on average 334.37  $\mu\text{mol/l}$  higher than Group Two. Better dietary control in Group Two may account for this discrepancy. Greater phenylalanine hydroxylase activity is probably not an explanation as historical phe levels did not differ between the groups. Though the between-group factor for phe was statistically significant, this finding would not materially affect the interpretation of the outcome data in relation to phe as analyses centred on the ANOVA interaction term. A difference in favour of Group Two might have emerged had concurrent phe been influencing test performance, but it is noteworthy that none of the between-group terms for the psychological measures was significant.

It could be argued that the WISC, Rivermead and Rutter Scales represent psychometric instruments that are insufficiently sensitive to detect subtle phe-related cognitive and behavioural changes. Waisbren *et al.* (1994) concluded that a small body of evidence exists pointing to executive cognitive functions possibly being specifically impaired in treated PKU. Executive functions include self-programming, foresight, sustained attention and set maintenance, and are associated with the prefrontal lobes. While the tests selected do not measure these skills directly, attentional and behavioural control are necessary substrates of all and, if adversely affected by phe, might have led to lowered performance scores. Similarly, impaired executive function such as attention deficit or hyperactivity might reasonably be expected to have been revealed in disordered behaviour, but this was not registered on the Problem Behaviour Checklist. In further support of the case for executive functions being uninfluenced by transient phe, Clarke *et al.* (1987) used an extensive neuropsychological battery consisting of several measures of executive function such as the Trail Making, Stroop and Verbal Fluency tests

and failed to find a relationship. The Rivermead was chosen for its ecological validity as it purports to measure everyday memory as opposed to laboratory performance. Its lack of sensitivity to elevated phe provides some confidence that recognition and recall skills required in daily living seem not to be impaired.

Perhaps the most striking result was the inability of the children to identify from their subjective state what phe condition had been imposed. Likewise the parents' random guessing about the nature of the phe supplement from the children's general demeanour reinforces the view that elevated phe was not potentiating any discernable change. Some parents and children were adamant they knew the supplement administered but, as around half were wrong, their convictions did little to alter the picture of non-discrimination.

The experimental, within-subject, crossover design adopted here is a powerful investigative methodology, yet no significant findings emerged. Absence of deleterious psychological effects from transitory hyperphenylalanaemia in classical PKU following 10 or more years of treatment adds some weight to the idea that the CNS may be sufficiently robust to withstand a relaxed or fully liberalized diet shortly after the first decade of life, though no firm statements can be made about age for cessation of treatment on the basis of a three-month monitoring period. A universally acceptable clinical policy on age for treatment discontinuation has not been formulated. Nonetheless, evidence is beginning to converge on 10 to 12 years as an age level beyond which risks to IQ and psychosocial functioning from prolonged hyperphenylalanaemia appear to diminish dramatically (Azen *et al.*, 1991; Beasley *et al.*, 1994). A three-month

challenge to the CNS by high circulating phe levels may be too short given the insidious nature of neurological deterioration in PKU and that a lengthier period of exposure may be required to provoke measurable cognitive and behavioural change. However, support for this view was not forthcoming from a recent study of longer-term consequences of diet discontinuation (Griffiths *et al.*, 1995). In this study, deleterious effects on mental and motor functioning of three or more years exposure to phe after treatment till age 10 were not clearly evidenced.

Guttler & Lou (1986) characterized PKU in terms of level of CNS maturation, dichotomizing the condition into juvenile and adolescent forms. In the juvenile form, hyperphenylalanaemia almost invariably produces mental retardation. In the post-dietary adolescent form, hyperphenylalanaemia is hypothesized by these authors as producing subtle neuropsychological defects. The idea of a 'diet discontinuation syndrome' was not borne out by the findings of the present study. All the children experienced a state of sustained phenylalanine intoxication but did not show accompanying impairment of psychological functioning. The theory that phe disturbs myelination in the immature nervous system has already gained some empirical support (Bick *et al.*, 1991; Cleary *et al.*, 1994), but the notion that post-treatment hyperphenylalanaemia in adolescence produces sufficient disturbance either in myelin or neurotransmitter biosynthesis to generate cognitive and behavioural abnormalities (Lou *et al.*, 1985) is more contentious. The negative findings of the present study are essentially in keeping with the arrested neuronal growth hypothesis of PKU (Bauman & Kemper, 1982), a corollary of which is that neurological structures that are permitted to reach maturity by the artificial creation of a low phe environment

are subsequently spared when that environment reverts to its natural high-phe state.

## **Conclusion**

The results did not suggest a dynamic relationship between level of circulating phenylalanine and psychological functioning in children with classical PKU over the age of 10 who were treated early and continuously. An absence of medium-term hyperphenylalaninaemic effects does not imply no adverse long-term consequences. However, well-controlled children who discontinue diet at age 10 and subsequently are exposed to high phe throughout adolescence or early adulthood do not appear to show significant deterioration in cognitive and motor ability (Schmidt *et al.*, 1987; Griffiths *et al.*, 1995). The question of whether subtle disturbance of executive function occurs in relation to phe level (Weglage *et al.*, 1996b) remains open and needs further research before a definitive policy statement can be made about treatment discontinuation in PKU.

## **Commentary**

In an early report, Smith *et al.* (1973) described a downward shift in the IQ distribution of 24 treated PKU patients, the mean for the group being 93 instead of the expected population norm of 100. A later analysis of data from the UK PKU Register appeared to confirm the conclusion that, as a diagnostic category, treated PKU appears to be characterized by lower than expected IQ,

discrepancies below the norm of around seven points being found at ages four to 14 years in a cohort of 375 born between 1972 and 1978 (Smith *et al.*, 1990a). In the present study group, slippage of IQ was not found. The mean IQ of 100.94 (SD 10.71) did not differ significantly from the population mean of 100 ( $Z = +0.09$ ,  $P = \text{N.S.}$ ), suggesting that the experimental sample was drawn from a different population than those patients in the 1970s cohort.

Consistently higher lifetime phe concentrations of between 100 and 200  $\mu\text{mol/l}$  in the UK Collaborative Study probably account for the IQ mismatch. Thus, while it may be true that IQ in a large random sample of PKU subjects falls short of normative expectations, a sample carefully selected for strict phe control in the early years of life can develop intellectually in a way that is indistinguishable from normal.

Progressive reductions in the recommended upper phe-level limit of the therapeutic range in early childhood have been made over the years to the point where the proposed German ceiling is now 240  $\mu\text{mol/l}$  for the birth to 10-year period. While in reality few patients have been able to attain the previous goal of keeping levels below 360  $\mu\text{mol/l}$  (Weglage *et al.*, 1993), it is nonetheless important for research samples to be selected for what is considered to be near-optimal control. If, by so doing, normal intellectual development can be demonstrated as the outcome, the rationale for strict recommendations is fortified and principles underlying patient management procedures thereby justified.

In attempting to resolve the issue of whether deviant personality deviant or frank psychopathology might be accompaniments of treated PKU, there is

some convergence between the findings of the present study and those of Weglage *et al.* (1996a). Weglage and his colleagues in Munster investigated psychosocial adjustment in a group of treated children with classical PKU whose average age, IQ and treatment history were similar to ours. The mean age of the Munster children was 10.00 years and their mean IQ 96. Their mean age for starting diet was 17 days and mean annual phe levels rose in a linear fashion from around 240  $\mu\text{mol/l}$  in infancy to around 480  $\mu\text{mol/l}$  at age 10. Weglage *et al.* tested this group on a self-report personality questionnaire (PFK 9-14) akin to the Children's Personality Questionnaire (Cattell, 1959), but found no significant abnormalities compared with population standards, nor a sex difference or a correlation between test scores and lifetime phe levels. The self-report approach differs from the parent rating scale we adopted. Nonetheless, neither method yielded data indicating that personal adjustment in well-controlled older children might be a function of previous phe level.

A matter of concern in the present study's findings, however, is the pattern of group x phase mean differences on the Problem Behaviour Checklist. Though failing to achieve statistical significance, the mean rating for both groups during the high phe phase was twice that of the rating during the low phe phase. The exact probability for the interaction *F*-ratio of 3.49 (*df* 1/14) is 0.08 and suggestive of a non-significant trend. The small sample size would not warrant a more liberal alpha level. Nonetheless, the question arises as to whether a larger sample with similar results might have raised the power of the test to beyond the range of rejection.

Little is known about factors determining adjustment outcome in treated PKU.

Weglage *et al.* (1996a) also reported personality data from on-diet adolescent samples and found a striking difference in both self-evaluations and parent ratings of psychosocial problems between the adolescents and the primary school children despite their treatment histories being comparable. By the age of 14, the patients were describing themselves as excessively retiring, diffident, introverted, inhibited, repressed, cautious and passive in relation to age norms; and parents were using terms like 'over-protected' and 'restricted' to describe their upbringing. Weglage *et al.* concluded that these signs of maladjustment could most likely be attributed to a combination of negative attitudes towards the rigorous diet, with its potential for creating peer-group isolation, and the stress of puberty. The argument was not altogether compelling, however, as he and his coworkers also found that social and emotional problem scores correlated negatively and significantly with dietary control. It will take further research to parcel out dietary and attitudinal factors. There is a moral and clinical imperative to finding a solution as they each lead to diametrically opposed management policies. If poor dietary control in later childhood is the culprit, then prolongation of the strict diet would be required. However, if the early beneficial effects of diet later become iatrogenic, so long as intellect is not compromised, prolongation of diet into adolescence may be counter productive for personal stability and social adaptation.

Apart from personality measures such as the above and occasional educational measures (Fishler *et al.*, 1987; Weglage *et al.*, 1993), few if any studies have incorporated tests that have high ecological validity, the great majority relying on laboratory instruments. While it may emerge that a limit-testing, laboratory approach such as choice reaction-time is necessary to detect the effects in the mature organism (de Sonneville, 1991), how these

might translate into daily living and interaction with the routine world remains to be defined. Arguably, the Rivermead Test includes a number of tasks that mimic those of everyday life but, despite being performed by a CNS bathed in circulating phe exceeding of 1000  $\mu\text{mol/l}$ , these functions did not appear to suffer in consequence.

In further support of the view that elevated phe levels are harmless in the short to medium term during the later stages of CNS maturation were the random subjective judgements of the patients and their parents about the condition imposed. As with the ecological validity approach, this is a new - if obvious - extension to previous experimental research in PKU and one whose results might cause clinicians and patients to pause over interpreting symptoms of malaise as phe-related phenomena. This brings PKU management and research into the domain of attribution theory (Jaspars *et al.*, 1983). Given the present results, clinicians might be wise to be cautious about accepting older patients' constructions when they attribute problems of cognition or adjustment to high phe levels. In the current climate of alarm about loss of dietary control in adulthood (Sullivan, 1996; Koch, 1997), there is a risk of singling out phe as the causal factor without applying Occam's razor and looking for simpler explanations such as personal adequacy or changes in external circumstances.

Experimental methodologies can only be short to medium term in their influence for ethical reasons. However, the oldest individuals amongst the post-screening, treated PKU population are now nearing the end of their third decade of life and opportunities will continue to grow for long-term follow-up of basic treatment questions.

In the meantime, knowledge about the optimal phe level for clinicians and parents to aim for during treatment remains limited and warrants continuing research. The Medical Research Council's Working Party on PKU (Medical Research Council, 1993a) recommended lowering the upper limit during the pre-school years from 600  $\mu\text{mol/l}$  to 360  $\mu\text{mol/l}$ , but the empirical foundation for this change in policy was only weakly articulated in a companion paper (Medical Research Council, 1993b). The rationale for a reduction in early phe levels from the more relaxed figure of 600  $\mu\text{mol/l}$  to the current recommendation of 360  $\mu\text{mol/l}$  was largely based on findings from IQ studies (Smith *et al.*, 1990b). However, neuropsychological studies of children with varying degrees of dietary control have raised the particular question of whether phe levels during treatment previously considered harmless might adversely affect prefrontal executive function (Welsh *et al.*, 1990). It is this issue that is addressed in the following two chapters.

## **Chapter 8**

### **ASSOCIATION BETWEEN PHENYLALANINE CONTROL DURING TREATMENT AND EXECUTIVE FUNCTION**

#### **Background to study**

The first instance in the literature of the notion that treated PKU may be associated with a specific deficiency in cognitive functioning was a report by Anderson *et al.* (1969). After matching for age, sex and IQ, these authors found that 10 children with classical PKU achieved significantly worse results than controls on an adaptation of the Continuous Performance Test (Rosvold *et al.*, 1956). The Continuous Performance Test is a vigilance measure that requires the subject to respond to designated target stimuli presented discretely and sequentially within a context of distractors. The orthodox version employs the letter *X* as the target signal, the letter *A* as a priming or alerting signal and other letters of the alphabet randomly interspersed as noise; thus, the subject's objective is to identify, usually by pressing a key, as many *AX* combinations in the stimulus series as possible and reject all others.

Instead of letters, Anderson *et al.* used pictures and found that the PKU children's overall hit-to-miss ratio was poorer than the comparison group's as was the slope of their performance decrement curve. They interpreted these results as evidence for impaired attention span in PKU but failed to explain the finding either on *a priori* grounds or by correlating task parameters with historical or concurrent phe levels in an endeavour to demonstrate a direct

biochemical effect. This omission is a serious one as their sample was heterogeneous for treatment factors: four subjects had come off diet between the ages of seven and twelve, one was treated very late and another never. These last two were, respectively, mildly and severely mentally handicapped.

Despite this study's limitations in regard to the representativeness of its sample, it raised an important issue. Hitherto, it had been assumed that excess phe gave rise to global intellectual dysfunction, but Anderson *et al.* suggested that more subtle deficits might arise in the treated condition. This concern was largely overlooked until the mid-eighties when several studies on the theme of specific cognitive impairment in treated PKU were published.

Pennington *et al.* (1985) tested a group of six early-treated PKU children aged 10 who had all terminated diet at six years of age and found they had particular problems with concept formation as measured by the Wisconsin Card Sorting Test, a difficulty often associated with damage to the frontal lobe of the brain (Duncan, 1986). The same year, Lou *et al.* (1985) described three adolescents whose simple reaction times increased while indicators of serotonin and dopamine decreased following dietary cessation, suggesting a possible inverse relationship between phe on the one hand, and neurotransmission and information processing on the other. Guttler and Lou enlarged on this theory the next year, proposing that subtle neuropsychological deficits might accompany dietary cessation in late childhood and adolescence (Guttler & Lou, 1986). They reasoned that biosynthesis of dopamine, serotonin and norepinephrine would be adversely affected by hyperphenylalaninaemia in post-treatment PKU in a similar way to that found in the untreated disease and that functional disorders such as rapid

decision-making would accompany changes in neurotransmitter metabolism.

A corollary of Guttler and Lou's (1986) theory is that neurotransmitter changes due to elevated phe concentrations and their attendant neuropsychological impairments are reversible by dietary regimen. Again using simple reaction time to gauge information-processing speed, Lou *et al.* (1987) investigated the influence of elevated phe levels by means of a repeated-measures design. They first measured reaction times in nine classical PKU patients aged 15 to 24 years while off diet then re-measured times after returning the group to diet. The mean score for the off-diet condition was 254 milliseconds and for the on-diet condition 245. From these results, the authors concluded that reducing post-treatment phe levels is psychologically beneficial and that high circulating phe in late adolescence and early adulthood 'impairs mental function' (sic). However, apart from being an over-generalization from a single, low-level index of cognition (*i.e.* reaction time) to intelligence overall, such a statement was not warranted by their data. Although no analysis was provided, raw scores were published and a correlated *t*-test by the author yielded a non-significant difference ( $t = 0.46, df 9, P = n.s.$ ).

A common fault in PKU research where matched control-group designs are employed is a failure to test the matching variables for non-significance. Such a weakness undermines the conclusion of Brunner and Berry (1987) who, like Anderson *et al.* (1969), used the Continuous Performance Test to quantify sustained attention in a group of 22 well-treated PKU adolescents. These workers devised a computerised version of the AX test (see also Halperin *et al.*, 1991) and gave it both to the patients and those of their siblings who were closest in chronological age. This procedure resulted in an age discrepancy

of three years between the groups, the mean age of the PKU group being 16.4 years and that of the sibling controls 13.4 years. Brunner and Berry reported no significant difference in hit, miss or false alarm scores on the Continuous Performance Test and concluded that the study furnished no evidence for hyperdistractibility in treated PKU. However, they made no mention of the risk of a Type II error and the possibility that the increased age of the PKU subjects might have neutralised the influence of the phe factor on test outcome. A simple test (*e.g.* Wilcoxon) of the age discrepancy between the experimental and control group would have sufficed. Had a reliable difference been found, ANCOVA could have provided the necessary statistical control for the confounding factor.

A study of the executive dysfunction theory by Welsh *et al.* (1990) illustrates the need for careful control of independent variables. She and her colleagues succeeded in matching a four-year-old, on-diet, PKU group with controls on age and WPPSI IQ but not on socio-economic status. When the groups were given tasks considered by the authors to be indices of executive ability and the results analyzed by ANCOVA with socio-economic status as the correlated variable, thumb-and-finger sequencing, semantic fluency and performance on the Tower of Hanoi task were significantly worse in the PKU children. Visual search and recognising familiar pictures were not. Therefore, when account was taken of possible confounding variables and the fact that the PKU children were tested while still on diet, some support for the executive dysfunction theory emerged. Further weight was given to the idea by the finding that a composite test score correlated negatively with mean lifetime phe level ( $r = -0.62$ ,  $P < 0.05$ ).

Welsh *et al.*'s study provided some of the best empirical evidence thus far that specific cognitive deficiencies might arise even in well-controlled PKU. The authors argued that the deficit was indicative of dysfunction in the frontal cortex of the brain, the tests which showed maximum disturbance being those most associated with motor sequencing, planning and set maintenance. They speculated that a biochemical deficiency centring on the fronto-cortical dopamine system underlay the cognitive impairment. This line of reasoning was theoretical and derived from the knowledge that, as the metabolic block in PKU occurs at that point in the pathway between phe and tyrosine, concentrations of metabolites beyond the block such as dopamine should be depleted.

While Welsh *et al.*'s dopamine account was purely conceptual and *a priori*, physiological data were collected by Diamond *et al.* (1994) from a rat study which suggested a link, possibly mediated by neurochemical changes in frontal cortex, between hyperphenylalaninaemia and maze performance. Diamond *et al.* injected an experimental group of rat pups with phe to a level six times that of normal at four days *post partum*, maintaining this regime on a daily basis until sacrifice around 50 days. In addition to injecting phe, these workers administered the PAH inhibitor alpha-methylphenylalanine. Experimental animals were compared at intervals with placebo controls on immediate and delayed alternation behaviour in a T-maze.

Diamond and her colleagues found only weak signs of impaired T-maze performance in infant pups, but stronger evidence for both immediate and delayed response inaccuracies in the juveniles, suggesting a cumulative phe effect over the developmental period. Biochemical analysis of brain

monoamine levels at biopsy yielded only one finding: a reduction of homovanillic acid in the experimental animals confined to the medial prefrontal cortex. Homovanillic acid is a metabolite of dopamine and thus a marker for activity in dopaminergic neurone systems. Norepinephrine (noradrenaline) showed no effect of hyperphenylalaninaemia.

As a conclusion to this comparative study, Diamond *et al.* provided both a theoretical account of the possible neurochemical and neuropsychological effects of elevated phe levels on brain function and a warning about acceptable phe levels during dietary therapy in humans. She and her co-authors suggested that excess phe has a selective negative influence on prefrontal cortex and specifically disrupts the dopamine system. Tyrosine is the immediate precursor of dopamine. In their view, tyrosine depletion, which is a direct consequence of PAH deficiency, results in dopamine reduction which, in turn, has maximal adverse effects on the prefrontal cortex because this area of the brain is especially rich in dopaminergic neurons. The behavioural outcome of a shortage of dopamine is malfunction of executive processes such as self-programming, sequencing and impulse inhibition.

Almost coincident with the rat study report, Diamond published the results of a parallel study into human PKU in which she claimed to have found evidence for previously accepted phe levels during treatment having a specific and malign influence on executive cognition (Diamond, 1994). Such a conclusion understandably generated consternation amongst clinicians managing the treatment of patients with PKU and the debate continues as to what phe levels might be considered safe at given ages throughout the developmental period.

In her single-authored study *Phenylalanine levels of 6-10 mg/dl may not be as benign as once thought*, Diamond (1994) made a case for reducing the upper limit of the therapeutic range. Until the early 1990s, and especially in the United States, 600  $\mu\text{mol/l}$  had been the generally agreed safe ceiling for phe during infancy and the pre-school years of treatment. (To convert mg/dl into  $\mu\text{mol/l}$ , multiply by 60.) Diamond's evidence was derived from scores on two tests: Piaget's A/not B task and a simplified Stroop measure.

The A/not B task follows a delayed, two-choice, alternation paradigm in which subjects are required to uncover a hidden toy in one of two wells on two consecutive trials; upon success, the side of hiding is switched and performance quantified by error (*i.e.* perseveration) scores. This task is similar to the delayed response technique used to study frontal lobe functioning in monkeys (Diamond & Goldman-Rakic, 1986). It purports to reveal the ability to inhibit the stronger response tendency and to reflect maturation of behavioural control systems in the frontal cortex (Kelly *et al.*, 1996). Diamond modified the traditional Stroop test by replacing the colour-word interference stimuli with pictures of the night sky, to which subjects had to respond with the word *day*, and with pictures of the sun, to which subjects had to respond with the word *night*. Thus, the stronger response had again to be suppressed.

Compared with control subjects, treated PKU children aged between six months and seven years did poorly on the A/not B and the Stroop-like tasks, but were not significantly different on six discriminatory tasks of spatial discrimination, recognition memory, visual comparison, delayed non-matching to sample, global-local decision-making and line bisection.

Diamond interpreted these findings as indicative of dopamine-related, prefrontal impairment along with selective sparing of the temporo-parietal cortex, despite treatment parameters in her largely pre-school PKU sample conforming to phe levels recommended at the time.

The impact this paper has made on thinking about phe control in early life is disproportionate to its substance. The acceptability of the conclusion is marred by the fact that no numerical data or statistical analyses were reported, nor IQ measures to allow the reader to determine whether the PKU and control groups were matched for general intelligence. Furthermore, Diamond did not provide a breakdown of test performance by age, saying only that three PKU subgroups were tested: infants, toddlers and pre- and early primary schoolers. The importance of this omission is that the report states that each subgroup was tested on a different number of measures at different intervals without making it clear whether the pattern of apparent impairment was consistent at each age level. Finally, no correlational data were presented to support the conclusion that the cognitive deficiencies were directly associated with elevated phe concentrations.

In view of these methodological shortcomings, and the last-mentioned in particular, we embarked on a study of the executive dysfunction hypothesis in which neuropsychological test performance in a group of treated 10 to 13 year-old children with PKU was correlated with annual phe concentrations across their lifetime and with concurrent levels. Our study strove to establish clearly the relationship between the independent variable, phe, and cognitive and personality outcome following an extensive treatment period. We selected experimental tests considered to be sensitive to impairment of the

prefrontal area of the brain and discriminative tests considered to be sensitive to posterior impairment in order to investigate the possible specificity of cognitive deficits. We also added a personality profile measure in an endeavour to extend knowledge about whether putative prefrontal cognitive impairment translates into personal and social maladjustment in treated PKU.

## Introduction

A safe upper limit for circulating phenylalanine (phe) during dietary treatment of classical phenylketonuria (PKU) in early childhood has yet to be established (Lancet, 1991). During the seventies and eighties, 600  $\mu\text{mol/l}$  was widely accepted in the USA and UK (Schuett & Brown, 1984; Smith *et al.*, 1973). The recommended pre-school level has since been revised downward in the UK (Medical Research Council, 1993a) and Germany (Burgard, 1994) to 360  $\mu\text{mol/l}$ . Findings from IQ (Beasley *et al.*, 1994) and neuropsychological studies have provided indirect support for these changes (Waisbren *et al.*, 1994). Pennington *et al.* (1985) and Welsh *et al.* (1990) claimed that executive skills - attention, planning, creativity and response inhibition - are especially vulnerable to elevated phe levels and Diamond (1994) specified that, in the early years of life, phe-levels in the range 360 to 600  $\mu\text{mol/l}$  may be especially harmful to the development of executive function. Her results suggested that infants and young children with phe levels in this range perform badly on executive tasks, but not on those of perception and memory. She proposed that executive functions are primarily mediated by the prefrontal cortex and that neurotransmission in this area of the brain is particularly susceptible to dopamine depletion (McDowell, 1996),

deficient dopamine being a consequence of the blockage in the metabolic pathway between phenylalanine and tyrosine.

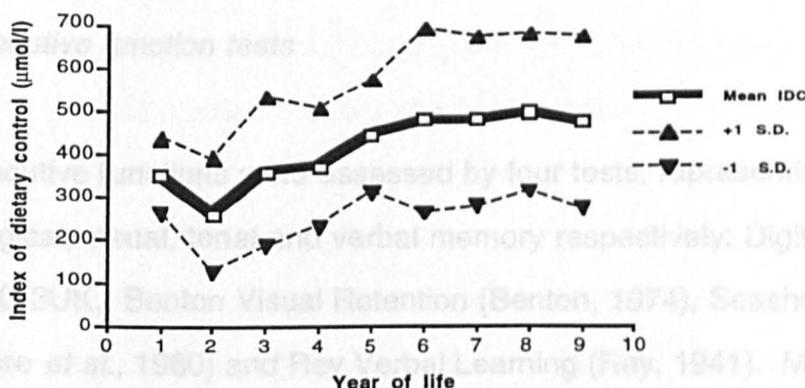
We investigated the executive dysfunction theory by means of a correlational study in which we tested a group of 10 to 13 year-old children with classical PKU on a battery of neuropsychological and personality measures. We selected children in whom individual dietary control in early life had varied around present UK recommendations, median annual levels for individuals in the first five years of life having ranged from 120 to 810  $\mu\text{mol/l}$  and in the next four years from 120 to 900  $\mu\text{mol/l}$ . At the time the study was conducted, levels ranged from 141 and 1053  $\mu\text{mol/l}$ . According to the theory, we hypothesized an inverse relationship between phe and performance on executive tasks, but not on non-executive tasks. In addition, on the assumption that personality disorders such as inattention, poor self-regulation and impulsivity reflect prefrontal lobe dysfunction (Hinshaw, 1994), we hypothesized a direct relationship between phe and measures of these variables.

## **Method**

### *Subjects*

The sample consisted of fifteen children (10 boys, 5 girls) from the West of Scotland PKU Register. All had peak blood phenylalanine levels in the post-natal period above 1200  $\mu\text{mol/l}$  (mean 1805, SD 591). Each had been diagnosed and placed on a low phe diet during the first four weeks of life (mean 12.07 days, SD 5.01). They were continuously treated thereafter.

Blood phe was measured by high-pressure liquid chromatography using an amino acid analyser. The mean lifetime phe level for the group was 411  $\mu\text{mol/l}$  (SD 178): 355  $\mu\text{mol/l}$  (SD 144) for the first five year period and 480  $\mu\text{mol/l}$  (SD 193) for the second four year period. At the time of psychological testing, mean phe (concurrent level) was 547  $\mu\text{mol/l}$  (SD 270). The index of dietary control (IDC) adopted was the median value of each subject's annual phenylalanine readings (Rupp and Burgard, 1995). Figure 15 shows mean IDCs with standard deviations for ages 1 to 9. The mean age of the sample when tested was 12.13 years (SD 1.21) and the mean IQ was 105.47 (SD 8.64).



**Figure 15.** Dietary control during first nine years of life ( $n=15$ ). Annual Index of Dietary Control (IDC) means and standard deviations (see text for explanation of IDC derivation and analysis).

#### Executive function tests

Executive functions were assessed by four tests: Mazes, Stroop, Letter

Cancellation and Design Fluency. The Mazes subtest from the WISC-3UK (Wechsler, 1992) was chosen as a measure of planning, the Stroop of response inhibition, Letter Cancellation of sustained attention and Design Fluency of creativity (Beaumont, 1983). The Stroop test was the orthodox word-colour interference version, the number-correct score being recorded (Stroop, 1935). The Letter Cancellation Test was a visual search task for 21 target letters irregularly located in an array of 1050 random distractor letters, a time limit of two minutes being set and hits per minute scored (Neisser, 1967). The Design Fluency test required construction of different designs from two curved and two straight pieces of plastic, the number of non-repeated designs created in three minutes being recorded (Griffiths, 1991).

#### *Non-executive function tests*

Non-executive functions were assessed by four tests, representing phonological, visual, tonal and verbal memory respectively: Digit Span from the WISC-3UK, Benton Visual Retention (Benton, 1974), Seashore Rhythm (Seashore *et al.*, 1960) and Rey Verbal Learning (Rey, 1941). Memory tests were chosen as control measures because of their relative immunity from the effects of prefrontal lesions (Beaumont, 1983). Benton VRT Forms C and D, with 10 second presentation and immediate recall, were given and the number correct out of 20 recorded. Sections A, B and C of the Rhythm subtest of the Seashore Measures of Musical Talents were administered to give a total-correct score out of 30. The 12-word list from the Rey Verbal Learning Test was presented repeatedly five times and the total number recalled out of 60 recorded.

### *Psychosocial functioning*

Maladjustment was measured by the child and adolescent forms of the Devereux Scales of Mental Disorders (Naglieri *et al.*, 1994), these parent-report questionnaires containing 110 items about overt behaviour rated on a 5-point Likert scale. The six scales - Conduct, Attention, Anxiety, Depression, Autism and Acute Problems - are factor-based and norm-referenced.

Devereux forms were completed by one of the children's parents on the day of testing and *T*-scores derived for the total scale and the six subscales.

### *General procedure and design*

Subjects were tested individually and given standardised instructions. Pre-training was given to ensure understanding of demand characteristics. Tests were presented with executive and non-executive measures alternating in the following sequence: Mazes, Digit Span, Letter Cancellation, Rey Verbal Learning, Design Fluency, Benton VRT, Stroop and Seashore Rhythm. The study conformed to a correlational design, though comparisons with non-PKU population norms were possible for Mazes, Digit Span and the Devereux Scales (*Z*-tests).

## **Results**

One child obtained a maximum score on the Stroop test and two on the Benton, otherwise floor and ceiling effects were absent.

Phenylalanine control was determined by analysing annual IDC distributions over the first nine years of life, this being the period for which complete historical phe data were available for every subject (see Figure 15). One-way analysis of variance was highly significant ( $F=3.57$ , d.f. 8/126,  $P<0.001$ ), indicating a gradual worsening of dietary control. In the first five-year period, one child's IDCs in Years 3 and 4 were 810 and 720  $\mu\text{mol/l}$  respectively, the others' staying between 120 to 600  $\mu\text{mol/l}$ . Between Years 6 and 9, about three-quarters maintained IDC levels below 600  $\mu\text{mol/l}$ , but some among the remainder occasionally reached 900  $\mu\text{mol/l}$ .

To provide a historical measure of dietary control, annual cumulative IDCs for the first nine years of life were calculated for each child. These data and concurrent levels were correlated with each of the psychological variables. Scaled scores were used for the WISC Mazes and Digit Span tests,  $T$ -scores for the Devereux Scales and raw scores for the rest. High scores on the cognitive tasks represented good performance, high scores on the personality measures maladjustment. The age range of the sample was 3.58 years. Test performance was not positively correlated with age except on Design Fluency ( $r=0.57$ , d.f. 13,  $P<0.05$ , Spearman, one-tailed). Partial correlations with age as the controlled variable were therefore performed on data from this test (Daniel, 1990).

Tests	Year of life and concurrent phe levels									
	1	2	3	4	5	6	7	8	9	Conc
WISC Mazes (E)	0.05	0.17	0.11	0.04	0.10	-0.03	-0.06	-0.12	-0.11	-0.29
Stroop (E)	-0.30	0.03	0.08	0.21	0.13	-0.03	-0.01	-0.03	-0.01	-0.11
Letter Cancellation (E)	0.02	0.11	0.20	0.21	0.15	0.22	0.05	0.12	0.10	-0.06
Design Fluency (E, P)	-0.24	-0.31	-0.40	-0.42	-0.42	-0.39	-0.39	-0.46*	-0.45*	-0.19
WISC Digits (NE)	0.14	-0.21	-0.26	-0.28	-0.38	-0.41	-0.35	-0.38	-0.31	-0.03
Benton VRT (NE)	-0.35	-0.21	-0.11	0.12	0.23	0.24	0.21	0.22	0.23	-0.13
Seashore (NE)	0.13	0.16	0.21	0.23	0.11	-0.02	0.01	-0.04	-0.05	-0.18
Rey Verbal Learning (NE)	-0.10	-0.15	-0.31	-0.33	-0.25	-0.20	-0.23	-0.27	-0.05	-0.18
Devereux Total	0.14	0.13	0.30	0.35	0.35	0.31	0.26	0.29	0.28	0.01
Conduct	0.12	0.22	0.43	0.42	0.40	0.32	0.24	0.28	0.30	0.09
Attention	0.19	0.16	0.37	0.48*	0.49*	0.49*	0.46*	0.49*	0.48*	0.13
Anxiety	0.28	0.18	0.09	-0.04	-0.09	-0.11	-0.13	-0.11	-0.14	-0.12
Depression	0.13	0.15	0.39	0.45	0.38	0.31	0.22	0.25	0.25	-0.03
Autism	-0.08	-0.19	-0.17	-0.11	-0.06	-0.07	-0.09	-0.05	-0.04	-0.14
Acute Problems	-0.00	0.04	-0.04	-0.01	0.14	0.17	0.21	0.22	0.21	0.02

\*  $P < 0.05$ , one-tailed

**Table 16.** Correlations and partial (P) correlations between annual and concurrent phenylalanine levels, and cognitive and personality variables. (E) denotes executive task, (NE) non-executive task.  $df$  13 throughout.

Table 16 shows correlations and partial correlations between annual cumulative IDCs for Years 1 to 9 and concurrent phe level, and the cognitive and personality measures. Pearson tests were used for normalised scores, Spearman tests for raw scores. Coefficients were non-significant except for positive correlations between Year 4 to Year 9 phe levels and Devereux Attention ( $P < 0.05$ ), and negative correlations between Year 8 and Year 9 phe

levels and Design Fluency ( $P < 0.05$ ). Subsidiary analyses revealed non-significant correlations between mean lifetime phe (Years 1 to 9), mean pre-school phe (Years 1 to 5), mean school-age phe (Years 6 to 9) and all the psychological measures except for Devereux Attention which correlated significantly with mean lifetime phe ( $r = 0.48$ ,  $df = 13$ ,  $P < 0.05$ ) and mean pre-school phe ( $r = 0.49$ ,  $df = 13$ ,  $P < 0.05$ ).

The positive correlations for Devereux Attention were in the theoretically predicted direction. However, regression analyses of Attention and IDCs for Years 4 to 9 isolated one subject whose score was an outlier and whose dietary control worsened between three and four years of age when her parents divorced. When this child was removed from a secondary analysis, coefficients between Attention and phe for Years 4 to 9, lifetime and pre-school phe were non-significant. Thus, the apparent association may have been due to a third factor, marital disharmony, affecting both phe-level and adjustment in a single subject whose data biased the group result. The negative correlations for Design Fluency were also as predicted but, in a matrix of 150 statistics (Table 16), two significant coefficients are well within chance expectations.

Mean Mazes and Digit Span scaled scores were 9.13 and 9.80 respectively. This difference was non-significant and neither test was significantly discrepant from the population norm of 10.

Phe-related differences between executive and control tests provide a critical test of the executive dysfunction hypothesis. However, Pearson coefficients for Digits-Mazes difference scores and historical phe levels were uniformly

non-significant, as was the coefficient for concurrent phe level.

The mean overall *T*-score (population mean 50, SD 10) for the Devereux Scales was 48.47 (SD 7.64). Mean *T*-scores with standard deviations in brackets for the subscales were: Conduct 49.60 (6.96), Attention 51.00 (10.08), Anxiety 46.60 (6.33), Depression 50.73 (14.25), Autism 48.80 (8.26) and Acute Problems 46.93 (3.65). *Z*-test comparisons with population norms revealed none as significantly different.

## **Discussion**

When chance and confounding factors were taken into consideration, results from correlational analyses and normative comparisons failed to provide support for the executive dysfunction hypothesis. Median annual phe levels in the first five years of life between 120 to 810  $\mu\text{mol/l}$ , in the next four years between 120 and 900  $\mu\text{mol/l}$  and concurrent with psychological testing between 141 and 1053  $\mu\text{mol/l}$  were not clearly associated with later executive impairment or behavioural maladjustment.

The Design Fluency test entails the production of novel spatial patterns and may be a task that is especially sensitive to prefrontal cortical impairment (Jones-Gotman & Milner, 1977). Success depends on intact executive functions such as set maintenance, ideational creativity and inhibition of previous responses. Despite this apparent construct validity, Design Fluency performance was not associated with concurrent phe level. The two significant correlations between Design Fluency and historical phe levels are

consistent with possible phe-related non-verbal fluency effects reported by Mazzocco *et al.* (1994) and Griffiths *et al.* (1995), however these isolated findings may be random. At  $P < 0.05$ , in a matrix of 150, eight significant coefficients would be expected by chance. Including the Design Fluency correlations, this was exactly the number found (see Table 16).

Small sample size and the absence of a control group were methodological weaknesses of the study, though these were partly offset by within-subject and population comparisons, none of which indicated either a general or specific neuropsychological or personality disorder. The lack of a positive association between norm-referenced Digits-Mazes difference scores and concurrent phe was particularly counter-theoretical because Diamond's (1994) neurotransmitter model explicitly predicts executive dysfunction as a consequence of elevated phe level, the underlying mechanism proposed being the reduction of prefrontal dopamine by phe.

Arguably, conduct disorder and personal disorganization are manifestations of executive dysfunction in the personality domain. Neither Conduct, Acute Problems nor any of the other Devereux subscale scores differed significantly from population norms. These findings, which imply that the sample was not maladjusted, contrast with those of Smith *et al.* (1988) who reported elevated anxiety and hyperactivity in eight-year-old PKU children. However, in their group, abnormal behaviour was predominantly associated with low IQ and pre-school phe levels above  $600 \mu\text{mol/l}$ . The absence of evidence for deviant personality development in the present study concurs with the findings of Weglage *et al.* (1994).

The neuropsychological test results are also consistent with those of Stemerdink *et al.* (1994) who found neither executive function nor memory deficits in a Dutch classical PKU sample with a comparable mean lifetime phenylalanine level of 381  $\mu\text{mol/l}$ . The present results differ from those of a recent German study by Weglage *et al.*, (1996) who reported a negative correlation between concurrent, but not lifetime, phe and Stroop performance. Though the German sample's yearly IDCs were only slightly higher (476  $\mu\text{mol/l}$  on average), as was its mean concurrent level (583  $\mu\text{mol/l}$ ), the lack of agreement may be accounted for by a considerably greater phe range and higher upper limit (85 to 1709  $\mu\text{mol/l}$ ). These authors also chose digit span as a control (non-executive) measure and, like the present study, found it to be neither developmentally abnormal, nor affected by lifetime or concurrent phe levels.

The present study does not disprove the executive dysfunction theory, but casts doubt on whether subtle impairment of mental ability or personality arises when average phe levels in the pre-school period are kept between 200 and 400  $\mu\text{mol/l}$ . Mean pre-school IDCs for the children in the present sample were close to 360  $\mu\text{mol/l}$  the lower limit of the range considered unsafe by Diamond (1994) and the figure currently recommended in the UK as the upper limit for phe control in the birth to five period (Medical Research Council, 1993a). De Sonnevile *et al.* (1990) found that patients with historical phenylalanine levels below 570  $\mu\text{mol/l}$  performed normally on a task involving sustained attention and response-inhibition, whereas those with levels above did not. This finding and those of the present study add weight to the view that when phe levels in early life are kept well below 600  $\mu\text{mol/l}$ , widespread deleterious effects on later cognitive functioning and adjustment do not occur.

Many studies of executive disorders in treated PKU are based on patients who discontinued treatment early or whose dietary control was exceptionally poor or unreported. In raising the issue, Pennington *et al.* (1985) found executive skill deficits at age 10 in a small sample of six patients. Though treated early, this group had discontinued treatment at six years of age. Nowadays, dietary restriction until 10 is considered desirable (Griffiths *et al.*, 1995) or even for life (Beasley *et al.*, 1994). Welsh *et al.* (1990) reported an inverse relationship between phenylalanine and executive tasks in a group of 11 four-year-olds, but even by this age some patients were returning phenylalanine levels in excess of 600  $\mu\text{mol/l}$  and one was recorded at 1074  $\mu\text{mol/l}$ . Mazzocco *et al.* (1994) found no reliable evidence for deficient executive functioning in early-treated, school-aged PKU children but failed to document historical and concurrent phe levels, making it impossible to translate this finding into recommendations for phe control during treatment. In settling the executive dysfunction issue, it is essential that studies clearly document treatment parameters, such as age of treatment onset and cessation, and quantify phe levels, rather than use categorical descriptions, and that further research is undertaken into the neuropsychological effects of phe levels above and below the 360  $\mu\text{mol/l}$  limit currently recommended in the UK for the pre-school period.

## **Commentary**

In the entire corpus of research on treated PKU, there is no evidence for a threshold effect of phe control on intellectual functioning (Beasley *et al.*,

1994). That is to say, there does not appear to be a clear value for the concentration of phe in blood below which CNS damage does not occur and above which it does. This implies that research studies into how phe control influences behaviour should adopt correlational (or pseudo-correlational designs such as ANOVA) wherever possible. This was the rationale for the correlational approach adopted in the present study. Diamond (1994) eschewed the correlational method in favour of a control group method in her study. However, lifetime phe data for her subjects, either for the group as a whole or for each subgroup, should have been presented and related to test outcomes. By so doing, the influence of historical and task-concurrent phe on executive function could have been determined in the fashion achieved by the present study.

Because of the apparent lack of a threshold effect, any figure that is recommended as a target for phe control at a given age will only be a guideline. What is well-recognized is the dynamic relationship between age and blood phe concentrations. For a given level of circulating phe, neurological damage will be greater the younger the child, the age factor being especially potent in infancy (Berman *et al.*, 1968). Age is presumed to reflect maturity of the CNS, and myelination in particular, so even moderately elevated phe levels will create more disruption to neuronal growth in the early weeks or months than the same levels in later childhood (Hommes, 1991). A possible reason for the disagreement between the findings of the present study and those of Diamond (1994) might lie in the poorer degree of control of her infant subjects or her group as whole in their early years compared with ours. She has acknowledged this possibility (Diamond, 1995) but, in the absence of published phe-control data in her study, the comparability of the

two sets of findings is compromised.

The phe control of the Glasgow patients in infancy and during the pre-school period was probably very close to a realistic optimum, though above current guidelines for the upper limit of the therapeutic range as currently laid down in the UK and Germany of 360  $\mu\text{mol/l}$  for the pre-school years. Weglage *et al.* (1993) have already hinted that the recommended targets may be idealistic, reporting that most patients cannot adhere to the German Paediatric Society's guideline of maintaining phe levels below 360  $\mu\text{mol/l}$  until age 10.

The degree of phe control achieved by Weglage *et al.*'s (1993) sample compared with the subjects in the present study was slightly worse. The table below shows mean indices of dietary control (median phe level per annum) for the Glasgow sample compared with the German sample at years 1, 5 and 10. The discrepancy in the figures provides some justification for the supposition that the Glasgow subjects were reasonably representative of the wider population of strictly treated PKU children. Individual patients who consistently achieve blood phe levels below 360  $\mu\text{mol/l}$  in the early years of life are rare. The sample in the present study only averaged levels around this figure. However, this pattern may be typical and closer to realistic levels of good dietary control than the target figures currently promulgated.

	Year of Life		
	1	5	10
Scotland (Glasgow, $n = 15$ )	348	444	470
Germany (Munster, $n = 34$ )	436	461	648

**Table 17.** Mean Index of Dietary Control, Glasgow versus Munster patients.

The mean IQ in Weglage *et al.*'s sample was 93.6 while that of the present sample was 105.47. Thus, the German PKU group was below the population norm for IQ while the Scottish group was above. The difference in favour of the Glasgow group may be attributable to the overall better dietary control of the Scottish children. Given that the Scottish children had IQs that were at least normal (and, if anything, slightly above) and that no convincing evidence was found either for specific executive deficits or a relationship between cognition generally and historical or concurrent phe, then these findings suggest that a satisfactory degree of dietary control was achieved by the Glasgow patients throughout the treatment age range.

It is a robust finding that, despite pressure from clinicians to the contrary, phe levels in treated PKU rise throughout childhood. Nonetheless, as the risk of CNS damage probably diminishes with age, this inexorable trend may be of little consequence in later childhood. The results of the present study might

be indicating that, so long as low phe levels in early life are maintained, a satisfactory intellectual outcome later on can be obtained. It is highly noteworthy that the average phe level for the group in the present study dipped to about 250  $\mu\text{mol/l}$  in the second year of life (see Figure 15). This may be an important characteristic of the sample that should not be overlooked. It is tempting to speculate that long-term intellectual benefit may be maximized if phe levels in infancy are kept even further below the 360  $\mu\text{mol/l}$  mark presently recommended. The Medical Research Council's Working Party on Phenylketonuria (Medical Research Council, 1993a) advised that this target should be aimed at for the pre-school years but there may be a case for setting a lower target for infancy (*i.e.* years 1 to 2).

In attempting to establish numerical upper limits for the therapeutic range that will yield an acceptable intellectual outcome, it is salutary to recognize that even a figure as low as 360  $\mu\text{mol/l}$  still represents a level of phe circulating in the blood that is six times the norm for the healthy population (Hsia, 1967) and that there is a strong correlation ( $r = 0.91$ ,  $P < 0.01$ ) between blood and brain phe (Pietz *et al.*, 1995). Thus, a potent case exists for reducing phe in infancy on theoretical grounds, because this is the developmental period when the CNS is most vulnerable; and also on practical grounds, because this is when diet can be influenced and manipulated most readily by caretakers. It is interesting that in Germany at the present moment there is a proposal to reduce the recommendation for the first 10 years of life to phe-target levels of between 40 and 240  $\mu\text{mol/l}$  (Burgard, personal communication), though these figures have not yet been formally adopted as national treatment policy.

A criticism of the present study is that no control group was built into the

design. This omission was partly redressed by the use of norm-referenced tests, *i.e.* WISC Digits and Mazes, and the Devereux Scales. However, the addition of a control group matched for age, sex and general ability would have aided interpretation of the findings and would have provided an extra safeguard against making a Type II error. A further criticism is that a measure of decision-making involving sustained attention was not included. In their review of the neuropsychological literature in relation to treated PKU, Waisbren *et al.* (1994) concluded that the kinds of tasks emerging in recent research that seem most sensitive to hyperphenylalaninaemia are those executive functions based on speed of information processing, such as choice reaction time.

With these considerations in mind, my colleagues and I constructed a control-group study to explore the frontal executive hypothesis further by means of a speeded vigilance task which was a modification of the Continuous Performance Test first described by Rosvold *et al.* (1956) and discussed at the start of this chapter.

## Chapter 9

### EXECUTIVE FUNCTION MEASURED BY VARIANTS OF THE CONTINUOUS PERFORMANCE TEST

#### Background to study

When designing this study, we decided to employ the orthodox AX version of the continuous performance test (CPT) as described in Chapter Eight because of the poorer than normal response decrement curve found in PKU by Anderson *et al.* (1969). However, as two subsequent investigations (Realmuto *et al.*, 1986; Brunner & Berry, 1987) based on the original format yielded equivocal data, we incorporated a variant of the CPT which increased attentional demands in an endeavour to increase the sensitivity of the instrument.

Realmuto *et al.* (1986) compared two off-diet adolescent groups on a computer-mediated version of the standard CPT (Klee & Garfinkel, 1983), one of which was given a phe-loaded supplement and the other a placebo. The phe-loaded group's average concentrations increased to above 2100  $\mu\text{mol/l}$  five hours after dosage while the placebo group's remained steady at around 1200  $\mu\text{mol/l}$ . Despite the temporary elevation in circulating phe, the experimental group showed no reliable change in either hits, errors of omission (misses) or commission (false alarms) on the CPT. Nonetheless, reaction times increased by 200 msec in the phe-loaded group and statistical analysis showed a trend towards some loss of information processing speed.

In a CPT study a year later, a similar, though at  $P = 0.14$ , very weak trend was found by Brunner and Berry (1987) on the miss-rate variable of an AX-type computerized task. In their study, off-diet, and therefore hyperphenylalaninaemic, adolescent PKUs made twice as many such errors as sibling controls.

The tenuous link between phe concentrations and performance on the CPT suggested by these studies begged the question whether a true vigilance deficiency exists but was not being revealed by the CPT because of insufficient sensitivity in the instrument. Thus, the standard AX paradigm chosen by Realmuto *et al.* and Brunner and Berry may have been too dull a tool for detecting inattention, especially in older PKU patients, and hence led these authors to make a Type II interpretive error.

Pursuing this line of reasoning, in the study reported below we took the computerized CPT as our measure of executive functioning but, in addition to the conventional AX version, increased task difficulty by the simple expedient of inserting a distractor stimulus (random letter) between the warning signal (letter A) and the target (letter X). This manipulation imposes additional demands on executive functions as the subject not only has to maintain a more complex mental set during execution of the task, but has a higher risk of failing to inhibit erroneous responses. In other words, in the A?X version an impulsive response to the next stimulus after the prime would always be scored as an error, whereas in the orthodox AX version there is a chance of the immediately succeeding stimulus being a target and therefore an impulsive response being registered as a hit rather than an error.

We designated the standard AX version of the CPT the *one-back* test to denote the serial position of the prime in relation to the target and the new A?X version, with its heavier attentional and memory demands, the *two-back* test.

To this author's knowledge, the two-back paradigm has not been used as a neuropsychological methodology in PKU research, nor was it mentioned by Corkum and Siegel (1993) in their review of task variables in thirteen CPT studies of various issues and populations.

The lack of a control group in the study reported in Chapter Eight and the lack of a correlational analysis in Diamond's (1994) seminal study are methodological deficiencies. These were remedied in the design of the study that forms the core of this chapter by testing healthy, matched controls in order to illuminate the effect of diagnostic category on one-back and two-back CPT performance. Also, within the PKU group, dietary control factors were quantified and related to outcome on the CPT in an endeavour to clarify whether the therapeutic range currently recommended in the United Kingdom (Medical Research Council, 1993b) confers protection against executive disorder.

## **Introduction**

Uncertainty prevails about whether subtle cognitive deficits are associated with treated phenylketonuria (PKU, McKusick 26160) in childhood. While it is universally accepted that loss of IQ can be minimized by implementation of

dietary prophylaxis as early as possible in the neonatal period (Barclay & Walton, 1988; Legido *et al.*, 1993), questions remain about the optimal age for cessation of the low phenylalanine diet and the maximum level of circulating phenylalanine (phe) allowable at different ages during treatment (Smith *et al.*, 1990; Editorial, 1991; Medical Research Council, 1993a). Following the introduction of mass newborn screening for PKU in the late sixties (Guthrie, 1996), an upper limit for the therapeutic range of 600  $\mu\text{mol/l}$  was widely accepted, especially in the United States (Williamson *et al.*, 1981). However, this figure has been challenged by Diamond (1994) who concluded that it is too high to prevent disorders of executive function in infants, toddlers and young children and that such groups are especially vulnerable because executive impairments may not be detectable by traditional intelligence tests.

Executive functions are higher, goal-directed mental activities that are organizational and supervisory in nature and heavily dependent on good attentional control. They entail planning, selecting or maintaining complex behavioural routines, often in novel contexts (Kelly *et al.*, 1996). Pennington and Ozonoff (1996) argued that executive functions in children are mediated by the prefrontal cortex of the brain as they are in adults and hypothesized that, where frank neurological damage is absent, deficits in executive functions may reflect dopamine depletion in this area. Welsh (1996) proposed that executive functions may be specifically at risk of impairment in treated PKU. The theoretical rationale she put forward for this view was that reduced tyrosine availability resulting from hepatic phenylalanine hydroxylase deficiency adversely affects dopaminergic neurotransmitter systems and, as the prefrontal dopamine system is especially sensitive to decreased tyrosine, executive dysfunctions would result. Accordingly, she argued, impairment of

executive functions represents a cognitive marker for tyrosine imbalance and would be seen in neuropsychological tasks involving planning, response inhibition and flexibility of thought and action. She drew upon empirical evidence from Davis *et al.* (1986), Pennington *et al.* (1985) and Welsh *et al.* (1990) to support the idea that executive functions may be impaired at least in young (*i.e.* preschool) children with treated PKU.

We tested the executive dysfunction hypothesis by means of a control-group study of primary school-age children with classical PKU who were treated early and whose mean phe concentrations for the preschool years remained close to 300  $\mu\text{mol/l}$ . Our aims were threefold: a) to establish whether executive skill was specifically deficient in this group compared with healthy children, b) to determine whether preschool or lifetime phe predicted later executive capability, and c) to investigate the relationship between task-concurrent phe and executive function. We chose the Continuous Performance Test (CPT) as the dependent experimental measure as it fulfils the criteria of a typical executive function task in having the intrinsic demand characteristics of sustained attention, set maintenance, decision-making and inhibition of competing but erroneous responses (Corkum & Siegel, 1993; Benton, 1994).

The prototype version of the Continuous Performance Test (Rosvold *et al.*, 1956) is a vigilance task that requires the subject to respond rapidly to the target letter X when immediately preceded by the priming letter A. Conventionally, the AX letter-combination is randomly located in a series of distractor letters. We designated this the one-back version to denote the position of the priming stimulus in relation to the target. In order to increase

the attentional demands of the task , we also employed a two-back version in which a distractor letter was interpolated between the priming and target stimuli. Clearly, n-back variations of this task are possible but, given that some of the children were only five years of age, we limited the priming letter to two serial positions back from the target to avoid possible floor effects.

## **Method**

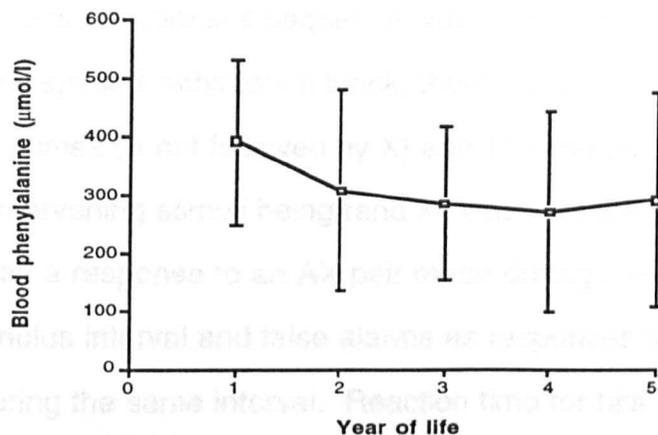
### *Subjects*

Eleven children (six boys and five girls) with Type 1 or 'classical' PKU (Guttler, 1980) made up the experimental sample. They were identified from the West of Scotland PKU Register as having had pretreatment phe levels in excess of  $1200\mu\text{mol/l}$  and as having been treated early (median day of life = 13, range 1 to 47) and continuously until the time of the study. None of the PKU children was mentally handicapped (mean IQ = 108.71, SD 12.30). Individual PKU subjects were paired with healthy controls, seven boys and four girls, from the normal primary school population on chronological age and receptive vocabulary, the latter being considered a non-executive ability that correlates highly with global IQ (Sattler, 1974). The median age of the PKU group was 8.83 years (range 5.11 to 11.92) and that for the control group also 8.83 (range 6.17 to 12.75). The age distributions did not differ significantly at the 0.05 alpha level ( $W = 29$ ,  $N = 10$ ,  $P = \text{n.s.}$ , Wilcoxon). Receptive vocabulary was measured by the British Picture Vocabulary Test (Dunn *et al.*, 1982), the standard score mean for the PKU group being 104.64 (SD 14.62) and that for the controls 101.45 (SD 15.40). This difference was not statistically significant

( $t = 1.47$ ,  $df = 10$ ,  $P = n.s.$ , correlated). Postcode-derived deprivation categories were used as an index of social class (Carstairs & Morris, 1992). The modal value on a 1 to 7 scale for the PKU group was 4 and that for the controls 3, 1 meaning least deprived. The groups were thus matched on size, chronological age and receptive vocabulary, and very closely matched on sex and social class. Pairing on age and vocabulary was achieved at the expense of an equal sex ratio, but gender has not been consistently demonstrated as a factor underlying intellectual outcome in treated PKU (Koff *et al.*, 1979).

### *Measurement of phenylalanine*

Historical and task-concurrent plasma phenylalanine concentrations were measured by high pressure liquid chromatography. For individual PKU children, the annual median value was taken as the index of dietary control (Rupp & Burgard, 1995). Figure 16 shows mean indices with standard deviations for the first five years of life for the group as a whole. The mean preschool (birth to age five) phe-level for the group was 342  $\mu\text{mol/l}$  (SD 126), the mean lifetime level 341  $\mu\text{mol/l}$  (SD 125) and the mean task-concurrent level 388  $\mu\text{mol/l}$  (SD 127).



**Figure 16.** Historical phe-level profile for PKU sample. Mean phe concentrations for first five years of life. Error bars are standard deviations.

#### *Executive function task*

A modification of the AX format of the CPT was adopted as a measure of executive function (Rosvold *et al.*, 1956). An Acorn Archimedes 310 personal computer was programmed in BASIC to deliver the stimuli which were upper-case letters of the alphabet, 7 mm high, placed centrally in a square frame with 38 mm sides. The frame was permanently displayed in the centre of the monitor screen throughout the test. Subjects sat at a distance of approximately 50 cm from the screen. The onset-offset duration for all stimuli was 500 msec and the interstimulus interval 1000 msec.

In the one-back version of the task, there were 560 stimuli. A target was defined as the letter X immediately preceded by the priming letter A. There were 80 targets in total, the subject being required to make a keyboard

response as quickly as possible to the presentation of an X contingent upon an A. The entire stimulus sequence was subdivided into four blocks of 140. Randomly spaced within each block, there were 20 targets (AX combinations), 10 false primes (A not followed by X) and 10 false targets (X not preceded by A), the intervening stimuli being random letters of the alphabet. A hit was defined as a response to an AX pair made during the immediately succeeding inter-stimulus interval and false alarms as responses to single A and X stimuli made during the same interval. Reaction time for hits was measured in centiseconds from the onset of the X stimulus in an AX pair to the depression of the space bar on the computer keyboard.

The parameters of the two-back task were identical to the one-back except that a distractor letter was interposed between the A and the X, the target combination becoming A?X instead of AX. Thus, the priming letter was two-back in the sequence from the target letter. Each task took 14 minutes from beginning to end. Losier *et al.* (1996), in a review of 26 studies employing the CPT with children, found the following modal values and ranges (in brackets): stimulus duration 200 msec (50-800), inter-stimulus interval 1500 msec (800-4000), time on task 12 minutes (9-27) and number of trials 350 (300-600). Thus, the settings chosen in the present study were within conventional limits and generally close to the mode. The number of trials (*i.e.* stimulus presentations) was the one parameter deliberately set at an above average level. The rationale for this decision was to raise the attentional demands of both versions of the task and, by compounding the extra difficulty of the two-back task with already protracted attentional maintenance, thoroughly test for weaknesses in vigilance performance. The program automatically recorded on-task performance and reported hit and false alarm totals and mean

reaction times for the one-back and two-back tests separately as well as a block-by-block breakdown.

### *Non-executive function task*

Forward recall of digits was selected as a non-executive, discriminative task on theoretical and empirical grounds. Immediate repetition of digit strings is a function strongly associated with posterior rather than prefrontal cortex mediation (Walsh, 1978) and short-term memory deficits for phonological information are not characteristic of treated PKU (Waisbren *et al.*, 1994). The test comprised two- to nine-long strings of the digits 1 to 9 ordered randomly. There were five trials at each level, each containing a different digit combination, making 40 in total. Individual digits were presented at a rate of one per second by means of an audio tape-recorder. An untimed oral response was required from subjects. A simple binary scoring scheme for success and failure was imposed. The discontinuation criterion was three or more failures within a five-trial level. The score recorded for each subject was the total number of trials correct out of 40.

### *Design and procedure*

The study conformed to a matched groups design. The test order for individual subjects was standardized, the one-back CPT being given first followed by the British Picture Vocabulary Scales, the two-back CPT and lastly recall of digits. Data were collected from subjects during a single session which had rest breaks between but not within test items. As many pretest practice trials were given for both versions of the CPT as were necessary to

convey the task requirements to individual subjects.

Approval for the study was obtained from the Glasgow Royal Hospital for Sick Children's Ethics Committee and permission granted by the local education authority to test the control children. Informed consent from the children and their carers was obtained by means of an explanatory letter.

## Results

The dependent variables for the one-back and two-back CPT were hit rate, number of false alarms and mean reaction time. Though the CPT program registered reaction times in centiseconds, these are reported below in milliseconds for ease of comprehension. Results for each version were analyzed by two-way ANOVA (Minitab) for group and time effects.

### *One-back continuous performance test*

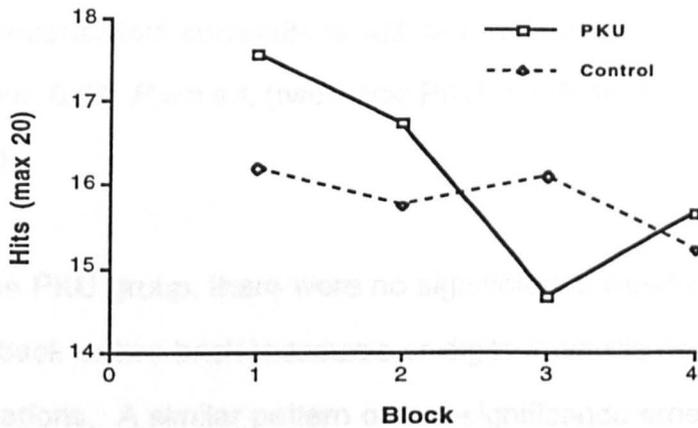
The overall mean hit rate out of 80 for the one-back CPT was 72.18 (SD 7.17) for the PKU group and 69.70 (SD 10.5) for the controls. The group effect was non-significant ( $F = 0.41$ ,  $df\ 1/20$ ,  $P = n.s.$ ) as were the time-on-task ( $F = 0.65$ ,  $df\ 3/20$ ,  $P = n.s.$ ) and interaction effects ( $F = 1.47$ ,  $df\ 3/20$ ,  $P = n.s.$ ). At 8.36 (SD 7.24), the overall mean false alarm frequency for the PKU children was less than that of the controls at 15.80 (SD 19.50), but the group difference was not statistically significant ( $F = 1.42$ ,  $df\ 1/20$ ,  $P = n.s.$ ) and neither was the time factor ( $F = 0.13$ ,  $df\ 3/20$ ,  $P = n.s.$ ), nor the interaction term ( $F = 0.55$ ,  $df\ 3/20$ ,  $P = n.s.$ ). The overall mean reaction time in milliseconds for the PKU group was

499 (SD 81) and that for the control group 462 (SD 92). Again the group effect was non-significant ( $F = 1.00$ ,  $df\ 1/20$ ,  $P = n.s.$ ) and no significant effects were found for time ( $F = 0.20$ ,  $df\ 3/20$ ,  $P = n.s.$ ) or the interaction ( $F = 0.29$ ,  $df\ 3/20$ ,  $P = n.s.$ ).

### *Two-back continuous performance test*

On the two-back version, the overall mean hit rates for the PKU and control groups were very similar at 64.55 (SD 7.54) and 63.20 (SD 16.90) respectively, and the group effect was not statistically significant ( $F = 0.81$ ,  $df\ 1/20$ ,  $P = n.s.$ ). However, a highly significant time-on-task effect emerged ( $F = 7.22$ ,  $df\ 3/20$ ,  $P < 0.001$ ), indicating a decrement in performance for both groups as the task proceeded. The interaction term was also significant ( $F = 5.50$ ,  $df\ 3/20$ ,  $P < 0.01$ ). Mean block-by-block changes in detection of target stimuli for the two groups are presented in Figure 17 which shows a precipitate decline in performance on block 3 followed by recovery during block 4 for the PKU group, while the opposite pattern was displayed by the control children. No further significant results were obtained from the two-back test. The overall mean false alarm rate for the PKU group was 9.64 (SD 5.89) which was considerably less than the controls' of 25.8 (SD 33.1), though the difference was not significant ( $F = 1.87$ ,  $df\ 1/20$ ,  $P = n.s.$ ). No significant effects on false alarms were found for time ( $F = 1.73$ ,  $df\ 3/20$ ,  $P = n.s.$ ) or the interaction ( $F = 0.91$ ,  $df\ 3/20$ ,  $P = n.s.$ ). The overall mean reaction time in milliseconds for the PKU children was 509 (SD 72) and that for the control children 481 (SD 14), but the difference was non-significant ( $F = 0.34$ ,  $df\ 1/20$ ,  $P = n.s.$ ). Similarly, reaction times were not significantly affected by time on task ( $F = 0.18$ ,  $df\ 3/20$ ,  $P = n.s.$ ) or the interaction between group and time ( $F$

= 0.42, df 3/20,  $P = \text{n.s.}$ ).



**Figure 17.** Mean hit rate per consecutive block of 140 trials on the two-back CPT. Illustrates significant performance decrement and group x time interaction. Between group factor was non-significant.

#### *Digits forwards test*

The mean score out of a maximum of 40 on the recall of digits test for the PKU group was 18.45 (SD 4.70) and for the control group 20.91 (SD 4.64). The difference between the distributions was not significant ( $t = 1.91$ , df 10,  $P = \text{n.s.}$ , related).

#### *Discussion*

##### *Correlational analyses*

The results from this small-sample study failed to support the view that As expected, within both groups there were significant correlations at  $P < 0.05$  (df 9, Pearson) between chronological age and one-back hits (PKU,  $r = 0.62$ ; control,  $r = 0.77$ ), one-back reaction time (PKU,  $r = -0.68$ ; control,  $r = -0.75$ ),

two-back hits (PKU,  $r = 0.86$ ; control,  $r = 0.69$ ), two-back reaction time (PKU,  $r = -0.75$ ; control,  $r = -0.60$ ) and digit recall (PKU,  $r = 0.81$ ; control,  $r = 0.65$ ), thus vindicating the use of a matched controls design. Only false alarms showed inconsistent correlations with age (one-back: PKU,  $r = -0.66$ ,  $P < 0.05$ ; control,  $r = -0.25$ ,  $P = \text{n.s.}$ ), (two-back: PKU,  $r = -0.48$ ,  $P = \text{n.s.}$ ; control,  $r = -0.64$ ,  $P < 0.05$ ).

Within the PKU group, there were no significant correlations between any of the one-back or two-back measures or digits forwards and mean lifetime phe concentrations. A similar pattern of non-significance emerged for correlations between the outcome variables and mean phe concentrations for the first five years of life and task-concurrent levels. Mean lifetime and concurrent phe level were highly positively correlated ( $r = 0.89$ ,  $df\ 9$ ,  $P < 0.001$ ). Interestingly, in the PKU group, only two-back hits (as opposed to false alarms and reaction time) and digit recall correlated significantly ( $r = 0.84$ ,  $df\ 9$ ,  $P < 0.01$ ) and in the control group only one-back hits and digit recall ( $r = 0.72$ ,  $df\ 9$ ,  $P < 0.05$ ), this inconsistent pattern of inter-test correlation suggesting poor concurrent validity and providing some evidence that the experimental and control tasks represented different cognitive domains.

## **Discussion**

The results from this small-sample study failed to support the view that continuously treated PKU in middle childhood is characterized by specific executive dysfunction, at least as gauged by the ability to sustain attention and decision-making, when average phe concentrations in the preschool

period are kept around the 300  $\mu\text{mol/l}$  mark. Neither the one-back nor the two-back versions of the CPT differentiated children with classical PKU who were treated early from healthy children of similar age and verbal ability in terms of target detection, inhibition of incorrect responses or speed of reaction to targets. Both CPT variants required prolonged periods of concentration and the case for the two-back version being the more taxing was strengthened by the finding of a significant vigilance decrement curve on this measure. Despite the increased attentional demands of the two-back test, no between-group differences emerged. Thus, it appeared that the on-diet PKU children could concentrate or, conversely, resist distraction as well as their non-PKU counterparts.

The significant interaction term in the analysis of the two-back, hit-rate data could not be ascribed to the above average phe levels that typify even treated PKU. (The non-PKU fasting norm falls within the 35 to 100  $\mu\text{mol/l}$  range (Smith & Brenton, 1995).) Figure 17 shows the means conforming to a crossover rather than a divergence pattern centring on block 3. While the PKU group's hit rate fell suddenly during this block compared with both the within-group trend established during blocks 1 and 2 and the performance of the controls, it recovered and indeed surpassed that of the controls during block 4.

The significant time effect found for hit rates on the two-back version of the CPT for both groups (Figure 17) clearly demonstrates the kind of temporal performance decrement curve customarily found in vigilance tasks (Mackworth, 1969). From this evidence, it is tempting to conclude that the two-back version of the CPT may be more sensitive to progressive failure of

attentional mechanisms than the one-back version and hence should be used whenever there is a need to explore subtle impairments of sustained attention. However, as the two-back version was confounded with test order and may have been more subject to general fatigue effects, the general superiority of this format as an investigative tool needs to be further examined.

The earliest example of the CPT being incorporated into PKU research methodology was a study by Anderson *et al.* (1969) who reported poorer hit rates and more errors than matched controls in a group of 11 PKU children. However, all their subjects were diagnosed before the advent of mass neonatal screening and their sample was extremely heterogeneous for treatment variables and IQ. Compared with our homogeneous group, two had never been on diet; three commenced later than one, six and eleven years of age respectively; and, with only one exception, the remainder began when they were more than two months old. The sample mean IQ was 77 and two subjects were severely mentally handicapped.

The present study concurs with those of Mazzocco *et al.* (1994) and Stemerding *et al.* (1994) in finding no significant executive function deficits in school-age subjects with early treated PKU. It extends these authors' work by relating performance on executive and non-executive tasks to lifetime, preschool and concurrent phe concentrations. As no significant correlations emerged between either historical or recent phe levels and executive or indeed non-executive function, it is a reasonable conclusion that the range of dietary control observed by the patients in our sample may have been protective against both global and specific cognitive deficits.

The results vindicate the policy adopted in the United Kingdom some years ago of aiming to reduce phe concentrations in the preschool period to below 360  $\mu\text{mol/l}$  (Medical Research Council, 1993b). At 342  $\mu\text{mol/l}$ , our subjects succeeded in maintaining their preschool average within this guideline. Though the group's standard deviation for the period was 126  $\mu\text{mol/l}$ , this degree of variance did not yield a phe-control effect in outcome correlations.

Poor self-management and conduct disorder may exemplify executive dysfunction in the personality as opposed to the cognitive domain.

Nonetheless, Weglage *et al.* (1996a) found no correlation between preschool phe and scores on a questionnaire measure of behavioural adjustment at 10 years of age in a group of 58 treated PKU children. The German children's average indices of dietary control between birth and five were close to ours at around 360  $\mu\text{mol/l}$ . Thus, Weglage *et al.*'s negative result regarding personality functioning provides further evidence for 360  $\mu\text{mol/l}$  being not only an attainable and realistic preschool target for phe control but one which does not appear to have demonstrable adverse consequences in the short-term.

In contrast to both the present study and their personality findings, Weglage *et al.* (1996b) reported significantly poor attention in 11-year-old, PKU subjects who were treated early compared with matched controls. These authors also found significant correlations between concurrent phe and attentional measures. In keeping with our data, they found no group effect on digit recall or associations between historical phe control and attention. As measures of attention, Weglage and his colleagues used the Stroop Task and the Test-d-2. The contradictory findings of the two studies may relate to the differential sensitivity of the instruments to phe. However, the older German

children had slightly poorer historical and considerably worse concurrent phe control (see Table 18). As average pre-school and school-age phe levels for our subjects fell within current UK guidelines of 360 and 480  $\mu\text{mol/l}$  respectively and no evidence for executive impairments was found, the case for the suitability of these benchmarks is strengthened.

		Mean	SD	Range
<b>Lifetime</b>	Present study	341	125	224-600
	Weglage <i>et al.</i>	476	144	279-818
<b>Concurrent</b>	Present study	388	127	193-562
	Weglage <i>et al.</i>	583	377	85-1709

**Table 18.** Treatment parameters of patients in the present study ( $n = 11$ , median CA = 8.83 years) compared with those studied by Weglage *et al.* (1996b) ( $n = 20$ , median CA = 10.92). Data are phe concentrations in  $\mu\text{mol/l}$ .

Knowledge is presently scant about the effects of longer-term exposure to elevated phe concentrations on executive skill in adolescence and beyond (Zetterstrom, 1995). Nonetheless, in a control group study, Griffiths *et al.* (1995) found no evidence for deficits in measures of executive function such as letter cancellation, verbal fluency and three-dimensional maze learning in patients who terminated diet abruptly at age 10 and who were

hyperphenylalaninaemic for three or more years thereafter.

Research into the executive dysfunction question would be served by further studies of infant PKU subjects in an endeavour to replicate Diamond's (1994) findings and animal studies are needed to explore the prefrontal dopamine model (McDowell, 1996). All studies of treated patients should integrate results from outcome measures with historical and concurrent phe levels. Further research using comparable measures on progressive age groups with well-defined treatment histories is needed before a definitive statement can be made about the executive dysfunction issue and optimal phe levels at different stages in the life cycle.

## **Commentary**

The term 'executive function' is of fairly recent origin, its provenance being the work of Stuss and Benson (1985). However, the concept dates back to the 19th century and localizationist theories about the role of the frontal lobes in cognition. Nowadays, 'executive' and 'frontal lobe' function are viewed as interchangeable, though it is important to recognize the metaphorical character of the 'frontal lobe' description (Pennington & Ozonoff, 1996). Far from being silent, the prefrontal cortex is now regarded as the seat of man's highest intellectual feats and the centre of creative, adaptive and purposeful behaviour (Hebb, 1945). Its activities are probably most closely linked to what Cattell (1971) described as *fluid* as opposed to *crystallized* intelligence. In essence, the cortex anterior to the central sulcus is responsible for the application of skill and knowledge, in contrast to the cortex posterior to this

division which is primarily devoted to input and storage of information. The acquisition of knowledge involves interaction between the two functionally distinct anterior and posterior regions (and, incidentally, provides the basis of a neuropsychological model for Piaget's developmental theory of epistemology).

Because of the organizational and metacognitive nature of executive functions, they remain not only difficult to define in the abstract but even more elusive to capture psychometrically. For instance, operationalizing executive constructs like creativity, foresight and flexibility of thought as tasks with quantifiable outcomes poses special problems compared with, say, measuring non-executive functions such as defining words or completing pictures. In general, assessment is presently better served by instruments measuring crystallized rather than fluid intelligence (Duncan, 1995). One reason is historical and to do with the dismissal of frontal function as being of minor significance by Hebb and Penfield in the forties, but principally the imbalance is due to controversy both now and in the past about what constitutes a valid measure of prefrontal activity.

From studies of impulsivity and perseveration in frontally damaged patients, some consensus has emerged that the ability to maintain mental set over time, resist distraction and inhibit maladaptive responses is a hallmark of many executive functions (Welsh & Pennington, 1988). Thus, the Continuous Performance Test, which was originally devised by Rosvold and his colleagues to discriminate neuropathological from non-organic disorder, has quickly found a central place in armoury of measures purporting to tap executive disorders. Because of its vigilance format, the CPT is nowadays

seen as a measure of sustained rather than selective attention and having a particular value in the diagnosis and monitoring of attention deficit disorder (Schachar *et al.*, 1993).

Neither of the studies of executive dysfunction by Anderson *et al.* (1969) and Diamond (1994) furnished sufficiently detailed data about phe levels in their PKU subjects to enable an analysis of whether a threshold value, either historical or concurrent, was identifiable above which attention and impulsivity were affected and below which were not. Thus the question remains largely open as to whether, when a certain level of circulating phe is reached, specific executive disorders are potentiated and, if so, what that precise point might be.

The two-back version of the CPT promises to be a useful tool for exploring this question. Indeed, with flexible programmable testing systems such as SuperLab (Haxby *et al.*, 1993), not only can the effect of altering stimulus variables such as duration or interval be studied but also that of increasing task difficulty. The elicitation of a vigilance decrement curve during a sustained attention task such as the CPT is usually taken as an index of validity. Task variables such as stimulus duration, inter-stimulus interval, event rate and whether targets are sequenced rather than single have a powerful influence on slope of the curve (Schacher *et al.*, 1988). In the present study, it is particularly noteworthy that only the two-back format yielded a significant decrement curve in the time allotted for testing. This begs the question whether a *three-back* paradigm would constitute an even more penetrating methodology for exploring the vigilance aspect of executive function in relation to phe level, though the likely interaction between task

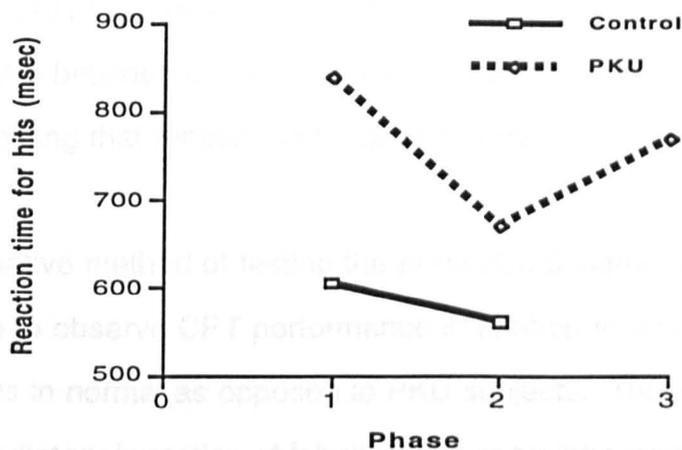
difficulty and developmental level would have to be carefully controlled and floor effects avoided.

The findings of the present study, with its well-controlled subjects, may be seen as corroborating Diamond's (1994) view, but they support more the corollary of her proposition, namely, that phe levels below 360  $\mu\text{mol/l}$  are *not* harmful, than the proposition itself which states that those above are. A natural extension to the work described here would be a study based on a sample with poorer control, preferably in the range specified by Diamond as harmful of 360 to 600  $\mu\text{mol/l}$ . To increase task difficulty and therefore the chances of limit-testing executive skills, a two-back or even three-back variant of the CPT could be employed. The clear hypotheses could then be tested that the performance decrement curves of healthy controls and PKU children with lifetime phe concentrations between 120 and 360  $\mu\text{mol/l}$  would be indistinguishable, but that those of children with concentrations between 360 and 600  $\mu\text{mol/l}$  would show significant impairment.

Another way of increasing vigilance task difficulty is to require subjects to respond selectively to targets and distractors. Thus, in addition to pressing a button when a target appears, they have to press another button in response to every distractor. This paradigm was adopted by de Sonneville who measured sustained attention by means of a Dot Pattern task in which the target stimulus was four dots and the distractors three and five dots. The positions of all the dots were random. Subjects had to press one of two buttons depending on their decision about the category to which the signals belonged and performance and reaction time were recorded automatically.

Twenty-two eight-year-olds with PKU achieved the same accuracy level as controls on this task but were significantly slower (de Sonneville *et al.*, 1990). However, within the PKU group, significantly worse accuracy was shown by those with concurrent phe concentrations above 570  $\mu\text{mol/l}$  compared with those below whose performance was normal. These findings provided some support for Diamond's view that phe concentrations in excess of 600  $\mu\text{mol/l}$  may be harmful but the study was marred by methodological flaws such as a failure to describe treatment variables fully, inappropriate statistical analyses (*t*-tests instead of ANOVA) and a mismatch between the PKU sample and controls on Performance IQ.

De Sonneville's Dot Pattern task was used as the investigative instrument in a subsequent experimental study in which 14 off-diet adults with PKU were returned to diet and finally taken off again (Schmidt *et al.*, 1994). Figure 18 is adapted from the data presented in this report. The authors claimed covariance between on-task phe and attention, reaction times on the dot detection test decreasing when phe was lower, then increasing when phe returned to baseline levels. The difficulty with this interpretation is that, potentially, practice and familiarity were confounded with the high to low phe transition (phases 1 to 2) and fatigue and boredom with the low phe to high phe transition (phases 2 to 3). The simple expedient of testing the controls a third time in parallel with the PKU group under the high phe (return to baseline) condition would have helped resolve the interpretive problem. However, these data are curiously absent in the results section of the report, the reaction-time means for hits being presented as in Figure 18.



**Figure 18.** Mean reaction times for hits on de Sonneville's Dot Pattern task as a function of dietary phe manipulation. For PKU sample, Phase 1 is high phe, Phase 2 low phe, and Phase 3 high phe. Adapted from Schmidt *et al.* (1994). Illustrates absent datum point for controls during phase 3.

Future research into the executive dysfunction issue would be better served by greater standardization of outcome measures and more careful definition of sample characteristics especially in relation to treatment variables. There is a great need for studies to be constructed from samples stratified for historical and concurrent phe control, although achieving large enough numbers for statistical analysis might necessitate collaboration and multi-centre data collection.

The study described in this chapter and that reported by Diamond (1994) illustrate that executive skill may vary according to lifetime or task-concurrent phe level and that the findings from these two studies are not necessarily

contradictory but may reflect subtle neuropsychological changes that occur with progressively poor phe control. A threshold effect somewhere between 300 and 600  $\mu\text{mol/l}$  is unlikely on theoretical grounds. A linear or curvilinear relationship between circulating phe and executive impairment is more likely, but something that remains to be clarified empirically.

An alternative method of testing the prefrontal dopamine biosynthesis model would be to observe CPT performance in relation to artificially elevated blood phe levels in normal as opposed to PKU subjects. These could be provoked either by dietary ingestion of labelled phe or by intravenous infusion. If hyperphenylalaninaemia in humans generally depresses dopamine synthesis in the prefrontal area of the brain, such a phenomenon should occur in non-PKU subjects when hepatic hydroxylation of phe to tyr is challenged to the point where the system becomes overloaded. Overloading would produce high blood phe concentrations. Though the blood-brain barrier diminishes cortical phe levels (Pietz *et al.*, 1995), the effects of overload should nonetheless be detectable in performance on measures such as the CPT if it is the case that, as Guttler and Lou (1986) proposed, dopaminergic neurone systems are functionally and temporarily compromised.

To the author's knowledge, such a study has not been conducted. However, Lou *et al.* (1987) adopted the opposite strategy of using large doses of tyr to counteract the hypothesized dampening effects of phe on prefrontal neurotransmitter metabolism in PKU. The results of tyr supplementation on continuous reaction times were equivocal with only the slowest showing improvement. The robustness of such an effect on neuropsychological performance remains seriously in doubt as it was not tested statistically by

Lou *et al.*, nor has it been independently replicated.

A further promising angle on the executive dysfunction question is the work, at present minimal, on late treatment of adult patients with PKU. Members of this population are mainly those mentally retarded adults, now generally into middle or old age, who were born before mass screening was initiated, or even before the possibility of treatment was demonstrated. They were never treated in infancy and the disease followed its natural course, resulting in severe intellectual impairment and usually institutionalization. A handful of case studies have been conducted to explore whether reducing dietary phe in untreated PKU can lead to behavioural improvement. There is no suggestion that introducing diet many years into the life cycle of the individual can reverse intellectual loss, but the dopamine-executive theory would suggest that lowering phe might dynamically ameliorate conduct disorder and poor impulse control.

Guttler and Lou (1986) argued that, in untreated PKU, high phe concentrations decrease tryptophan in the CNS and interfere with the conversion of tyrosine into dopamine and norepinephrine, but that the temporal dynamics of these processes are short and reversible.

Administering treatment late in PKU provides a test of this view and also of the executive dysfunction theory. The high level of comorbidity between attention deficit, conduct disorder and hyperkinesis is currently regraded by some as a marker for severe executive dysfunction (Moffit, 1993). Although frank attention deficit hyperactivity disorder is not commonly reported as a clinical feature of treated PKU, even when dietary control is weak, in the untreated, mentally handicapped patient poor self-control, oppositional behaviour and

impulsivity are frequent symptoms that may be attributable to hyperphenylalaninaemia and resultant deranged metabolism in prefrontal cortex. If reducing phe by dietary management mitigates these behaviours in untreated PKU, such a finding would provide indirect support for a frontal neurotransmission mechanism.

At the time of writing, the balance of empirical evidence on this issue is, if anything, slightly against late treatment having beneficial behavioural effects. Harvey and Kirk (1995) reduced blood phe in a 30-year-old untreated male patient to around 300  $\mu\text{mol/l}$  by dietary manipulation from a pre-diet level of 1900  $\mu\text{mol/l}$ . They reported no systematic change for the better in behaviour during the on-diet period. Only slight improvement occurred following dietary inception, which may have been a Hawthorne rather than a phe effect. The case for social and material reinforcement being the active factor rather than phe was strengthened by the additional observation that the patient's behaviour also improved on return to his liberalized, and presumably more varied and palatable, normal diet.

Three independent reports cited by Harvey and Kirk showed diet-related improvement in individual patients, while a further four, in addition to their own, did not. In the untreated patient, gross rather than subtle executive dysfunction is the norm. If research consistently shows that drastic reduction of blood phe in the *untreated* PKU population consistently fails to produce observable behavioural improvement, then the status of the metabolic theory of executive dysfunction must become severely diminished and its fundamental assumptions called into question.

## **Chapter 10**

### **OVERVIEW: TREATMENT AND POLICY**

Gene therapy or enzyme replacement therapy are theoretical alternatives to the current costly and time- and effort-consuming dietary method for preventing mental handicap in PKU. However, while the hope of developing a substitute for the missing enzyme was expressed over twenty years ago (Gearheart & Litton, 1975), it has yet to be fulfilled. The practicalities of enzyme replacement by means of orally administered medication have so far foundered on the problems of creating a satisfactory transport system and animal studies of techniques involving injection and implantation of active enzyme have failed (Danks & Cotton, 1987). Gene insertion therapy is now a real possibility and would represent a 'one-shot' or 'magic bullet' cure. The first disease being treated with gene therapy is adenosine deaminase deficiency and cystic fibrosis is another likely candidate (Batshaw & Perret, 1992). However, the prospects for genetic remediation of the PAH deficiency in PKU are as yet impossible to gauge (Kent, 1994). Liver transplantation has never been seriously mooted as a treatment for PKU for ethical and practical reasons. There can be little moral justification for replacing a healthy organ whose functions aside from the manufacture of PAH are normal and in which the effects of PAH deficiency can be successfully circumvented by dietary adjustment. Furthermore, successful transplantation depends on availability and compatibility of donor organs and rejection of a vital organ such as the liver carries with it an unacceptable risk to the life of the patient.

For the foreseeable future, then, dietary prophylaxis, with all its demands for intra-familial social and behavioural reorganization, will remain the only practicable treatment option in severe PKU. In the 1950s, Bickel's method was truly revolutionary and the first instance in the history of medicine of the natural course of a neuropathic disease leading to mental retardation being halted. While it is broadly true to say that, as a therapeutic procedure, restriction of dietary phe is effective, refinement of the parameters of treatment must remain the goal of research. Taking an overview of the present state of knowledge about treatment factors in childhood PKU, it appears that dietary restriction of phe over the first decade of life produces diminishing intellectual benefit while the personal and social costs to the individual of adhering to the diet increase concomitantly. As demonstrated by meta-analyses such as those by Berman *et al.* (1968) and Hanley *et al.* (1971) on the relationship between age of treatment commencement and intellectual outcome, it is clear that the neurotoxic effects of phe are particularly savage in infancy and it is now universally accepted that treatment should begin as soon as is practicable in the neonatal period. However, there are signs that the issue as to when treatment may be safely terminated will become increasingly enmeshed with the issue of whether the harshness of the regimen justifies the cognitive outcome in the teenage years and beyond. Hence, risk-benefit analysis has more-and-more become the perspective on management of the adolescent and adult in recent years. This more complex question about the balance between the demands of continuing treatment and the profitability of so doing will foreseeably be the one that will exercise clinical research in the future. Data from across the life spans of individuals with PKU will undoubtedly be required to enable a definitive statement to be made about whether treatment cessation at any age is safe for the integrity of the CNS.

The studies in this volume were designed to illuminate the effects of the two remaining 'pure' treatment factors - age of cessation and degree of control - in those age groups from which satisfactory research samples can now be constructed, namely, school-age children, adolescents and young adults. As any derangement of mental and motor functioning at these developmental levels is likely to be subtle rather than gross (Guttler & Lou, 1986), a neuropsychological approach was adopted. The neuropsychological approach centres on the use of psychometric instruments that are known or suspected of being able to provide quantitative evidence about elementary aspects of brain function (Griffiths, 1989; Griffiths, 1996). While in the 1980s IQ findings and in the 1990s MRI evidence made the most powerful impact on the shaping of treatment policy in PKU, the neuropsychological approach has gradually become more influential (Burgard, 1996). The principal reason for the move away from traditional IQ tests, with their high proportion of long-term or crystallized memory items, is the relative dullness of the latter in detecting the effects of subtle neurotoxicological and biochemical factors. The impetus in recent years has been towards the development of more sensitive neuropsychological measures many of which derive from clinical research on subjects with identifiable neurological lesions (Benton, 1994).

### **Neuropsychological approach to treatment discontinuation issue**

The treatment discontinuation study described in Chapter Three followed the approach of Brunner *et al.* (1987) in selecting neuropsychological as opposed to IQ measures. Control group comparisons were necessary as few

developmental neuropsychological measures are norm-referenced. Careful matching of experimental and control subjects yielded little in the way of firm evidence that three or more years of diet-induced hyperphenylalaninaemia led to neuropsychological deficit, this result providing additional support for the view that dietary liberalization at 10 years of age does not potentiate widespread cognitive impairment. However, IQ measures would have further contributed to knowledge about the intellectual consequences of lifting of the restricted diet and also about whether the off-diet group were deviant in relation to population norms, and it is a minor shortcoming of this study that IQ data were not collected.

The apparently phe-related poor Design Fluency finding in the treatment discontinuation study raised the question of whether executive functions might be specifically vulnerable to elevated and prolonged phe concentrations in post-treatment PKU. It was not possible to explore this issue further in subsequent discontinuation studies because of the change in policy adopted by the Glasgow clinic in the early ninties of encouraging children and adolescents with PKU to remain on a phe-restricted diet indefinitely. The negative results of the phe manipulation method reported in Chapters Five, Six and Seven, the further negative results from the correlational approach reported in Chapter Eight and finally the lack of a group effect in the control group study reported in Chapter Nine summate to undermine the notion that either global or specific intellectual deficit is an outcome in PKU when the disease is treated early and for at least a decade. Though, little is known about the influence of many years' exposure to high phe concentrations on neuropsychological test performance in adulthood after abrupt dietary cessation, sufficient numbers of older patients who voluntarily ceased dieting

in their teens should become available in the near future to permit treatment-outcome research, probably on a multi-centre basis.

Apart from the post-treatment study described in Chapter Three, the remainder all focussed on subjects who remained, at least nominally, on diet. The experimental studies reported in Chapters Six and Seven, though artificially elevating phe for longer than most of their kind, could not be said to have involved chronically extended hyperphenylalaninaemia. Nonetheless, the evidence from these studies was weighted against the adolescent period being one of CNS susceptibility to the malign effects of phe. The negative results derived from temporarily raising phe concentrations in adolescence lent further support to the view that the period of irreversible vulnerability of the CNS to phe was passed and suggested that a treatment policy of dietary liberalization or at least considerable relaxation may be pursued with impunity after the first decade of life.

### **Selection of tests**

In retrospect, the omission of a choice reaction time measure in the neuropsychological test batteries employed in the studies described in Chapters Three, Six and Eight was a weakness as this variable appears to be the only one frequently associated with functional impairment in treated PKU. A meta-analysis by Burgard (*personal communication*) of 21 neuropsychological studies in PKU reviewed by Waisbren *et al.* (1994) revealed highly conflicting results with evidence presented for and against a variety of cognitive and motor functions being impaired. However, abnormally

long response times on choice-reaction tasks emerged as a consistent feature across the groups studied, a mean increase of 111 msec for an increase in mean phe concentration of 966  $\mu\text{mol/l}$  being found. This increment in CRT is about half again the average for the normal population but has to be viewed against a very substantial rise in phe level. Deficiencies in CRT were apparently reversible by dietary restriction of phe and thus the incorporation of such a measure in the experimental manipulation studies described in Chapters Six and Seven would have possibly shed light on the temporary, or otherwise, nature of this phenomenon and perhaps added to the empirical foundation of the prefrontal dopamine theory.

There is no standard choice reaction task, but de Sonneville's Dot Pattern Exercise (de Sonneville, 1993) and the Continuous Performance Task as used in Chapter Nine conform to a CRT paradigm in being both speeded and having a response inhibition component. The negative results from the one-back and two-back variants of the CPT suggest that mean lifetime phe levels between 300 and 400  $\mu\text{mol/l}$  may confer immunity to slowing of reaction time. However, a longitudinal study of patients who were well-controlled in early life would be necessary to establish if later loss of control leads to poorer speed of psychomotor responding. As a general point, little is known or even conjectured about what practical implications such a minor deficiency might have for daily functioning.

Taylor (1988) has argued that neuropsychological tests represent the sharpest tools for uncovering subtle intellectual deficiencies occasioned by brain disease. In PKU research, treatment outcome studies have incorporated a wide range of neuropsychological instruments. However, as

only choice reaction time appears to emerge as a consistently deficient function throughout the age range in treated patients, why then has the clinical community in recent years sided so strongly with the policy of dietary prolongation? The answer to this question can best be understood historically and, in particular, by viewing the apparently contradictory evidence furnished by IQ, executive function and MRI studies.

### **Background to treatment cessation policy: IQ, MRI and executive function studies**

In the absence of empirical knowledge about treatment duration, dietary cessation between four and six years of age was commonplace in the 1970s, especially in the USA (Schuett *et al.*, 1980; Schuett & Brown, 1984). Above all, it was Cabalska *et al.*'s (1977) report describing average falls of 14 and 11 IQ points four years after treatment cessation at ages three and four respectively that caused a major revision of the policy to terminate diet in the preschool period. Others' findings (*e.g.* Koch *et al.*, 1982; Seashore *et al.*, 1985) led to a gradual upward revision of the recommended age for treatment discontinuation.

In England the policy, hitherto adopted on the grounds of caution rather than of science, had always been one of maintaining treatment until age eight (Smith *et al.*, 1978). Only the French have been consistently out of step with the trend towards extending the treatment period. In France, the restricted diet was traditionally discontinued around age five. Saudubray *et al.* (1987) reported a mean IQ of 101 at 11 years in a group who discontinued at five

years of age and interpreted this finding as supporting the French policy which, by international standards, would now be regarded as excessively early. Nevertheless, though the mean IQ of his sample nearly equated with the population norm, standard deviations over the period increased from 10 to 21, suggesting that at least some children deteriorated coincidentally with hyperphenylalaninaemia in the primary school period.

Treatment in Scotland was more conservative than that of England and certainly that of France in following a policy of abrupt cessation at 10 years of age. In Heidelberg, Germany, under Bickel's direction, strict diet was continued for as long as families could adhere to it and, when difficulties of compliance occurred, the diet was relaxed rather than fully liberalized.

Worldwide controversy over age of treatment cessation was not settled by Waisbren *et al.*'s (1980) review of 19 outcome studies when diet was stopped between ages four and eight. Waisbren and her co-authors concluded that children's intellectual performance decreased in about half the studies while the others showed no change. Holtzman *et al.* (1986) analyzed IQ and scholastic attainments in treated children at age 10, some of whom stopped at six and others at eight, concluding that the recommendation be made to clinicians to maintain patients on diet until at least eight years of age.

From the late seventies onwards there was thus emerging evidence for increasing the age for stopping treatment and that four to six years of age was too young to prevent later loss of IQ. Taken alone the IQ evidence might have led to a worldwide discontinuation policy that was in keeping with existing practice in Scotland of maintaining strict control until age 10.

Despite this apparent convergence of research opinion, in the late 1980s MRI studies showing abnormal brain structures become a second strand of evidence influencing commentators' judgements about what constitutes a safe age for treatment discontinuation. Villasana *et al.* (1989) were the first to report MRI anomalies in treated PKU, their subject being an 18 year-old who had been diagnosed and treated from birth. White matter changes were not diffuse but centred on the periventricular area. As the patient had discontinued diet at six, the complex of signs - an IQ of 81, fine motor incoordination, tremor and localized white matter lesions - implied that this was too early an age for cessation of treatment

Shaw *et al.* (1990) corroborated the above findings. In a patient of similar age (19 years), they found posterior, periventricular abnormalities on MRI. IQ was slightly lower at 75. As treatment commencement was mildly delayed at six weeks and concluded early at six years, the abnormal neurological and intellectual outcome may again have represented a combination of treatment factors. In the same year, Pearsen *et al.* (1990) published the results from MRI scans of 14 patients with classical PKU, nine of whom were treated early and continued diet at least until age five. Again, white matter abnormalities were found principally in areas immediately posterior to the cerebral ventricles. The cerebral cortex (grey matter), basal ganglia, cerebellum and brainstem were unaffected. In this group study, no correlation was found, however, between severity of white matter abnormality and IQ. Indeed, two cases with the worst-rated MRI abnormalities had IQs in the normal range.

Abnormal MRI scans were obtained by Thompson *et al.* (1990) from a group

of seven classical PKUs that were heterogeneous for treatment variables. Later, in a larger group of 25, all of whom were treated early and until at least seven, Thomson and his colleagues found periventricular and subcortical white matter lesions on MRI (Thomson *et al.*, 1993). Severity was unrelated to age of treatment, dietary control after eight and IQ. Time off diet was predictive, however, leading these authors to conclude that older patients on normal diet are at risk of CNS changes.

Finally, neuropsychological test findings obtained by Pennington and his co-workers pointed to the possibility of executive functions being specifically impaired in treated PKU (Pennington *et al.*, 1985; Welsh *et al.*, 1990). Though based on small samples that were heterogeneous for treatment history (see Chapter Eight), this view gathered momentum in the PKU community. Not only did it call into question the validity of IQ tests as phe-sensitive outcome measures but it concurred with the theory that prefrontal dopamine is selectively diminished even by the mildly elevated phe concentrations characteristic of treated PKU (Guttler & Lou, 1986; Diamond *et al.*, 1992; Diamond & Herzberg, 1996). Perhaps more than any other, the report by Diamond *et al.* (1994) of an association between executive dysfunction and concentrations in the 360 to 600  $\mu\text{mol/l}$  range in the very early years of treatment led to concern, particularly in the USA, about reducing phe levels to below 360  $\mu\text{mol/l}$  in the preschool period.

### **Dietary treatment policy in the UK**

In 1993 a working party of the UK Medical Research Council produced a

much-quoted report in which recommendations for dietary management of children with PKU were made explicit (Medical Research Council, 1993a). These included the following specific guidelines for all patients found to have neonatal, pretreatment phe concentrations above 600  $\mu\text{mol/l}$ . They have been widely incorporated into clinical practice throughout the UK at the time of writing.

- 1) Treatment should begin within 20 days of birth.
- 2) Phe concentrations should be kept above 120  $\mu\text{mol/l}$  irrespective of age, but below 360  $\mu\text{mol/l}$  during the pre-school period, below 480  $\mu\text{mol/l}$  until adolescence and below 700  $\mu\text{mol/l}$  thereafter.
- 3) Restriction of dietary phe should continue until adult life.

The importance of these recommendations for treatment policy was that they embodied a complete reversal of previous theory and practice in the UK. Until this time it had always been assumed that an age could be identified at which the nervous system became impervious to the harmful effects of hyperphenylalaninaemia and at which dietary restriction could therefore be stopped. This *volte face* in thinking was further bolstered by a report by Beasley *et al.* (1994) which went even further in recommending lifelong dietary restriction, a conclusion that was echoed by Potocnik and Widhalm (1994). The diet-for-life policy is one that has always been pursued in Eire (Naughten *et al.*, 1987) and in many parts of Holland (Slijper *et al.*, 1988). Despite the opportunities for international comparisons, for example between the extreme positions of the Irish and the French, few such studies have been undertaken until very recently (*e.g.* Burgard *et al.*, 1997).

The weakness of the 1993 MRC recommendations lay in a poor exposition of their rationale. A reappraisal of the IQ, neuropsychological, MRI and executive function evidence on which they were founded is needed and now, four years on, it is possible to review the recommendations in the light of the evidence presented above and that derived from other sources.

## **Reappraisal of the evidence**

### *IQ data*

When longer follow-up periods of IQ outcome from national collaborative studies became possible, Levy and Waisbren (1994) conducted a second major review of the literature and concluded that there was little evidence for loss of IQ when diet is carried on until at least 10 years. Furthermore, Beasley *et al.* (1994) analyzed the UK Collaborative Study database and found that dietary control after eight years had little predictive value for IQ compared with before. In the same year, Potocnik and Widhalm (1994), after reviewing 12 studies from the eighties and nineties, concluded more specifically that treatment termination at or near age 10 does not lead to subsequent loss of IQ.

Having transformed their IQ data into up-to-date estimates to compensate for normative changes over the quarter-century or so during which test results were submitted from clinics throughout the UK to the Collaborative Study's headquarters (Flynn, 1984), Smith and her colleagues concluded that the period between eight and ten years of age appeared to be a time during or

after which the CNS was considerably less vulnerable to the neurotoxic effects of excess circulating phe than before (Smith *et al.*, 1990a; Smith *et al.*, 1990b; Beasley *et al.*, 1994). Earlier support for this view had been forthcoming from Cerone *et al.* (1986) and from the German Collaborative Study (Schmidt *et al.*, 1987), the Heidelberg workers reporting an IQ range of 108 to 124 in a well-treated adolescent PKU group who were permitted a relaxed - as opposed to a liberalized - diet at age 10. Abrupt cessation around 10 years of age rather than graduated reduction of phe restriction was a characteristic of a subgroup within the US Collaborative Study (Azen *et al.*, 1991), the mean IQ for these children being normal at 101 when tested on the WISC-R, which was restandardized in North America in 1974.

Despite the lack of evidence for intellectual deficit following treatment cessation at the end of the first decade of life, the MRC Working Party formulated a policy for indefinite dietary continuation - as, paradoxically, did Potocnik and Widhalm (1994), despite having collated evidence to suggest the opposite! With hindsight, the principal reasons for this were undoubtedly the abnormal signs shown by MRI scans of the brain in treated PKU and the suspicion that executive functions were especially vulnerable to elevated phe concentrations. Following Villasana *et al.*'s (1989) seminal study, a succession of studies were published yielding evidence upholding the idea that even well-controlled levels of phe during the developmental period might lead to subtle white matter changes in subcortical structures (*e.g.* McCombe *et al.*, 1992; Thompson *et al.*, 1990; Thompson *et al.*, 1991; Toft *et al.*, 1994).

## *MRI data*

These unusual MRI patterns were troubling for policy-making in relation to the management of PKU. The specificity and consistency of the results were highly relevant to the most widely accepted mechanism of neural damage in untreated PKU, namely dysmyelination (Hommes *et al.*, 1982). Both human autopsy evidence (Malamud, 1966) and findings from the brains of PAH deficient mice (Hommes, 1994) suggested that hyperphenylalaninaemia in early life increases myelin turnover which is not compensated for by increased myelin synthesis. This in turn leads to loss of neurones and decreased connectivity and, since connectivity is the basis for complex behavioural organization, mental dysfunction results.

The question arises, however, whether this neuropathological mechanism, which appears so devastating during foetal and infant growth in PKU, loses its potency when neural structures become mature. An extended period of phe restriction permits the CNS to myelinate and arborize in a non-hostile biochemical environment. If the mechanism is then allowed to operate, does it cause white matter abnormalities which have functional significance? White matter abnormalities so far found in treated PKU are indicative of changes in white matter water content that are suggestive of increased myelin turnover, but they tend to be subtle and localized. One of the firsts hints that these structural anomalies may be benign and of little clinical relevance came from Thompson *et al.*, (1993) who reported finding no correlation between the degree of abnormality on MRI scans and IQ in a group of 25 early-treated subjects.

Corroborative evidence for a lack of association between neurological abnormality and intelligence was provided by an MRI study of 77 post-treatment adolescents and adults (Cleary *et al.*, 1994). While every scan was abnormal, the extent of the abnormality correlated randomly with IQ, leading these authors to conclude that, although the brain in PKU after treatment until late childhood is typified by deviant MRI scans, the phenomenon may have no significance for intactness of cognitive function.

Other reports of MRI data confirmed the essential picture, namely, that children treated early, well and until at least middle childhood (nine to ten years) show localized white matter abnormalities suggestive of disturbed myelination (Bick *et al.*, 1991; Toft *et al.*, 1994), but the clinical meaning of hyperintensity on the MRI scan and whether it translates into psychological malfunction remains unclear. The emerging view is that structural abnormality as gauged by the MRI scan may not imply a commensurate risk to intelligence. However, few studies of this issue using neuropsychological as opposed to IQ measures have been conducted so far.

Lou *et al.* (1982) concluded that neurological lesions as revealed by MRI had few functional consequences. Results from a neuropsychological battery that comprised paired-associate learning, free-recall of word lists, digit span, digit-symbol and trail-making were uniformly normal, though only nine of the 25 subjects had stopped treatment. Warrington and James's Metric Figures Test was the sole exception, but the authors failed to correlate concurrent or lifetime phe with performance on this perceptual test.

It is too early to dismiss MRI anomalies as being of no consequence in treated

PKU as the long-term effects of dietary cessation after the first decade are as yet largely unknown. Water density as revealed by MRI reflects white matter deposition which in turn reflects myelin synthesis. Given the demonstrated and rapid disruption of myelination in the immature brain by phe (Hommes, 1991), an open-minded view is required about the possibility of this process continuing thereafter. For example, in the mature organism, slow and insidious phe-related myelin changes may occur and harmful effects arise over a more protracted time-scale than that observed in childhood.

### *Executive test data*

The impact made by the executive dysfunction hypothesis on thinking about control of treatment variables in PKU is disproportionate to its empirical substance. Policy makers have perhaps too hastily latched on to the simple notion that mildly elevated phe concentrations potentiate subtle information processing deficits characteristic of prefrontal dysfunction without either appraising or weighing the supportive evidence. Unfortunately, thus far, the area is strewn with methodologically questionable investigations and contradictory findings. These problems were described in Chapter Eight. The publication by Diamond *et al.* (1994) has been influential in a way unwarranted by its content. More of an abstract than a report containing sufficient detail to permit replication, it lacks sample-matching data and an elementary analysis of the association between historical phe levels and performance on executive tasks. The study described in Chapter Eight strove to remedy these deficiencies and failed to provide confirmatory evidence of a link between phe concentrations that were three to five times greater than normal and executive and personality disorders.

In a subsequent study in which, instead of executive function, visual contrast sensitivity was measured as an index of dopaminergic functioning in the retinal as opposed to prefrontal neuronal systems, Diamond and Herzberg (1996) reported significantly impaired contrast sensitivity in 12 PKU subjects compared with healthy controls. However, even superficial examination of this report reveals the omission of historical phe data, thus making it impossible to relate outcome on the perceptual test with lifetime phe control. Furthermore, closer scrutiny shows that the effect may have been principally due to the controls having supranormal ability on the visual contrast test rather than the PKU subjects having subnormal ability. Thus, when average scores from the 12 PKU children in Diamond and Herzberg's sample were compared with those obtained by Scharre *et al.* (1990) from a similar-aged sample of 61 healthy children, they were not consistently worse.

In fairness to Diamond's position about phe concentrations in the 360 to 600  $\mu\text{mol/l}$  range in the early years of life possibly having an adverse effect on development of executive function, the samples on which the studies in Chapters Eight and Nine were based had rather better lifetime phe control than this and individual readings either as a totality or a even a majority did not fall consistently within this band. These considerations emphasize the need for further work on the executive dysfunction hypothesis based on samples stratified for level of historical phe control.

## **Caveats about interpretations of studies**

Research findings in PKU rapidly affect parental perceptions of the disease and its management as organizations such as the National Society for Phenylketonuria in the UK are in close contact with their American and European counterparts and information about new developments is often shared and disseminated through magazines, conferences and websites.

Particularly where studies of dietary continuation are based on samples of adolescent subjects, there is a danger of findings being misinterpreted by both commentators and parents. An example of ambiguity in findings leading to a risk of false conclusions being drawn is a study by Weglage *et al.* (1996). These authors found that early and continuously treated 10-year-olds with PKU showed no sign of psychosocial maladjustment, a piece of evidence that is in agreement with the lack of behavioural and personality deviance found in the 11-year-olds described in Chapter Seven. However, Weglage *et al.* discovered that the picture deteriorated in adolescence and that 14- to 16-year-olds with PKU were significantly maladjusted compared with healthy controls. In the younger group, correlations between phe and personality measures were non-significant but in the older subjects significant correlations emerged. The simplistic interpretation might be that the personality findings were manifestations of elevated phe and thus that loss of dietary control in adolescence leads to a decline in personal and social adjustment. Such a view would be in keeping with the executive dysfunction hypothesis.

Correlation does not imply a causal relationship, however. A third factor might

have potentiated correlated changes in the other two. It is noteworthy that when Weglage *et al.* surveyed their adolescent sample, 77 per cent indicated difficulties with staying on diet and an enormous 94 per cent wished to stop immediately. There were thus strong signs of conflict between the motivation of the clinicians and parents, on the one hand, and the teenage patients, on the other. Construing the correlation as a causal link between elevated phe and disordered conduct may be spurious. An alternative view is that negative and hostile attitudes towards the treatment might have generated both loss of control and a general heightening of maladjusted behaviour in many of the Weglage *et al.*'s mid-adolescent subjects.

In instances such as this where interpretation is unclear and causal sequences are ambiguous, it is essential to seek corroboration. In the experimental manipulation studies described in Chapters Seven and Eight where subject, carers and researchers were blind to alterations in phe concentrations, no cognitive or personality effects were demonstrated, at least over three months. Thus, the theory that from early adolescence onwards phe directly affects psychological variables is weakened. The alternative view is that the turmoil of adolescence and puberty has widespread repercussions on adherence to diet and general behaviour and that opposition to dietary maintenance at this age results in both loss of dietary control and changes for the worse on indices of integration and adjustment.

A further illustration of this point is provided by the correlational analyses performed on the data reported in Chapter Eight. A naive interpretation might be that phe was positively correlated with attention deficit. However, more thorough scrutiny revealed that a third factor, marital breakdown, appeared to

be influencing both phe control and concentration in one subject and that, on removing that subject from the analysis, the significant correlation disappeared.

### **Overstating research findings and overgeneralizing conclusions**

An example of lack of balance in evaluating research findings and of proclaiming a misleading message to the PKU community is provided by Koch *et al.* (1996) of the US Collaborative Study on PKU. These authors compared two early-treated groups, one of which maintained the phe-restricted diet until 22 years of age on average while the other stopped at a little under eight years of age on average. The WAIS-R administered at age 22 and 24 years respectively revealed a mean IQ of 104 (range 74 to 123) for the on-diet group and a mean IQ of 92 (range 69 to 116) for the off-diet group. Furthermore, 78 per cent of the on-diet subjects went to college compared with only 28 per cent of the off-diet group. Koch *et al.* concluded that these results affirmed the need for dietary phe restriction into adulthood. Koch *et al.* also implied that the earlier policy in the US of terminating diet at six years of age was misguided, the principal reason being cost to health insurance companies rather than clinical need, and argued that funds should be sought to support the provision of dietary supplements into adulthood.

This strongly political message was published in a summary article by the American PKU Society's newsletter under the emotive title 'Off-diet young adults with PKU: lives in danger' and broadcast worldwide by Schuett on the Internet without criticism or commentary. While Koch *et al.*'s argument may

provide ammunition for pressure groups campaigning for better financial support for an expensive treatment (the cost of the PKU diet is currently about \$5000 *per annum*), the reasoning behind the conclusion drawn for prolongation of the diet into adulthood is seriously flawed.

First, no statistical test of the IQ group-difference was performed. Second, dietary cessation and control were confounded as the off-diet group contained six times as many subjects as the on-diet group who had phe concentrations above 1200  $\mu\text{mol/l}$  during treatment. Third, no attempt was made to correlate historical phe levels with adult IQ. And fourth, the authors failed to acknowledge the possibility that the non-phe factors, such as low parental IQ and education, which predispose PKU subjects to lower IQs may also predispose families to poor treatment compliance and early cessation of diet.

### **Non-treatment factors**

The empirical studies described in earlier chapters focussed on treatment factors. However, non-treatment factors have a powerful influence on intellectual outcome in PKU and it is important when interpreting research findings that the two are clearly separated. Using multiple linear regression, Chang *et al.* (1983) demonstrated that parental IQ and educational level were significant predictors of IQ in the on-diet PKU child. Later, Legido *et al.* (1993) confirmed this finding. Moreover, these authors reported that parental IQ was the strongest predictor of IQ in the treated child. Few studies have quantified socio-economic status (SES) in families with a PKU child. An exception is Smith *et al.*'s (1990b) analysis of data from the UK collaborative study which

showed a strong effect of social class on IQ, the most able children being associated with classes one and two in the Registrar General's scheme. Such a statistical manipulation is generally overlooked in research into treatment factors in PKU but the social background and educational status of parents deserve greater consideration as factors in the outcome equation. They clearly have no direct influence on metabolic processes and the issue is not that the intelligence of parents and offspring are generally positively correlated (for instance, Williamson *et al.* (1981) showed that maternal IQ predicts the IQ of PKU and non-PKU siblings equally well). What matters is that parental ability probably dictates to some extent level of understanding of the condition, knowledge about biochemical mechanisms and appreciation of the constituents of dietary intake in both carers and patient which, in turn, may affect treatment compliance and the eventual neurological and psychological status of the child (Russell *et al.*, 1988; Ferguson & Griffiths, 1997).

### **Design of future research studies**

Over the years since the inception of newborn screening for PKU, arguably sufficient expertise has been amassed to identify mistakes made in research designs and to enable them to be avoided in the future. Research evidence accumulated thus far points to severity, age of treatment commencement, quality of early phe control and age of treatment cessation as being the four active phe-related factors that determine cognitive outcome in treated PKU. Operational definitions of these factors remain to be refined. In so doing, retrospective and prospective research designs must endeavour not to confound these factors or to control statistically for confounded effects if these

are unavoidable.

One methodology for guiding future research into treatment factors is to concentrate on samples that are considered by contemporary standards to be optimally treated in order to establish whether recommended treatment practices are truly prophylactic. This approach would entail selecting groups of patients with PKU whose treatment conforms to criteria currently promulgated as optimal by different national bodies and comparing results on neuropsychological and personality outcome variables in such groups with those derived either from healthy controls matched for non-metabolic independent variables or from PKU comparison groups that differ only on the factor under scrutiny. Research procedures in PKU over the past thirty years have largely been based on between-group models - for instance, severity has been investigated by contrasting the intellectual performance of mildly hyperphenylalaninaemic children with that of classical PKUs (Costello *et al.*, 1994) - but comparison of optimally treated patients with non-PKU controls and sub-optimally treated controls largely remains to be done. This omission is understandable, not only because of the circularity of the process, but also because German treatment recommendations were only made explicit in 1990 (Stellungnahme APS, 1990) and British recommendations only published in 1993 (Medical Research Council, 1993a). American guidelines on treatment targets have never been formally announced.

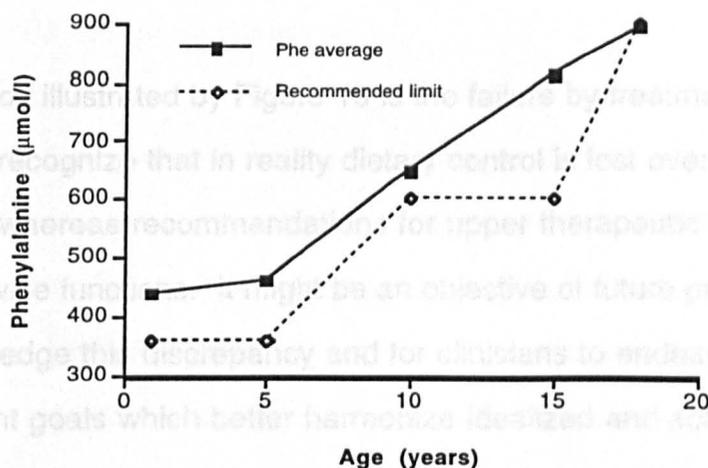
Such an approach would necessitate rather more careful selection of samples in future than perhaps has been the case in the past. An attempt to compile a sample based on present UK guidelines for optimal treatment was made in the CPT study described in Chapter Nine and executive dysfunction as

measured by sustained attention and decision-making was not found in this group. In general terms, the sample selection process entails operationally defining inclusion-exclusion criteria and is dependent on having detailed historical records available for potential subjects. Thus, computerized databases with serial test data, both biochemical and psychological, are highly desirable prerequisites for research. If not already established, existing clinical services could facilitate the development of local research datasets by immediate implementation of computer storage and analysis programs.

One difficulty of applying the criteria currently recommended as the basis for, say, treatment policy in the UK is that control levels refer to ranges within which all phe readings for the individual patient ideally should remain. For instance, the therapeutic range at present recommended in the UK (Medical Research Council, 1993a) is 120 to 360  $\mu\text{mol/l}$  for the pre-school period but an unresolved problem is a definition of what proportion of phe readings per unit of time falling within this range constitutes acceptable compliance with the standard. In the studies reported above, the annual median of phe readings was adopted as the index of dietary control according to the conclusions of Rupp and Burgard (1995), with means of annual medians calculated to summarize group or individual compliance during specified developmental periods or indeed over subjects' lifetimes. However, in studies where samples are compiled on the basis of their conformity with a broadcast standard, an index of compliance would be beneficial. In this way, a single value could represent the degree to which a sample of patients approximates to recommended parameters. Unfortunately, it would be difficult to escape incorporating an element into a formula for an index of compliance, such as a percentage threshold, the value of which would have to be set arbitrarily.

## Dietary control, outcome and policy

Despite the change in UK treatment policy, the reality of dietary control in treated PKU is that average phe levels drift inexorably upward. Figures 8 and 14 demonstrate this phenomenon. Notwithstanding the current German policy of striving to maintain phe concentrations below 363  $\mu\text{mol/l}$  up to the age of years, below 606  $\mu\text{mol/l}$  until age 15 and below 909  $\mu\text{mol/l}$  thereafter (Stellungnahme APS, 1990), Weglage *et al.* (1993) have shown that actual levels creep upward and generally surpass recommended limits across the age range. Figure 19 illustrates this trend.



**Figure 19.** Mean indices of dietary control for 34 Munster patients from birth to adolescence compared with German Paediatric Society's recommended upper limits. After Weglage *et al.* (1993).

It is important to recognise that the curve depicted in Figure 19 represents mean values. The fact of there being variation in phe control around these figures signifies that some patients have better control and others worse. Thus, although policy nowadays is for indefinite dietary continuation, research opportunities are nonetheless available for comparing outcome in poorly controlled groups with that in better controlled groups or for conducting correlational studies of historical phe and outcome as was done in Chapter Eight and in a cluster analysis study by Burgard *et al.* (1996). What may be lacking in the future is the opportunity that was exploited in Chapter Three of measuring the effect of a sudden change in the independent variable due to treatment cessation at a predetermined age. Experimental manipulation of phe, the method chosen in Chapters Six and Seven, is an alternative approach but one in which ethics dictate much shorter periods of hyperphenylalaninaemia.

A paradox illustrated by Figure 19 is the failure by treatment policy-makers in PKU to recognize that in reality dietary control is lost over childhood in a linear fashion whereas recommendations for upper therapeutic limits are expressed as stepwise functions. It might be an objective of future practice to acknowledge this discrepancy and for clinicians to endeavour to construct treatment goals which better harmonize idealized and actual dieting behaviour. The difficulty with using stepped rather than a progressive phe targets in treatment is that, with the passage of time along a step, the trend will be for target and achieved phe levels to diverge. As they become more discrepant, there is a risk that motivation will be diminished. This is a 'knowledge of results' or reinforcement phenomenon which could easily be counteracted by the practitioner adopting a target model that was linear or

curvilinear across the age range. For instance, to reflect better what is now suspected about the toxicity of phe in relation to the maturity of the CNS, a negatively accelerated target curve could be adopted, the demands of which would be for very strict dietary control in infancy, minimal relaxation of phe restriction during the pre-school period, moderate relaxation during the primary school years and near-liberalization in adolescence and adulthood. In practice, targets could be expressed and adjusted annually as increasing acceptable phe levels so that given targets would be operative for given yearly ages.

### **Aims of future research**

Three inter-related research areas are identifiable as extensions of the issues raised in the foregoing chapters.

The first is the need for follow-up neuropsychological studies of off-diet PKU adults who were treated early and strictly and who discontinued after the first decade. A lifespan developmental approach is required as there are as yet no grounds for making the assumption that lack of evidence for cognitive or motor deterioration in early adulthood implies protection from the neurotoxic effects of phe in middle or old age. Knowledge derived from charting the progress of this group would inform the 'diet-for-life' issue.

The second is the need for independent studies to corroborate or annul the view that applying current treatment recommendations for dietary management, such as those formally articulated in the UK and in Germany,

will prevent concurrent or subsequent cognitive impairment. In particular, the question of the relationship between age level and therapeutic range in childhood needs further exploration in order ultimately to formulate policy and specify optimal limits.

The third is the need for greater illumination of the executive dysfunction hypothesis. It remains unclear whether prefrontal functions are hyper-sensitive to moderate elevations in phe level. For instance, there may be phe thresholds at varying ages at which specific impairment of executive functions arises. It is foreseeable that solution of the executive dysfunction problem will be hampered by lack of agreement about what psychometric instruments might be maximally sensitive to deficits in this domain or, more fundamentally, what exact cognitive operations fall under the rubric of 'executive function'.

## **Conclusions**

The results of the studies reported above generally failed to support the theses: a) that cessation of dietary treatment for PKU at the end of the first decade of life leads to subsequent neuropsychological impairment, b) that temporary elevation of phe as an interruption to normal prophylactic therapy in adolescence produces immediate deficits in cognition and personality, and c) that specific disorders of executive function arise when phe concentrations are maintained within or close to limits currently advised in the UK.

Arguably, negative findings constitute the most satisfactory outcome in preventive medicine. Restriction of dietary phe in PKU is designed to arrest

the natural course of a disease that otherwise would lead to severe mental impairment. The evidence obtained from the studies described here suggests that even subtle disorders of information processing can be avoided if circulating phe concentrations are kept at low levels in the early years of life.

The findings concur with a small but growing body of neuropsychological evidence that derives from research studies that are characterized by careful quantification of treatment factors and background phe levels. Henceforth, research methodologies based on more rigorous control and separation of treatment factors than has been achieved in the past are desirable.

The import of attributions made by patients and their families to phe should be viewed with caution by clinicians. Evidence was found that neither school-age children nor their parents could reliably perceive artificially-induced hyperphenylalaninaemia on the basis of information stemming from subjective state or behaviour.

The implications of the present research for treatment policy in PKU are that the case for lifelong phe restriction in patients with the classical form of the disease may be over-stated. The recommendation of the MRC Working Party on Phenylketonuria that dietary restriction of phe be prolonged into adolescence and adulthood may have been premature (Medical Research Council, 1993a). Data presented in this volume suggest that considerable relaxation of the restricted diet after the first decade of life may be acceptable and not deleterious to long-term cognitive and motor functioning. Furthermore, at the time they were published, the MRC Working Party's pronouncements on the optimal therapeutic range for phenylalanine

concentrations lacked a sound empirical basis and even precise age ranges for recommended concentrations were left unspecified. The results from the studies reported above suggest that adherence to the MRC guideline of maintaining concentrations within the 120 to 360  $\mu\text{mol/l}$  band during the preschool period prevents neuropsychological dysfunction, at least in the short-term, but further evidence is needed about functional outcome in patients whose preschool phe control has strayed into the 360 to 600  $\mu\text{mol/l}$  range or beyond during this period.

## REFERENCES

- Alvord, E.C., Stevenson, L.D., Vogel, F.S. & Engle, R.L. (1950) Neuropathological findings in phenylpyruvic oligophrenia (phenylketonuria). *Journal of Neuropathology and Experimental Neurology*, **9**, 298-307.
- Anderson, V.E., Siegel, F.S., Fisch, R.O. & Wirt, R.D. (1969) Responses of phenylketonuric children on a continuous performance test. *Journal of Abnormal Psychology*, **74**, 358-362.
- Andresen, J. (1991) PKU patients - problems in adolescence. *Proceedings of the 5th Annual ESPKU Meeting, De Haan, Belgium*, 10-15. Liverpool: SHS International.
- Azen, C.G., Koch, R. & Friedman, E.G. (1991) Intellectual development in 12-year-old children treated for phenylketonuria. *American Journal of Diseases of Children*, **145**, 35-39.
- Barbizet, J. (1970) *Human Memory and its Pathology*. San Francisco: Freeman. Pp. 95-106.
- Barclay, A. & Walton, O. (1988) Phenylketonuria: implications for initial serum phenylalanine levels on cognitive development. *Psychological Reports*, **63**, 135-142.
- Batshaw, M.L. & Perret, Y.M. (1992) *Children with Disabilities, 3rd Edition*. Baltimore: Brookes. Pp. 159-170.
- Bauman, M.L., & Kemper, T.L. (1982) Histoanatomy of the brain in untreated human phenylketonuria. *Acta Neuropathologica*, **58**, 55-63.
- Beasley, M.G., Costello, P.M. & Smith, I. (1994) Outcome of treatment in young adults with phenylketonuria detected by routine neonatal screening between 1964 and 1971. *Quarterly Journal of Medicine*, **87**, 155-160.

Beaumont, J.G. (1983) *Introduction to Neuropsychology*. Oxford: Blackwell. Chapters 3 and 4.

Behbehani, A.W. & Langenbeck, U. (1982) A combined study of neurophysiological, biochemical and psychological parameters in children with phenylketonuria. *Journal of Inherited Metabolic Disease*, **5** (Supplement 1), 29-30.

Behbehani, A.W., Vollrath, M., Matschke, I. & Langenbeck, U. (1986) Termination of strict diet in phenylketonuria. *Journal of Inherited Metabolic Disease*, **9** (Supplement 2), 218-222.

Belmaker, R.H. & Fleischmann, A. (1995) Transcranial magnetic stimulation: a potential new frontier in psychiatry. *Biological Psychiatry*, **38**, 419-421.

Bender, D.A. (1993) *Introduction to Nutrition and Metabolism*. London: UCL Press. Pp. 292-294.

Benton, A.L. (1974) *The Revised Visual Retention Test, 4th Edition*. New York: The Psychological Corporation.

Benton, A.L. (1994) Neuropsychological assessment. *Annual Review of Psychology*, **45**, 1-23.

Berman, P.W., Waisman, H.A. & Graham, F.K. (1966) Intelligence in treated phenylketonuric children: a developmental study. *Child Development*, **37**, 731-736.

Berman, P.W., Waisman, H.A. & Graham, F.K. (1968) Effectiveness of dietary treatment in phenylketonuria: what is the proof? *Developmental Medicine and Child Neurology*, **10**, 411-418.

- Berry, H.K., O'Grady, D.K., Perlmutter, L.J. & Boffinger, M.K. (1979) Intellectual development and academic achievement of children treated early for phenylketonuria. *Developmental Medicine and Child Neurology*, **21**, 311-320.
- Bick, U., Fahrendorf, G., Ludolph, A.C., Vassallo, P., Weglage, J., & Ullrich, K. (1991) Disturbed myelination in patients with treated hyperphenylalaninaemia: evaluation with magnetic resonance imaging. *European Journal of Pediatrics*, **150**, 185-189.
- Bickel, H. (1980) Phenylketonuria: past, present and future. *Journal of Inherited Metabolic Disease*, **3**, 123-132.
- Bickel, H. (1996) The first treatment of phenylketonuria. *European Journal of Pediatrics*, **155** (Supplement 1), S2-S3.
- Bickel, H., Beckers, R.G. & Wamberg, E. (1973) Collective results of mass screening for inborn metabolic errors in eight european countries. *Acta Paediatrica Scandinavica*, **62**. 413-416.
- Bickel, H., Gerrard, J. & Hickmans, E.M. (1953) Influence of phenylalanine intake on phenylketonuria. *Lancet*, **2**, 812-813.
- Bickel, H., Gerrard, J. & Hickmans, E.M. (1954) The influence of phenylalanine intake on the chemistry and behaviour of a phenylketonuric child. *Acta Paediatrica*, **43**, 64-68.
- Brenton, D. (1996) Effects of long-term dietary treatment versus long-term diet discontinuation. *Paper presented at 10th Annual ESPKU Conference, Aegerisee, Switzerland.*
- Brody, T. (1994) *Nutritional Biochemistry*. London: Academic Press. Pp. 15-19.

Brooks, P.H. & Baumeister, A.A. (1977) A plea for consideration of ecological validity in the experimental psychology of mental retardation. *American Journal of Mental Deficiency*, **81**, 407-416.

Brown, E.S. & Warner, R. (1976) Mental development of phenylketonuric children on or off diet after the age of six. *Psychological Medicine*, **6**, 287-296.

Brunner, R.L. & Berry, H. (1987) Phenylketonuria and sustained attention: the continuous performance test. *International Journal of Clinical Neuropsychology*, **9**, 68-70.

Brunner, R.L., Jordan, M.K. & Berry, H.K. (1983) Early-treated phenylketonuria: neuropsychologic consequences. *Journal of Pediatrics*, **102**, 831-835.

Burgard, P. (1991) The intellectual development of children with phenylketonuria. *Proceedings of the 5th Annual ESPKU Meeting, De Haan, Belgium..* Liverpool: SHS International. Pp. 2-5.

Burgard, P. (1994) Results of the German PKU collaborative study: implications for PKU treatment. *Proceedings of the 8th Annual ESPKU Meeting, Fulda, Germany.* Liverpool: SHS International. Pp. 14-15.

Burgard, P (1996) Neuropsychological investigations in PKU. *Paper presented at 10th Annual ESPKU Conference, Aegerisee, Switzerland.*

Burgard, P., Rey, F., Rupp, A., Abadie, V. & Rey, J. (1997) Neuropsychologic functions of early treated patients with phenylketonuria, on and off diet: results of a cross-national and cross-sectional study. *Pediatric Research*, **41**, 368-374.

Cabalska, B., Duczynska, N., Borzymowska, J., Sorska, K., Koslacz-Folga, A. & Bozkowa, K. (1977) Termination of dietary treatment in phenylketonuria. *European Journal of Pediatrics*, **126**, 253-262.

Cairns, E. & Cammock, T. (1978) Development of a more reliable version of the Matching Familiar Figures Test. *Developmental Psychology*, **14**, 555-560.

Carstairs, V. & Morris, R. (1992) *Deprivation and Health in Scotland*. Aberdeen: Aberdeen University Press. Pp. 285-301.

Cattell, R.B (1959) *The Children's Personality Questionnaire*. Champaign, Illinois: IPAT.

Cattell, R.B (1971) *Abilities: Their Structure, Growth and Action*. Boston: Houghton Mifflin.

Centerwall, W.R. & Centerwall, S.A. (1961) *Phenylketonuria*. Washington, DC: US Department of Health, Education and Welfare.

Cerone, R., Scalisi, S., Schiaffino, M.C. & Romano, C. (1986) Effects of stopping phenylalanine restricted diet on intellectual progress of children with phenylketonuria. *Journal of Inherited Metabolic Disease*, **9** (Supplement 2), 223-224.

Chang, P.-N., Cook, R.D. & Fisch, R.O. (1983) Prognostic factors of the intellectual outcome of phenylketonurics: on and off diet. *Journal of Psychiatric Treatment and Evaluation*, **5**, 157-163.

Chang, P.-N. & Fisch, R.O. (1976) Observation of behavior and personality characteristics of phenylketonurics according to their dietary duration. *Psychological Reports*, **39**, 835-841.

Claridge, G. (1970) *Drugs and Human Behaviour*. London: Allen Lane. Chapter 4.

Clark, B., Cockburn, F & Tyfield, L. (1995) *The Child with PKU*. Milton Keynes: National Society for Phenylketonuria.

Clarke, J.T.R., Gates, R.D., Hogan, S.E., Barrett, M. & MacDonald, G.W. (1987) Neuropsychological studies on adolescents with phenylketonuria returned to phenylalanine-restricted diets. *American Journal on Mental Retardation*, **92**, 255-262.

Cleary, M.A., Walter, J.H., Wraith, J.E., Jenkins, J.P.R., Alani, S.M., Tyler, K., & Whittle, D. (1994) Magnetic resonance imaging of the brain in phenylketonuria. *Lancet*, **344**, 87-90.

Cleland, C.C. (1978) *Mental Retardation*. Englewood Cliffs, NJ: Prentice-Hall. Pp. 243-245

Cole, M. & Bruner, J.S., (1971) Cultural differences and inferences about psychological processes. *American Psychologist*, **26**, 867-876.

Corkum, P.V. & Siegel, L.S. (1993) The continuous performance test: a valuable tool for use with children with attention-deficit-hyperactivity disorder. *Journal of Child Psychology and Psychiatry*, **34**, 1217-1239.

Costello, P.M., Beasley, M.G., Tillotson, S.L. & Smith, I. (1994) Intelligence in mild atypical phenylketonuria. *European Journal of Pediatrics*, **153**, 260-263.

Corsellis, J.A.N. (1953) The pathological report of a case of phenylpyruvic oligophrenia. *Journal of Neurology, Neurosurgery and Psychiatry*, **16**, 139-141.

Cotton, R.G.H. (1990) Heterogeneity of phenylketonuria at the clinical, protein and DNA levels. *Journal of Inherited Metabolic Disease*, **13**, 739-750.

Daniel, W.W. (1990) *Applied Non-Parametric Statistics, 2nd Edition*. Boston: PWS-Kent. Pp. 395-400.

Danks, D.M. & Cotton, R.G.H. (1987) Future developments in phenylketonuria. *Enzyme*, **38**, 296-301.

Davis, D.D., McIntyre, C.W., & Murray, M.E. (1986) Cognitive styles in children with dietary-treated phenylketonuria. *Educational and Psychological Research*, **6**, 9-15.

Davison, A.N. (1973) Inborn errors of amino acid metabolism affecting myelination of the central nervous system. In, F.A. Hommes & C.J. Van Den Berg (Eds.), *Inborn Errors of Metabolism*. London: Academic Press. Pp. 55-67.

Demellweek, C. (1996) Memory impairment in early-treated off-diet adolescents and adults with phenylketonuria (PKU). *Inborn Errors Review Series*, **6**, 6 (abstract). Liverpool: SHS International.

Devinsky, O. (1993) Electrical and magnetic stimulation of the central nervous system: historical overview. *Advances in Neurology*, **63**, 43-50.

Diamond, A. (1994) Phenylalanine levels of 6-10 mg/dl may not be as benign as once thought. *Acta Paediatrica*, **83** (Supplement 407), 89-91

Diamond, A. (1995) Children treated early and continuously for PKU: deficits in cognition and vision , and why. *Paper presented at Department of Child Health, Royal Hospital for Sick Children, Glasgow.*

Diamond, A., Ciaramitaro, V., Donner, E., Hurwitz, W., Lee, E., Grover, W. & Minarcik, C. Prefrontal cortex cognitive deficits in early-treated PKU. *Society of Neuroscience Abstracts*, **18**, 1063.

Diamond, A., Ciaramitaro, V., Donner, E., Djali, S. & Robinson, M.B. (1994) An animal model of early-treated PKU. *Journal of Neuroscience*, **14**, 3072-3082.

Diamond, A. & Goldman-Rakic, P.S. (1986) Comparative development in human infants and infant rhesus monkeys on cognitive functions that depend on prefrontal cortex. *Society of Neurosciences Abstracts*, **12**, 742.

Diamond, A. & Herzberg, C. (1996) Impaired sensitivity to visual contrast in children treated early and continuously for phenylketonuria. *Brain*, **119**, 523-538.

Diamond, A., Hurwitz, W., Lee, E.Y., Grover, W. & Minarcik, C. (Unpublished) Cognitive deficits on frontal cortex tasks in children with early-treated PKU: results of two years of longitudinal study.

Dobson, J.C., Williamson, M.L., Azen, C. & Koch, R. (1977) Intellectual assessment of 111 four-year-old children with phenylketonuria. *Paediatrics*, **60**, 822-827.

Duncan, J. (1986) Disorganization of behavior after frontal lobe damage. *Cognitive Neuropsychology*, **3**, 271-290.

Duncan, J. (1995) Attention, intelligence and the frontal lobes. In, M.S. Gazzaniga, (Ed.), *The Cognitive Neurosciences*. Cambridge, Massachusetts: MIT Press. Pp. 721-733.

Dunn, L.M., Dunn, L.M., Whetton, C. & Pintillie, D. (1982) *British Picture Vocabulary Scale*. Windsor: NFER-Nelson.

Editorial. (1991) Phenylketonuria grows up. *Lancet*, **337**, 1256-1257.

Egger, J., Carter, C.M., Graham, P.J., Gumley, D., Soothill, J.F. (1985) Controlled trial of oligoantigenic treatment in the hyperkinetic syndrome. *Lancet*, **14**, 540-545.

Eisensmith, R.C. & Woo, S.L.C. (1991) Phenylketonuria and the phenylalanine hydroxylase gene. *Molecular Biology and Medicine*, **8**, 3-18.

Faust, D., Libon, D. & Pueschel, S. (1986) Neuropsychological functioning in treated phenylketonuria. *International Journal of Psychiatry in Medicine*, **16**, 169-177.

Feinberg, S.B. & Fisch, R.O. (1962) Roentgenologic findings in growing long bones in phenylketonuria. *Radiology*, **78**, 394-399.

Ferguson, L. & Griffiths, P. (1997) Parents' and children's knowledge of PKU. *NSPKU News and Views*, **81**, 10-11.

Fishler, K., Azen, C., Henderson, R., Friedman, E.G. & Koch, R. (1987) Psychoeducational findings among children treated for phenylketonuria. *American Journal of Mental Deficiency*, **92**, 65-73.

Flynn, J.R. (1984) The mean IQ of Americans: massive gains 1932-1978. *Psychological Bulletin*, **95**, 29-51.

Folling, A. (1971) The original detection of phenylketonuria. In, H. Bickel, F. Hudson & L.I. Woolf (Eds.), *Phenylketonuria and some other inborn errors of amino acid metabolism*. Stuttgart: Thieme. Pp. 1-10.

Folling, I. (1994) The discovery of PKU by Dr Asbjorn Folling: Norway, 1934. *Acta Paediatrica*, **83** (Supplement 407), 4-10.

Frankenburg, W.K., Goldstein, A.D. & Olsen, C. (1973) Behavioral consequences of increased phenylalanine intake by phenylketonuric children. *American Journal of Mental Deficiency*, **77**, 524-532

Fulton, T.R., Triano, T., Rabe, A. & Loo, Y.H. (1980) Phenylacetate and the enduring behavioral deficit in experimental phenylketonuria. *Life Sciences*, **27**, 1271-1281.

Garrod, A.E. (1909) *Inborn Errors of Metabolism*. London: Frowde, Hodder & Stoughton.

Gearheart, B.R. & Litton, F.W. (1975) *The Trainable Retarded*. St. Louis: Mosby.

- Gordon, N. (1976) *Paediatric Neurology for the Clinician*. London: Heinemann. Pp.42-43.
- Gravetter, F.J. & Walnau, L.B. (1992) *Statistics for the Behavioral Sciences, 3rd Edition*. St. Paul, West. Pp. 233-234.
- Griffiths, P. (1989) Investigative methods in paediatric neuropsychology. In, J.R. Crawford & D.M. Parker (Eds.), *Developments in Clinical and Experimental Neuropsychology*. New York: Plenum. Pp. 247-258.
- Griffiths, P. (1991) Word-finding ability and design fluency in developmental dyslexia. *British Journal of Clinical Psychology*, **30**, 47-60.
- Griffiths, P. (1996) Paediatric neuropsychology. *The Psychologist*, **9**, 511-514.
- Griffiths, P., Paterson, L. & Harvie, A. (1995) Neuropsychological effects of subsequent exposure to phenylalanine in adolescents and young adults with early-treated phenylketonuria. *Journal of Intellectual Disability Research*, **39**, 365-372.
- Guthrie, R. (1996) The introduction of newborn screening for phenylketonuria: a personal history. *European Journal of Pediatrics*, **155** (Supplement 1), S4-S5.
- Guthrie, R. (1980) Organization of a regional newborn screening laboratory. In, H. Bickel, R. Guthrie & G. Hammersen (Eds.), *Neonatal Screening for Inborn Errors of Metabolism*. Berlin: Springer. Pp 259-265.
- Guthrie, R. & Susie, (1963) A simple phenylalanine method for detecting PKU in large populations of newborn infants. *Pediatrics*, **32**, 338-343.
- Guttler, F. (1980) Hyperphenylalaninaemia: diagnosis and classification of the various types of phenylalanine deficiency in childhood. *Acta Paediatrica*, Supplement **280**, S1-S80.

Guttler, F. (1984) Phenylketonuria: 50 years since Folling's discovery and still expanding our clinical and biochemical knowledge. *Acta Paediatrica Scandinavica*, **73**, 705-716.

Guttler, F. (1996) Molecular studies in PKU: implications for treatment. *NSPKU News and Views*, **80**, 6.

Guttler, F. & Lou, H. (1986) Dietary problems of phenylketonuria: effect on CNS transmitters and their possible role in behaviour and neuropsychological function. *Journal of Inherited Metabolic Disease*, **9** (Supplement 2), 169-177.

Halperin, J.M., Sharma, V., Grenblatt, E. & Schwartz, S.T. (1991) Assessment of the continuous performance test: reliability and validity in a nonreferred sample. *Psychological Assessment*, **3**, 603-608.

Halvorsen, S., Heyerdahl, S., Kindt, E., Motzfeldt, K., Vandvik, I.H. & Aarseth, J. (1989) Effect of gradual liberalization of diet on cognitive functions and behavior in PKU. *Pediatric Research*, **26**, 514.

Hanley, W.B., Feigenbaum, A., Clarke, J.T., Schoonheydt, W. & Austin, V. (1993) Vitamin B12 deficiency in adolescents and young adults with phenylketonuria. *Lancet*, **342**, 997.

Hanley, W.B., Linsao, L.S. & Netley, C. (1971) The efficiency of dietary therapy for phenylketonuria. *Canadian Medical Association Journal*, **104**, 1089-1091.

Harris, H. (1970) Genetical theory and the inborn errors of metabolism. *British Medical Journal*, **1**, 321.

Harvey, E.L. & Kirk, S.F. (1995) The use of a low phenylalanine diet in response to the challenging behaviour of a man with untreated phenylketonuria and profound learning disabilities. *Journal of Intellectual Disability Research*, **39**, 520-526.

Haxby, J.V., Parasuraman, R., Laloude, F. & Abboud, H. (1993) Superlab: general purpose Macintosh software for human experimental psychology and psychological testing. *Behavior Research Methods, Instruments and Computers*, **25**, 400-425.

Hebb, D.O. (1945) Man's frontal lobes: a critical review. *Archives of Neurology and Psychology*, **54**, 10-24.

Heaton-Ward, W.A. & Wiley, Y. (1984) *Mental Handicap, 5th Edition*. Bristol: Wright. Pp. 38-40.

Hinshaw, S.P. (1994) *Attention Deficits and Hyperactivity in Children*. London: Sage. Pg. 52.

Hogan, S.E., Gates, R.D., MacDonald, G.W. and Clarke, J.T.R. (1986) Experience with adolescents with phenylketonuria returned to phenylalanine-restricted diets. *Journal of the American Dietetic Association*, **86**, 1203-1207.

Holtzman, N.A., Kronmal, R.A., Doornick, W. van, Azen, C. & Koch, R. (1986) Effect of age at loss of dietary control on intellectual performance and behaviour of children with phenylketonuria. *New England Journal of Medicine*, **314**, 593-598.

Hommes, F.A. (1991) On the mechanism of permanent brain dysfunction in hyperphenylalaninaemia. *Biochemical Medicine and Metabolic Biology*, **46**, 277-287.

Hommes, F.A. (1994) Loss of neurotransmitter receptors by hyperphenylalaninaemia in the HPH-5 mouse brain. *Acta Paediatrica*, Supplement **407**, 120-121.

Hommes, F.A., Eller, A.G. & Taylor, E.H. (1982) Turnover of the fast components of myelin and myelin proteins in experimental hyperphenylalaninaemia: relevance to termination of dietary treatment in human phenylketonuria. *Journal of Inherited Metabolic Disease*, **5**, 21-27.

Hsia, D.Y.-Y. (1967) Phenylketonuria 1967. *Developmental Medicine and Child Neurology*, **9**, 531-540.

Jaspars, J., Fincham, F.D. & Hewstone, M. (1983) *Attribution Theory and Research*. London: Academic.

Jervis, G.A. (1939) The genetics of phenylpyruvic oligophrenia. *Journal of Mental Science*, **85**, 719-762.

Jervis, G.A. (1953) Deficiency of phenylalanine oxidizing system. *Proceedings of the Society of Experimental and Biological Medicine*, **82**, 514-520.

Jones-Gotman, M. & Milner, B. (1977) Design fluency: the invention of nonsense drawings after focal cortical lesions. *Neuropsychologia*, **15**, 653-674.

Kang, E.D., Sollee, N.D. & Gerald, P.S. (1970) Results of treatment and termination of the diet in phenylketonuria (PKU). *Pediatrics*, **46**, 881-884.

Kaplan, A.R. (1969) Phenylketonuria: a review. In, M. Manosevitz, G. Lindsay & D. Thiessen (Eds.), *Behavioral Genetics: Method and Research*. New York: Academic Press. Pp. 424-438.

Kaufman, S. (1977) Phenylketonuria: biochemical mechanisms. *Advances in Neurochemistry*, **2**, 1-13.

Kelly, T.P., Borrill, H.S. & Maddell, D.L. (1996) Development and assessment of executive function in children. *Child Psychology and Psychiatry Review*, **1**, 46-51.

Kent, A. (1994) The gap between finding the gene and having a cure. *NSPKU News and Views*, **70**, 19.

Kinney, H.C., Brody, B.A., Kloman, A.S. & Gilles, F.H. (1988) Sequence of central nervous system myelination in human infancy. *Journal of Neuropathology and Experimental Neurology*, **47**, 217-234.

Klee, S. & Garfinkel, B.D. (1983) The computerized Continuous Performance Test. *Journal of Abnormal Child Psychology*, **1**, 487-496.

Koch, R., Azen, C., Friedman, E.G., Fishler, K., Bauman-Frischling, C. & Lin, T. (1996) Care of the adult with phenylketonuria. *European Journal of Pediatrics*, **155** (Supplement 1), S90-S92.

Koch, R., Azen, C.G., Hurst, N., Friedman, E.G. & Fishler, K. (1987) The effects of diet discontinuation in children with phenylketonuria. *European Journal of Pediatrics*, **146** (Supplement 1), A12-A16.

Koch, R., Friedman, E.G., Williamson, M.L. & Azen, C.G., (1982) Preliminary report on the effects of diet discontinuation in phenylketonuria. *Journal of Inherited Metabolic Disease*, **5** (Supplement 1), 63-64.

Koch, R., Friedman, E.G., Azen, C.G., Williamson, M.L. & Donnell, G.N. (1985) Report from the United States collaborative study of children treated for phenylketonuria (PKU). In, H. Bickel & U. Wachtel (Eds.), *Inherited Diseases of Amino Acid Metabolism*. Stuttgart: Thieme. Pp. 134-148.

Koch, R. & Wenz, E. (1987) Phenylketonuria. *Annual Review of Nutrition*, **7**, 117-135.

Koff, E., Kammerer, B., Boyle, P. & Pueschel, S.M. (1979) Intelligence and phenylketonuria: effects of dietary termination. *The Journal of Pediatrics*, **94**, 534-537.

Kolb, B. & Wishaw, I.Q. (1996) *Fundamentals of Human Neuropsychology*, 4th Edition. New York: Freeman. Pp. 367-368.

Krause, W., Halminski, M., McDonald, L., Dembure, P., Salvo, R., Freides, D. & Elsas, L. (1985) Biochemical and neuropsychological effects of elevated plasma phenylalanine in patients with treated phenylketonuria. *Journal of Clinical Investigation*, **75**, 40-48.

Legido, A., Tonyes, L., Carter, D., Schoemaker, A., Di George, A., & Grover, W.D. (1993) Treatment variables and intellectual outcome in children with classic phenylketonuria. *Clinical Pediatrics*, **32**, 417-425.

Levy, H.L. & Waisbren, S.E. (1994) PKU in adolescents: rationale and psychosocial factors in diet continuation. *Acta Paediatrica*, Supplement **407**, 92-97.

Lidsky, A.S., Law, M.L., Morse, H.G., Kao, F.T., Rabin, M., Ruddle, F. & Woo, S.L.C. (1985) Regional mapping of the phenylalanine-hydroxylase gene and the phenylketonuria locus in the human genome. *Proceedings of the National Academy of Sciences of the United States of America*, **82**, 6221-6225.

Loo, Y.H., Fulton, T.R., Hyde, K.R. & Wisniewski, H.M. (1984) Biochemical indices of neuronal development in experimental phenylketonuria: high affinity transport systems and gangliosides. *Developmental Neuroscience*, **6**, 235-245.

Losier, B.J., McGrath, P.J. & Klein, R.M. (1996) Error patterns on the continuous performance test in non-medicated and medicated samples of children with and without ADHD: a meta-analytic review. *Journal of Child Psychology and Psychiatry*, **37**, 971-987.

Lou, H.C., Guttler, F., Lykkelund, C., Bruhn, P. & Niederwieser, A. (1985) Decreased vigilance and neurotransmitter synthesis after discontinuation of dietary treatment for phenylketonuria in adolescents. *European Journal of Pediatrics*, **144**, 17-20.

Lou, H.C., Lykkelund, C., Gerdes, A.-M., Udesen, H. & Bruhn, P. (1987) Increased vigilance and dopamine synthesis by large doses of tyrosine or phenylalanine restriction in phenylketonuria. *Acta Paediatrica Scandinavica*, **76**, 560-565.

Mackworth, J. (1969) *Vigilance and Habituation*. Harmondsworth: Penguin. Pp. 13-45.

Malamud, N. (1966) Neuropathology of phenylketonuria. *Journal of Neuropathology and Experimental Neurology*, **25**, 254-268.

Marholin, D., Pohl, R.E., & Steward, R.M. (1978) Effects of diet and behavior therapy on social and motor behavior of retarded phenylketonuric adults. *Pediatric Research*, **12**, 179-187.

Mazzocco, M.M.M., Nord, A.M., Doornick, W. van, Greene, C.L., Kover, C.G. & Pennington, B.F. (1994) Cognitive development among children with early-treated phenylketonuria. *Developmental Neuropsychology*, **10**, 133-151.

McCombe, P.A., McLaughlin, D.B., Chalk, J.B., Brown, N.N., McGill, J.J. & Pender, M.P. (1992) Spasticity and white matter abnormalities in adult phenylketonuria. *Journal of Neurology, Neurosurgery and Psychiatry*, **55**, 359-361.

McDonald, J.D. (1994) The PKU mouse project: its history, potential and implications. *Acta Paediatrica*, **83**, 122-123.

McDowell, S.K. (1996) A role for dopamine in executive function deficits. *Journal of Head Trauma Rehabilitation*, **11**, 89-92.

Medical Research Council (1981) Routine neonatal screening for phenylketonuria in United Kingdom 1964-78. *British Medical Journal*, **282**, 1680-1684.

Medical Research Council (1993a) Recommendations on the dietary management of phenylketonuria. *Archives of Disease in Childhood*, **68**, 426-427.

Medical Research Council (1993b) Phenylketonuria due to phenylalanine hydroxylase deficiency: an unfolding story. *British Medical Journal*, **306**, 115-119.

Michals, K., Azen, C., Acosta, P., Koch, R. & Matalon, R. (1988) Blood phenylalanine levels and intelligence of 10-year-old children with PKU in the National Collaborative Study. *Journal of the American Dietetic Association*, **88**, 1226-1229.

Michals, K., Dominik, M., Schuett, V., Brown, E. & Matalon, R. (1985) Return to diet therapy in patients with phenylketonuria. *The Journal of Pediatrics*, **106**, 933-936.

Miller, E. (1977) *Abnormal Ageing: The Psychology of Senile and Presenile Dementia*. London: Wiley. Pp. 110-116.

Moffitt, T.E. (1993) The neuropsychology of conduct disorder. *Development and Psychopathology*, **5**, 135-151.

Moncrieff, A. & Wilkinson, R.H. (1961) Relaxation of phenylalanine restriction in the treatment of PKU. *British Medical Journal*, **1**, 763-767.

Morgan, C.T. (1965) *Physiological Psychology, 3rd Edn*. New York: McGraw-Hill.

Murphy, D. (1963) Termination of dietary treatment of phenylketonuria. *Irish Journal of Medical Science*, **6**, 355-358.

Naglieri, J.A., LeBuffe, P.A. & Pfeiffer, S.I. (1994) *Devereux Scales of Mental Disorders*. San Antonio: The Psychological Corporation.

National Society for Phenylketonuria (1996) Dietary Information for Treatment of Phenylketonuria. Milton Keynes: NSPKU.

Naughten, E.R., Kiely, B., Saul, I. & Murphy, D. (1987) Phenylketonuria: outcome and problems in a 'diet-for-life' clinic. *European Journal of Pediatrics*, **146** (Supplement 1), A23-A24.

Neisser, U. (1967) *Cognitive Psychology*. New York: Appleton-Century-Crofts. Pp. 65-72.

Nyhan, W.L. (1981) Inheritable abnormalities of amino acid metabolism. In, M. Florkin & E.H. Stotz (Eds.), *Comprehensive Biochemistry*. Amsterdam: Elsevier. Pp. 313-326.

Pearson, K.D., Gean-Marton, A.D., Levy, H.L. & Davis, K.R. (1990) Phenylketonuria: MR imaging of the brain with clinical correlation. *Radiology*, **177**, 437-440.

Pennington, B.F. (1991) *Diagnosing Learning Disorders*. New York: Guilford. Pp. 3-22.

Pennington, B.F., Doornick, W.J. van , McCabe, L.L. & McCabe E.R.B. (1985) Neuropsychological deficits in early-treated phenylketonuric children. *American Journal of Mental Deficiency*, **89**, 467-474.

Pennington, B.F & Ozonoff, S. (1996) Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry*, **37**, 51-87.

Penrose, L. & Quastel, J.H. (1937) Metabolic studies in phenylketonuria. *Journal of Biochemistry*, **31**, 266-269.

Pietz, J., Kreis, R., Boesch, C., Penzien, J., Rating, D. & Herschkowitz, N. (1995) The dynamics of brain concentrations of phenylalanine and its clinical significance in patients with phenylketonuria determined by in vivo <sup>1</sup>H magnetic resonance spectroscopy. *Pediatric Research*, **38**, 657-663.

Pietz, J., Schmidt, E., Matthis, P., Kobialka, B., Kutscha, A. & Sonnevile, L. de (1993) EEGs in phenylketonuria: follow-up to adulthood and short-term diet-related changes in EEGs and cognitive function. *Developmental Medicine and Child Neurology*, **35**, 54-64.

Potocnik, U. & Widhalm, K. (1994) Long-term follow-up of children with classical phenylketonuria after diet discontinuation: a review. *Journal of the American College of Nutrition*, **13**, 232-236.

Pueschel, S.M., Fogelson-Doyle L., Kammerer, B. & Matsumiya, Y. (1983) Neurophysiological, psychological and nutritional investigations during discontinuation of the phenylalanine-restricted diet in children with classic phenylketonuria. *Journal of Mental Deficiency Research*, **27**, 61-67.

Realmuto, G.M., Garfinkel, B.D., Tuchman, M., Tsai, M.Y., Chang, P.N., Fisch, R.O. & Shapiro, S. (1986) Psychiatric diagnosis and behavioral characteristics of phenylketonuric children. *Journal of Nervous and Mental Disease*, **174**, 536-540.

Rey, A. (1941) L'examen psychologique. *Archives de Psychologie*, **28**, 112-164.

Rosvold, H.E., Mirsky, A.F., Sarason, I., Bransome, E.D. & Beck, L.H. (1956) A continuous performance test of brain damage. *Journal of Consulting and Clinical Psychology*, **20**, 343-350.

Rupp, A. & Burgard, P. (1995) Comparison of different indices of dietary control in phenylketonuria. *Acta Paediatrica*, **84**, 521-527.

Russell, F.F., Mills, B.C. & Zucconi, T. (1988) Relationship of parental attitudes and knowledge to treatment adherence in children with PKU. *Pediatric Nursing*, **14**, 514-516.

Rutter, M., Tizard, J., & Whitmore, K. (1970) *Education, Health and Behaviour*. 1970. London: Longman. Pp. 412-421.

Rylance, G. (1989) Outcome of early detected and early treated phenylketonuria patients. *Postgraduate Medical Journal*, **65** (Supplement 2), S7-S9.

Rylance, G., Asplin, D., Hall, S.K., Denmeade, R., Wills, D. & Booth, I.W. (1996) Variability in plasma phenylalanine concentrations over 24 hours in PKU. *Inborn Errors Review Series*, **6**, 5 (abstract). Liverpool: SHS International.

Salway, J.G. (1994) *Metabolism at a Glance*. Oxford: Blackwell. Chapter 22.

Sattler, J.M. (1974) *Assessment of Children's Intelligence*. Philadelphia: Saunders. Pp. 236-240.

Sattler, J.M. (1982) *Assessing Children's Intelligence and Special Abilities, 2nd Edition*. Boston: Allyn & Bacon. Pp. 161-163.

Saudubray, J.-M., Rey, F., Ogier, H., Abadie, V., Farriaux, J.P., Ghisolfi, J., Guibaud, P, Rey, J. & Vidailhet, M. (1987) Intellectual and school performances in early-treated classical PKU patients. *European Journal of Pediatrics*, **146** (Supplement 1), A20-A22.

Schachar, R., Tannock, R. & Logan, G. (1993) Inhibitory control, impulsiveness and attention deficit hyperactivity disorder. *Clinical Psychology Review*, **13**, 721-739.

Scharre, J.E., Cotter, S.A., Block, S.S. & Kelly, S.A. Normative contrast sensitivity data for young children. *Optical and Visual Science*, 1990, **67**, 826-832.

Schmidt, H., Burgard, P., Pietz, J. & Rupp, A. (1996) Intelligence and professional career in young adults treated early for phenylketonuria. *European Journal of Pediatrics*, **155** (Supplement 1), S97-S100.

Schmidt, H., Mahle., M., Michel, U., & Pietz, J. (1987) Continuation versus discontinuation of low-phenylalanine diet in PKU adolescents. *European Journal of Pediatrics*, **146** (Supplement 1), A17-A19.

Schmidt, E., Rupp, A., Burgard, P., Pietz, J., Weglage, J. & Sonnevile, L. de (1994) Sustained attention in adult phenylketonuria: the influence of the concurrent phenylalanine blood-level. *Journal of Experimental and Clinical Neuropsychology*, **16**, 681-688.

Schor, D.P. (1983) PKU and temperament: rating children 3 through 7 years old in PKU families. *Clinical Pediatrics*, **22**, 807-811.

Schuett, V.E. (1996) British and European guidelines for the treatment of PKU. *Word Wide Web*, [http:// www.wolfenet.com/~kronmal/guide.htm](http://www.wolfenet.com/~kronmal/guide.htm).

Schuett, V.E. & Brown, E.S. (1984) Diet policies of PKU clinics in the United States. *American Journal of Public Health*, **74**, 501-503.

Schuett, V.E., Brown, E.S. & Michals, K. (1985) Reinstitution of diet therapy in PKU patients from twenty-two US clinics. *American Journal of Public Health*, **75**, 39-42.

Schuett, V.E., Gurda, R.F. & Brown, E.S. (1980) Diet discontinuation policies and practices of PKU clinics in the United States. *American Journal of Public Health*, **70**, 498-503.

Scriver, C.R., Byck, S., Prevost, L. & Hoang, L. (1996) The phenylalanine hydroxylase locus: a marker for the history of phenylketonuria and human genetic diversity. *CIBA Foundation Symposia*, **197**, 73-90.

Scriver, C.R. & Clow, C.L. (1980) Phenylketonuria: epitome of human biochemical genetics. *New England Journal of Medicine*, **303**, 1336-1342.

Seashore, C.E., Lewis, D. & Saetveit, D.L. (1960) *Seashore Measures of Musical Talents*. New York: The Psychological Corporation.

Seashore, M.R., Friedman, E., Novelly, R.A. & Bapat, V. (1985) Loss of intellectual function in children with phenylketonuria after relaxation of dietary phenylalanine restriction. *Pediatrics*, **75**, 226-232.

Shah, S.N., Peterson, N.A. & McKean, C.M. (1972) Abnormalities of myelin synthesis in phenylketonuria. *Journal of Neurochemistry*, **19**, 2369-2376.

Shaw, D.W.W., Weinberger, E. & Maravilla, K.R. (1990) Cranial MR in phenylketonuria. *Journal of Computer Assisted Tomography*, **14**, 458-460.

Slijper, F.M.E., Huisman, J., Hendriks, M.M.T., Kalverboer, A.F. & Schot, L. van der (1988) The effects of diet discontinuation in children with phenylketonuria. *Archives of Disease in Childhood*, **63**, 445.

Smith, I. (1985) The hyperphenylalaninaemias. In, I. Smith, J.K. Lloyd & C.R. Scriver (Eds.), *Genetic and Metabolic Disease in Paediatrics*. London: Butterworths. Pp. 166-210.

Smith, I. & Beasley, M. (1989) Intelligence and behaviour in children with early-treated phenylketonuria. *European Journal of Clinical Nutrition*, **43**, 1-5.

Smith, I., Beasley, M.G. & Ades, A.E. (1990a) Effect on intelligence of relaxing the low phenylalanine diet in phenylketonuria. *Archives of Disease in Childhood*, **65**, 311-316.

Smith, I., Beasley, M.G. & Ades, A.E. (1990b) Intelligence and quality of dietary treatment in phenylketonuria. *Archives of Disease in Childhood*, **65**, 472-478.

Smith, I., Beasley, M.G., Wolff, O.H. & Ades, A.E. (1988) Behaviour disturbances in 8-year-old children with early-treated phenylketonuria. *Journal of Paediatrics*, **112**, 403-408.

Smith, I. & Brenton, D.P. (1995) Hyperphenylalaninaemias. In, J. Fernandes, J.-M. Saudubray & G. Van Der Berghe (Eds.), *Inborn Metabolic Diseases, 2nd Edn.* Berlin: Springer-Verlag. Pp. 147-160.

Smith, I., Glossop, J. & Beasley, M. (1990) Foetal damage due to maternal phenylketonuria. *Journal of Inherited Metabolic Disease*, **13**, 651-657.

Smith, I., Lobascher, M., Stephenson, J.E., Wolff, O.H., Schmidt, H., Grubel-Kaiser, S. & Bickel, H. (1978) Effect of stopping low-phenylalanine diet on intellectual progress of children with phenylketonuria. *British Medical Journal*, **2**, 723-726.

Smith, I., Lobascher, M. & Wolff, O.H. (1973) Factors influencing outcome in early-treated phenylketonuria. In, J.W.T. Seakins (Ed.), *Treatment of Inborn Errors of Metabolism.* Edinburgh: Livingstone. Pp. 41-49.

Smith, I. & Wolff, O.H. (1978) *MRC/DHSS Phenylketonuria Register, Newsletter Number 5.*

Sonneville, L.M.J. de (1991) The importance of dietary control of early-treated phenylketonuria in relation to sustained and complex information processing. *Proceedings of the 5th Annual ESPKU Meeting, de Haan, Belgium.* Pp. 6-9.

Sonneville, L.M.J. de, (1993) SVAT: a computer-based approach to development and disorders of information processing. In, F.J. Maarse, A.E. Akkerman, N. Brand, L.J.M. Mulder & M. van der Stelt (Eds.), *Computers in Psychology.* Lisse: Swets & Zeitlinger. Pp. 268-276.

Sonneville, L.M.J. de, Schmidt, E., Michel, U. & Batzler, U. (1990) Preliminary neuropsychological test results. *European Journal of Pediatrics*, **149** (Supplement 1), S39-S44.

Spreen, O. & Benton, A. (1969) *Neurosensory Center Comprehensive Examination for Aphasia.* Victoria, BC: University of Victoria Press.

Stellungnahme der Arbeitsgemeinschaft für Padiatrische Stoffwechselstörungen (APS) zur diätetischen Behandlung der Phenylketonurie. (1990) *Monatsschrift Kinderheilkunde*, **138**, 636.

Stemerding, B.A., Molen, M.W. van der, Kalverboer, A.F., Meere, J.J. van der, Hendrix, M.W., Huisman, J., Schot, L.W.A. van der & Slijper, F.M.E. (1994) Information processing deficits in children with early and continuously treated phenylketonuria? *Acta Paediatrica*, **83** (Supplement 407), 106-107.

Stroop, J.R. (1935) Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, **18**, 643-662

Stuart-Hamilton, I. (1996) *Dictionary of Psychological Testing, Assessment and Treatment*. London: Jessica Kingsley.

Stuss, D. (1987) The neuropsychology of the frontal lobes. *BNI Quarterly*, **3**, 28-33.

Stuss, D. & Benson, D.F. (1986) *The Frontal Lobes*. New York: Raven.

Sullivan, D. (1996) Experiences with adult PKU. *Paper presented at 10th Annual ESPKU Meeting, Aegerisee, Switzerland*.

Taylor, H.G. (1988) Neuropsychological testing: relevance for assessing children's learning disabilities. *Journal of Consulting and Clinical Psychology*, **56**, 795-800.

Thompson, A.J., Smith, I., Brenton, D. & Youl, B.D. (1990) Neurological deterioration in young adults with phenylketonuria. *Lancet*, **336**, 602-605.

Thompson, A.J., Smith, I., Kendall, B.E., Youl, B.D. & Brenton, D. (1991) Magnetic resonance imaging changes in early-treated patients with phenylketonuria. *Lancet*, **337**, 1224-1225.

Thompson, A.J., Tillotson, S., Smith, I., Kendall, B., Moore, S.G. & Brenton, D.P. (1993) Brain MRI Changes in phenylketonuria - associations with dietary status. *Brain*, **116**, 811-821.

Thompson, R.J. & O'Quinn, A.N. (1979) *Developmental Disabilities*. New York: Oxford University Press. Pp. 33-36.

Toft, P.B., Lou, H.C., Kragheloh-Mann, I., Andersen, J., Guttler, F., Guldborg, P. & Henriksen, O. (1994) Brain magnetic resonance imaging in children with optimally controlled hyperphenylalaninaemia. *Journal of Inherited Metabolic Disease*, **17**, 575-583.

Tyfield, L.A., Osborn, M.J. & Holton, J.B. (1991) Molecular heterogeneity at the phenylalanine hydroxylase locus in the population of the south-west of England. *Journal of Medical Genetics*, **28**, 244-247.

Vasta, R. (1979) *Studying Children*. San Francisco: Freeman. Chapters 4 & 5.

Villasana, D., Butler, I.J., Williams, J.C. & Roonta, S.M. (1989) Neurological deterioration in adult phenylketonuria. *Journal of Inherited Metabolic Disease*, **12**, 451-457.

Vorhees, C.V., Butcher, R.E., & Berry, H.K. (1981) Progress in experimental phenylketonuria: a critical review. *Neuroscience and Biobehavioural Reviews*, **5**, 177-190

Waisbren, S.E., Brown, M.J., Sonnevile, L.M.J. de, & Levy, H.L. (1994) Review of neuropsychological functioning in treated phenylketonuria. *Acta Paediatrica*, **83** (Supplement 407), 98-103.

Waisbren, S.E., Mahon, B.E., Schnell, R.R. & Levy, H.L. (1987) Predictors of intelligence quotient and intelligence quotient change in persons treated for phenylketonuria in early life. *Pediatrics*, **79**, 351-355.

Waisbren, S.E., Schnell, R.R. & Levy, H.L. (1980) Diet termination in children with phenylketonuria: a review of psychological assessments used to determine outcome. *Journal of Inherited Metabolic Disease*, **3**, 149-153.

Walsh, K. W. (1978) *Neuropsychology: a clinical approach*. Edinburgh: Churchill Livingstone. Pp. 172-187.

Walter, J.H. (1995) Late effects of phenylketonuria. *Archives of Disease in Childhood*, **73**, 485-486.

Wechsler, D. (1987) *Wechsler Memory Scale-Revised*. London: The Psychological Corporation.

Wechsler, D. (1992) *Wechsler Intelligence Scale for Children-Third UK Edition*. London: The Psychological Corporation.

Weglage, J., Funders, B., Ullrich, K., Rupp, A. & Schmidt, E. (1996a) Psychosocial aspects in phenylketonuria. *European Journal of Pediatrics*, **155** (Supplement 1), S101-S104.

Weglage, J., Funders, B., Wilken, B., Schubert & Ullrich, K. (1993) School performance and intellectual outcome in adolescents with phenylketonuria. *Acta Paediatrica*, **81**, 582-586.

Weglage, J., Pietsch, M., Funders, B., Koch, H.G., & Ullrich, K. (1996b) Deficits in selective and sustained attention processes in early-treated children with phenylketonuria. *European Journal of Pediatrics*, **155**, 200-204.

Weglage, J., Rupp, A. & Schmidt, E. (1994) Personality characteristics in patients with phenylketonuria treated early. *Pediatric Research*, **35**, 611-613.

Weglage, J., Ullrich, K., Pietsch, M., Funders, B., Zass, R. & Koch, H.G. (1996c) Untreated non-phenylketonuric-hyperphenylalaninaemia: intellectual and neurological outcome. *European Journal of Pediatrics*, **155** (Supplement 1), S26-S28.

Welsh, M.C. (1996) A prefrontal model of early-treated phenylketonuria. *European Journal of Pediatrics*, **155** (Supplement 1), S87-S89.

Welsh, M.C. & Pennington, B.F. (1988) Assessing frontal lobe functioning in children. *Developmental Neuropsychology*, **4**, 199-230.

Welsh, M.C., Pennington, B.F., Ozonoff, S., Rouse, B. & McCabe, E.R.B. (1990) Neuropsychology of early-treated phenylketonuria: specific executive function deficits. *Child Development*, **61**, 1697-1713.

Williamson, M.L., Koch, R., Azen, C. & Chang, C. (1981) Correlates of intelligence test results in treated phenylketonuric children. *Pediatrics*, **68**, 161-167.

Williamson, M.L., Koch, R., Azen, C.G. & Friedman, E.G. (1985) Diet termination: matched pair control between PKU children and their non-PKU siblings. In, H. Bickel & U. Wachtel (Eds.), *Inherited Diseases of Amino Acid Metabolism*. Stuttgart: Thieme. Pp. 150-160.

Wilson, B., Cockburn, J., & Baddeley, A. (1991) *Manual of The Rivermead Behavioural Memory Test*. Bury St. Edmunds: Thames Valley Test Company.

Wolff, O.H. (1981) Routine neonatal screening for phenylketonuria in the United Kingdom 1964-78. *British Medical Journal*, **282**, 1680-1684.

Woo, S.L.C. (1989) Molecular basis and population genetics of phenylketonuria. *Biochemistry*, **28**, 1-7.

Wood, A.C., Friedmann, C.J & Steisel, I.M. & (1967) Psychological factors in phenylketonuria. *American Journal of Orthopsychiatry*, **37**, 671-679.

- Woolf, L.I. (1967) A critical assessment of the dietary treatment of phenylketonuria. In, *Phenylketonuria and Allied Metabolic Diseases*. Washington, DC: US Department of Child Health, US Government Printing Office.
- Woolf, L.I. (1979) Late onset phenylalanine intoxication. *Journal of Inherited Metabolic Disease*, **2**, 19-20.
- Woolf, L.I., Griffiths, R. & Moncrieff, A. (1955) Treatment of phenylketonuria with a diet low in phenylalanine. *British Medical Journal*, **1**, 57-64.
- Yakovlev, L. & Lecours, A.R. (1967) The myelogenetic cycles of regional maturation of the brain. In, A. Minkowski (Ed.), *Regional Development of the Brain in Early Life*. Oxford: Blackwell. Pp. 3-70.
- Yaremko, R.M., Harari, H., Harrison, R.C. & Lynn, E. (1982) *Reference Handbook of Research and Statistical Methods in Psychology*. New York: Harper & Row. Pp. 199.
- Zametkin, A.J., Karoum, F. & Rapoport, J.L. (1987) Treatment of hyperactive children with D-phenylalanine. *American Journal of Psychiatry*, **144**, 792-794.
- Zangwill, O. (1946) Some clinical applications of the Rey-Davis performance test. *Journal of Mental Science*, **92**, 19-34.
- Zetterstrom, R. (1995) Editorial comment on phenylketonuria. *Acta Paediatrica*, **84**, 716-718.

## APPENDICES

### Appendix One

#### Chapter Three - Post-treatment hyperphenylalaninaemia

#### RAW DATA - PHENYLALANINE IDCs ( $\mu\text{mol/l}$ )

S (Gp)	Year of Life										Post Diet
	1	2	3	4	5	6	7	8	9	10	
<b>PKU young, on-diet</b>											
1	480	240	360	240	420	360	240	480	360	360	
2	360	660	480	360	540	300	540				
3	300	240	120	240	300	120	240				
4	360	180	120	120	120						
5	240	120	120	240	120	120	240	240	120		
6	600	300	240	360	480	240					
7	480	360	360	240	360	480	180	480	600		
8	360	360	540	540	360	480					
9	300	120	240	360	360	360	360	480	360	360	
10	480	240	240	240	360	360	600	360	420		
<b>PKU old, off-diet</b>											
1	480	240	260	300	280	260	440	520	520	540	1014
2	240	240	260	300	280	240	480	480	480	480	773
3	300	360	300	360	360	360	360	480	360	600	791
4	100	100	120	150	120	120	240	240	240	480	720
5	280	240	240	300	280	260	440	520	510	540	900
6	480	240	240	360	360	120	480	480	480	480	1174
7	300	240	240	180	120	180	240	240	360	480	900
8	240	240	240	240	360	360	240	360	480	600	1200
9	240	240	480	480	300	360	600	660	480	480	1204
10	290	240	260	290	270	460	900	1200	1200	720	1200

## Appendix Two

### Chapter Three - Post-treatment hyperphenylalaninaemia

#### RAW DATA - SEX, AGE & NEUROPSYCHOLOGICAL TESTS

S (Gp)	Sex (M/F)	Age (yrs)	React Time (secs)	Peg Trans (secs)	Match Figs (0-10)	Letter Canc (n/min)	Verb Flu (n/3m)	Des Flu (n/3m)	Rey Verb (0-75)	Rey Lab (0-20)
<b>PKU young, on-diet</b>										
1	M	9.75	2.11	63	7	5.00	29	12	54	13
2	M	6.67	3.31	86	4	2.00	12	8	30	12
3	M	6.42	3.31	114	5	3.73	8	16	26	4
4	M	5.58	3.83	109	3	3.00	11	13	45	17
5	M	8.83	2.56	77	7	3.25	28	14	37	15
6	M	5.58	3.47	114	4	2.50	13	12	39	5
7	M	8.50	2.53	68	4	3.00	18	13	41	9
8	F	6.17	5.28	94	4	2.50	12	8	47	12
9	F	9.42	2.95	67	7	3.25	33	10	46	4
10	F	8.42	3.78	93	4	3.75	24	13	59	14
<b>CONTROL young</b>										
11	M	9.75	2.97	66	7	3.75	20	7	55	18
12	M	6.75	3.32	77	3	2.50	16	13	33	6
13	M	6.33	2.81	74	6	2.00	20	13	39	9
14	M	5.75	2.82	85	3	3.50	9	12	45	8
15	M	8.75	2.46	61	6	4.75	23	15	47	18
16	M	5.67	3.35	93	4	1.25	5	10	39	8
17	M	8.50	2.09	63	7	6.39	29	14	51	18
18	F	6.17	3.90	89	4	3.00	12	11	32	12
19	F	9.33	2.56	59	4	3.75	38	13	49	17
20	F	8.42	3.25	60	4	4.25	18	14	40	17
<b>PKU old, off-diet</b>										
21	M	22.67	2.52	49	8	5.70	26	18	49	15
22	M	18.83	2.40	48	6	4.50	37	20	50	9
23	M	17.25	2.74	48	9	10.00	28	12	43	15
24	M	28.42	3.27	47	6	8.20	42	19	45	17
25	M	21.00	3.75	47	5	15.50	27	17	43	18
26	M	22.33	1.70	46	6	5.00	32	16	51	9
27	M	13.67	2.59	54	9	8.43	39	25	53	14
28	F	13.58	1.57	60	8	4.75	48	17	61	19
29	F	20.17	1.85	58	7	3.75	42	17	41	19
30	F	23.67	1.32	58	8	7.81	17	21	31	10
<b>CONTROL old</b>										
31	M	21.75	2.56	38	9	4.50	46	26	60	18
32	M	17.75	1.39	48	9	12.90	48	27	49	16
33	M	17.25	1.80	46	9	8.09	32	22	53	19
34	M	27.92	2.79	47	10	19.60	41	22	55	16
35	M	21.42	1.87	55	10	4.75	46	19	53	18
36	M	21.42	2.77	50	10	12.74	43	27	61	18
37	M	13.67	2.06	46	8	8.69	22	28	50	19
38	F	13.58	2.20	48	6	3.00	28	20	42	17
39	F	19.67	2.63	41	8	9.95	42	28	68	17
40	F	24.17	1.96	50	9	13.33	41	26	60	18

### **Appendix Three**

#### **Chapters Six and Seven - Experimental phe manipulation**

#### **RAW DATA - SEX & AGE**

<b>S</b>	<b>Gp</b>	<b>Sex (M/F)</b>	<b>Age (yrs)</b>
1	1	F	12.17
2	1	F	11.00
3	1	F	16.08
4	1	M	11.25
5	1	M	12.17
6	1	M	16.42
7	1	M	16.25
8	1	M	11.00
9	1	M	13.75
10	1	F	10.83
11	2	M	11.25
12	2	M	10.83
13	2	F	10.92
14	2	M	14.67
15	2	F	11.63
16	2	M	11.33

## **Appendix Four**

### **Chapters Six and Seven - Experimental phe manipulation**

#### **RAW DATA - PHENYLALANINE IDCs ( $\mu\text{mol/l}$ )**

---

<b>S</b>	<b>Gp</b>	<b>Year of Life</b>									
		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>
1	1	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	360	420
2	1	360	240	480	360	600	720	720	900	900	900
3	1	480	480	360	180	360	480	360	360	360	240
4	1	240	120	120	240	120	120	240	240	120	120
5	1	480	240	360	240	420	360	240	480	360	360
6	1	480	360	480	360	360	600	360	480	480	750
7	1	120	240	360	360	360	480	360	480	480	720
8	1	480	360	360	240	360	480	180	480	600	540
9	1	240	360	720	420	420	600	900	540	600	900
10	1	480	240	240	240	360	360	600	360	420	480
11	2	360	240	240	360	600	480	420	600	480	480
12	2	240	240	360	480	600	900	600	660	360	299
13	2	360	360	810	720	480	600	540	660	600	720
14	2	300	120	240	360	360	360	360	480	360	360
15	2	300	120	240	360	360	360	360	480	360	360
16	2	240	120	360	360	360	240	360	360	480	360

---

**Appendix Five**

**Chapters Six and Seven - Experimental phe manipulation**

**RAW DATA - PHE CONCENTRATIONS x PHASE ( $\mu\text{mol/l}$ )**

---

<b>S</b>	<b>Gp</b>	<b>Phase</b>		
		<b>Baseline</b>	<b>Low</b>	<b>High</b>
1	1	472	614	844
2	1	1217	1303	1875
3	1	1064	1013	1285
4	1	209	386	1345
5	1	306	343	1558
6	1	1008	1067	1516
7	1	999	1074	1294
8	1	1030	1109	1035
9	1	1236	1047	1233
10	1	687	760	1560
11	2	471	396	1087
12	2	507	868	902
13	2	300	344	998
14	2	459	455	741
15	2	506	745	1385
16	2	258	282	933

---

## Appendix Six

### Chapter Six - Experimental phe manipulation

#### RAW DATA - NEUROPSYCHOLOGICAL TESTS

S	Gp	Matching Figures (0-10)			Rey Verbal (0-60)			Digits (0-16)			Paired Associates (0-30)		
		Base	Low	High	Base	Low	High	Base	Low	High	Base	Low	High
1	1	6	6	6	41	55	52	8	9	9	28	24	29
2	1	3	4	6	52	57	56	8	8	8	28	29	24
3	1	7	7	8	52	60	59	11	11	10	29	29	26
4	1	7	8	8	54	55	50	8	8	10	30	30	30
5	1	7	7	6	49	58	51	11	12	10	28	27	27
6	1	8	8	8	54	52	57	11	13	13	28	28	24
7	1	7	8	8	49	55	56	9	9	9	26	25	25
8	1	5	5	5	46	55	43	11	11	11	25	21	23
9	1	6	7	8	50	50	56	11	10	11	25	28	28
10	1	7	8	7	51	59	55	13	13	12	26	28	27
11	2	5	4	6	27	36	42	6	8	5	28	29	25
12	2	7	6	6	47	50	50	8	8	7	23	21	26
13	2	6	6	3	47	54	54	9	9	8	25	26	27
14	2	6	5	4	43	47	44	10	7	7	28	29	28
15	2	4	7	6	49	52	55	10	8	11	20	26	23
16	2	6	8	7	49	57	58	14	12	13	28	28	23

S	Gp	Block Tapping (0-16)			Rey Labyrinth (0-20)			Pegboard (secs)			Steadiness (0--)		
		Base	Low	High	Base	Low	High	Base	Low	High	Base	Low	High
1	1	6	7	5	18	19	16	152	153	162	31	1	9
2	1	8	10	6	17	18	16	164	162	169	10	18	18
3	1	11	11	11	19	19	20	146	129	131	21	5	3
4	1	7	6	7	20	15	17	157	168	169	40	34	40
5	1	6	6	6	15	15	15	153	141	159	32	11	14
6	1	9	10	8	6	13	18	174	176	172	14	26	17
7	1	9	10	8	18	20	16	167	159	169	14	9	16
8	1	8	7	8	18	18	19	167	156	171	16	11	17
9	1	8	9	6	17	9	10	160	157	154	19	11	14
10	1	7	7	7	17	15	16	164	158	164	15	19	31
11	2	9	9	8	15	20	11	159	180	180	29	13	35
12	2	7	7	7	16	18	9	167	160	160	15	18	33
13	2	5	7	3	7	13	7	143	146	146	10	4	9
14	2	8	6	10	19	14	17	151	154	146	14	17	15
15	2	9	8	9	13	17	19	139	131	131	17	18	13
16	2	10	10	9	20	20	14	167	154	154	11	11	43

## Appendix Seven

### Chapter Seven - Experimental phe manipulation

#### RAW DATA - WISC, Rivermead Memory & Problem Behaviours

S	Gp	Similarities		Vocabulary		Block Design		Object Assembly		Rivermead BMT		Problem Behaviours	
		(1-19)		(1-19)		(1-19)		(1-19)		(0-90)		(0-80)	
		Low	High	Low	High	Low	High	Low	High	Low	High	Low	High
1	1	10	10	9	8	9	10	13	13	67.5	62.5	2	3
2	1	13	12	11	10	9	7	8	17	56.5	59.0	25	58
3	1	12	19	10	15	14	15	12	16	77.5	80.0	14	2
4	1	12	12	10	11	11	11	13	16	66.5	63.0	2	15
5	1	13	12	10	11	12	11	12	14	67.5	70.5	5	39
6	1	11	11	11	9	9	6	10	8	72.5	57.5	1	4
7	1	11	10	8	5	10	5	11	11	66.0	62.5	14	3
8	1	10	10	8	8	10	8	10	11	60.0	64.0	4	4
9	1	9	8	9	9	9	9	7	10	58.5	61.0	0	8
10	1	16	15	12	12	12	9	12	9	78.0	85.0	5	22
11	2	9	8	7	6	11	13	11	13	51.0	45.5	6	30
12	2	13	10	10	5	10	9	9	4	64.0	59.0	3	3
13	2	10	10	9	9	5	5	11	10	64.5	61.5	10	8
14	2	8	9	5	6	10	10	13	9	59.5	59.0	4	3
15	2	13	10	8	8	8	10	11	10	52.5	61.0	6	9
16	2	14	14	12	11	8	10	9	13	67.0	71.0	5	11

**Appendix Eight**

**Chapter Eight - Executive Dysfunction, Correlational Study**

**RAW DATA - Sex and Age**

---

<b>S</b>	<b>Sex (M/F)</b>	<b>Age (Yrs)</b>
1	M	12.92
2	M	12.50
3	M	10.75
4	M	10.50
5	F	12.67
6	M	10.50
7	M	10.83
8	M	12.83
9	F	10.25
10	M	13.83
11	F	13.58
12	F	12.58
13	M	13.00
14	M	12.67
15	F	12.58

---

**Appendix Nine**

**Chapter Eight - Executive Dysfunction, Correlational Study**

**RAW DATA - Annual Indices of Dietary Control**

---

**IDC ( $\mu\text{mol/l}$ )**

---

<b>S</b>	<b>Yr 1</b>	<b>Yr 2</b>	<b>Yr3</b>	<b>Yr 4</b>	<b>Yr 5</b>	<b>Yr 6</b>	<b>Yr 7</b>	<b>Yr 8</b>	<b>Yr9</b>
1	360	240	240	360	600	504	420	600	480
2	240	240	360	480	600	900	600	660	360
3	360	120	360	480	600	720	900	600	822
4	360	180	210	300	480	600	600	540	331
5	360	240	480	360	600	720	720	900	900
6	360	360	240	240	480	300	360	272	298
7	240	600	600	360	480	240	540	360	408
8	240	120	120	240	120	120	240	240	120
9	360	360	360	540	360	540	480	562	523
10	480	240	360	240	420	420	240	480	360
11	300	120	240	360	360	360	360	480	360
12	360	360	810	720	480	600	540	660	600
13	240	120	360	360	360	240	360	360	480
14	480	360	360	240	360	480	180	480	600
15	480	240	240	240	360	420	600	240	420

---

## **Appendix Ten**

### **Chapter Eight - Executive Dysfunction, Correlational Study**

#### **RAW DATA - Cognitive Variables**

---

<b>S</b>	<b>Mazes (1-19)</b>	<b>Stroop (0-36)</b>	<b>Cancel (Hits/m)</b>	<b>Des Flu (Des/3m)</b>	<b>Digits (1-19)</b>	<b>Benton (0-20)</b>	<b>Seashore (0-30)</b>	<b>Rey Verb (0-60)</b>
1	12	31	10.0	15	6	20	24	47
2	7	31	10.5	20	6	18	23	46
3	12	34	8.0	14	9	20	26	45
4	5	28	7.5	12	7	13	19	34
5	6	31	7.5	15	10	15	21	41
6	17	31	6.5	16	8	12	21	45
7	10	35	10.0	14	6	19	26	44
8	7	34	11.5	18	10	20	25	51
9	7	36	8.0	12	12	19	28	40
10	11	31	12.0	20	9	17	28	42
11	6	35	11.0	18	10	18	25	45
12	9	33	12.0	12	7	18	26	40
13	12	32	3.5	26	19	15	26	37
14	8	32	11.0	13	9	15	22	43
15	8	32	6.5	21	19	15	29	48

---

**Appendix Eleven**

**Chapter Eight - Executive Dysfunction, Correlational Study**

**RAW DATA - Personality Variables**

---

**Devereux Scales of Mental Disorders (T-Scores)**

---

<b>S</b>	<b>Dev</b>	<b>Tot</b>	<b>Conduct</b>	<b>Attent</b>	<b>Anxiety</b>	<b>Depress</b>	<b>Autism</b>	<b>Acute</b>	<b>Pr</b>
1	59	54	60	48	64	65	58		
2	40	40	40	40	40	40	40	45	
3	45	42	58	41	41	44	44	49	
4	59	56	65	60	48	65	65	51	
5	51	58	52	48	58	44	44	46	
6	42	48	44	40	40	44	44	45	
7	46	49	47	49	43	45	45	49	
8	49	45	47	53	62	44	44	45	
9	42	44	47	41	40	47	47	46	
10	46	46	48	56	47	41	41	44	
11	42	44	44	40	40	50	50	45	
12	66	64	76	49	92	52	52	46	
13	46	55	43	40	43	59	59	44	
14	52	55	54	47	58	51	51	45	
15	42	44	40	47	45	41	41	46	

---

*Appendix Twelve*

**Chapter Nine - Executive Dysfunction, Control Group Study**

**RAW DATA - Sex, Age, British Picture Vocabulary Scales, Digits Forwards and Deprivation Category**

---

<b>S</b>	<b>Sex (M/F)</b>	<b>Age (Yrs)</b>	<b>BPVS (Std Sc)</b>	<b>Digits (0-40)</b>	<b>DEPCAT (0-7)</b>
----------	----------------------	----------------------	--------------------------	--------------------------	-------------------------

---

**PKU Group**

1	M	11.92	125	22	3
2	M	10.83	126	28	3
3	F	11.42	104	23	3
4	M	9.08	100	18	4
5	F	8.83	101	18	4
6	M	7.25	111	20	6
7	M	7.08	107	16	4
8	F	10.75	93	18	6
9	F	7.33	85	15	6
10	F	5.11	117	11	4
11	M	6.42	82	14	4

---

**Control Group**

1	M	12.75	124	26	2
2	F	10.42	133	27	2
3	M	11.58	106	22	3
4	M	9.17	95	26	3
5	F	8.83	100	15	3
6	M	7.17	103	18	3
7	M	7.25	95	15	3
8	F	10.67	87	24	3
9	F	6.83	81	18	3
10	M	6.17	102	16	3
11	M	6.42	90	23	3

---

**Appendix Thirteen**

**Chapter Nine - Executive Dysfunction, Control Group Study**

**RAW DATA - PKU Group Indices of Dietary Control (Years 1 to 5)  
and Test-Concurrent Phenylalanine Concentration**

---

Phenylalanine ( $\mu\text{mol/l}$ )						
S	Yr 1	Yr 2	Yr 3	Yr 4	Yr 5	Conc
1	360	600	480	420	480	501
2	480	240	120	120	120	193
3	360	360	360	480	480	544
4	660	120	240	120	120	401
5	480	240	300	420	360	475
6	480	660	480	600	685	562
7	120	240	120	180	139	240
8	360	360	390	240	240	437
9	240	240	120	120	133	336
10	420	120	306	124	284	312
11	360	240	240	180	195	269

---

**Appendix Fourteen**

**Chapter Nine - Executive Dysfunction, Control Group Study**

**RAW DATA - Continuous Performance Test, One-Back Version,  
Hits, False Alarms and Reaction Time per Trial Block**

S	Hits (0-20)				False Alarms (0-20)				Reaction Time (Centisec)			
	B1	B2	B3	B4	B1	B2	B3	B4	B1	B2	B3	B4
<b>PKU Group</b>												
1	19	20	18	18	1	1	1	1	41	43	43	48
2	19	20	20	20	0	0	1	0	37	39	35	36
3	20	20	20	20	0	1	0	0	37	38	40	40
4	20	19	20	17	1	1	1	5	46	49	52	47
5	17	18	16	17	1	5	2	2	60	66	52	53
6	14	16	15	10	1	1	2	5	51	60	67	66
7	17	19	19	18	3	4	9	9	86	53	55	45
8	20	19	20	20	0	0	0	1	48	46	44	47
9	18	20	15	17	3	5	2	0	49	51	57	50
10	18	18	16	19	1	3	1	2	48	52	46	51
11	18	17	16	17	8	1	3	4	55	53	54	54
<b>Control Group</b>												
1	20	20	20	20	0	0	0	0	37	37	37	37
2	20	20	20	20	0	0	0	0	41	41	41	42
3	17	20	20	20	13	12	10	10	37	39	46	40
4	20	20	20	20	2	1	2	2	37	34	34	36
5	18	16	16	18	5	0	1	1	41	40	49	43
6	14	16	15	20	2	5	0	2	43	47	39	42
7	19	15	19	17	1	3	7	1	48	55	39	40
8	20	19	19	17	0	3	3	3	41	45	44	44
9	15	13	9	12	5	4	1	1	52	59	53	61
10	13	14	13	14	15	16	16	15	59	69	64	65
11	18	16	18	17	2	2	2	6	62	54	58	56

**Appendix Fifteen**

**Chapter Nine - Executive Dysfunction, Control Group Study**

**RAW DATA - Continuous Performance Test, Two-Back Version,  
Hits, False Alarms and Reaction Time per Trial Block**

S	Hits (0-20)				False Alarms (0-20)				Reaction Time (Centisec)			
	B1	B2	B3	B4	B1	B2	B3	B4	B1	B2	B3	B4
<b>PKU Group</b>												
1	18	17	17	19	3	0	2	1	38	40	36	39
2	20	19	18	17	0	1	2	0	41	44	47	43
3	20	19	16	19	8	1	4	1	37	49	45	53
4	20	18	16	16	1	0	3	2	51	52	54	63
5	16	14	13	14	1	1	2	2	54	49	47	56
6	17	18	14	15	2	3	4	6	51	59	69	63
7	18	17	14	15	2	7	5	6	51	55	47	49
8	16	20	15	18	0	0	0	2	51	54	48	49
9	17	13	15	12	2	10	2	2	40	55	47	42
10	14	15	12	13	2	1	6	2	68	56	57	67
11	17	14	11	14	2	0	3	2	55	50	58	54

**Control Group**

1	19	20	20	20	1	0	1	2	36	37	34	37
2	20	19	20	20	0	0	0	1	43	45	38	38
3	19	19	20	17	3	0	2	1	35	36	34	36
4	20	19	20	18	1	2	1	2	30	31	35	36
5	14	15	15	13	8	4	7	7	41	61	49	58
6	11	10	13	9	8	4	8	11	34	40	52	43
7	17	16	16	15	2	5	3	5	47	46	42	48
8	20	19	16	19	0	6	4	4	43	46	59	53
9	10	9	7	10	0	1	5	5	63	72	65	78
10	18	19	18	18	1	3	4	5	48	48	45	48
11	10	8	12	8	11	18	13	15	130	63	55	56