The Genetic Determinants of Lamotrigine Dosing in Epilepsy

Matthew John Grant

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Master of Philosophy

University of Liverpool, Department of Neurological Science,

August 2010

Abstract

Introduction: Epilepsy is a common serious neurological disorder affecting approximately 50 million people worldwide. The mainstay of treatment is the use of antiepileptic drugs (AEDs), to which approximately 70% of people with epilepsy respond.[1, 2]

Lamotrigine (LTG) was first licensed for use in the UK in 1991 and has become one of the most commonly prescribed AEDs in the developed world. It has proven efficacy for various epilepsy types, is employed at first-line treatment for partial-onset and generalised seizures, may be used for atonic, tonic, and atypical absence seizure types, and also has licensing for the treatment of Lennox-Gastaut syndrome.[1, 2] The individual dose of LTG required to achieve seizure freedom varies considerably,[3] which may be at least partially explained by inter-individual variation in pharmacokinetic and pharmacodynamic profiles.

Two proteins appear to be involved in the pharmacokinetics of LTG. UDP-glucuronosyl-transferase 1A4 (encoded by *UGT1A4*) is the enzyme responsible for the hepatic metabolism of LTG and organic cation transporter 1 (OCT1; encoded by *SLC22A1*) has very recently been implicated in the transport of LTG at the level of hepatocytes and the blood-brain-barrier. The pharmacodynamic profile of LTG is characterised by inhibition of neuronal voltage-gated sodium channels. The three predominant sodium channel alpha-subunits in mammalian brain are encoded by the genes *SCN1A*, *SCN2A* and *SCN3A*. We have assessed whether genetic variation in these five genes is associated with the maintenance dose of LTG when successfully employed as monotherapy in people with newly-diagnosed epilepsy.

Methods: A total of 96 individuals with newly diagnosed epilepsy (50% male, mean age 42 years, range 9 to 83 years) who had been seizure-free for at least 12 months on an unchanged dose of LTG (median dose = 200 mg/day, range = 50 to 675 mg/day) were included in the analysis. A total of 173 single nucleotide polymorphisms (SNPs) across five genes (*SCN1A*, *SCN2A*, *SCN3A*, *UGT1A4*, *SLC22A1*) were genotyped and subjected to univariate and multivariate-regression analyses to identify associations with LTG maintenance dose. Multiple demographic factors (including age, gender, epilepsy type, previous AED use, and number of seizures prior to treatment) were also included as co-variates.

Results: Univariate analyses revealed associations between maintenance dose and genotype at fourteen SNP loci, although these did not remain significant after correction for multiple comparisons using false-discovery rate. Multivariate regression analysis generated several models that associated

with the observed variation in LTG maintenance dose. The most explanatory of these accomplished an r^2-value of 0.606 ($p<0.001$) incorporating polymorphisms from *SLC22A1, SCN1A, SCN2A* and *SCN3A*.

Conclusions: Genetic variation in *SLC22A1, SCN3A* and *SCN2A*, and to a lesser extent *SCN1A*, affect LTG maintenance dose requirement. These results require validation in a larger, independent cohort.

Acknowledgements

First of all I would like to thank Dr Graeme Sills and Professor Tony Marson for acting as supervisors for my year of research. They have provided me with invaluable guidance throughout the MPhil process, and for this I cannot thank them enough.

I would like to thank all members of the Neuroscience Research Team who have made this year possible; thanks to Pete Dixon for his assistance in data collection, Andrea Jorgensen for her input into the statistical analysis, and Professor Munir Pirmohammed for the use of the laboratory in the Wolfson Centre, University of Liverpool.

I would especially like to thank Kanvel Shazadi for all her help during the past twelve-months. Kanvel has supported me throughout this project, and without her continued input the process would have been a lonely and far less enjoyable one.

Finally I would like to thank my parents, family and friends for their support throughout the undertaking of this project.

A big thank you to everyone involved.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1. Epilepsy

The signs and symptoms of epilepsy have been described by scholars since ancient times, with the word 'epilepsy' derived from the Greek verb **ελαμβανειν** (epilamvanein) meaning "to be seized", "to be taken hold of" or "to be attacked".[4]

One of the earliest descriptions of epilepsy was recorded over 3,000 years ago, and was attributed to actions of the god of the moon.[5] Since then, beliefs surrounding the cause of epilepsy have developed from notions such as 'divine punishment for sinners' by the ancient Greeks, to the first formal description of epilepsy as a disease by Hippocrates some 2,500 years ago.[6]

Theories regarding divine aetiologies have waxed and waned over time. However, by the late 19th century, investigations into the aetiology of epilepsy had found that epileptic seizures could be recorded and evoked in the cerebral cortices of animals.[5] This marked a movement towards modern-day thinking about epilepsy, and by 1929 there were reports of using scalp-electrodes to measure brain electrical activity and classify epileptic seizures.[7]

Today, epilepsy is considered not as a single disease or syndrome, but rather as a broad category of symptom complexes that arise from disordered brain functions of multiple aetiologies.[4] A seizure is the manifestation of abnormal, synchronous mal-coordinated discharges of cortical neurones, and epilepsy is the predisposition for recurrence of such seizures.

Epilepsy can affect all members of society, with significant implications for sufferers' psychological, physical and social wellbeing. Epilepsy has been shown to carry a heightened risk of injury and sudden death,[8] in addition to widely-acknowledged psycho-social problems resulting from prejudice, stigma and discrimination.[9]

1.1.1. Incidence

In industrialised countries, the annual incidence rate of epilepsy is estimated at 40-70 per 100,000 people.[10-12] Epilepsy has been shown to have a predilection for lower-socioeconomic groups, although the true extent of this relationship remains unclear, as does the directionality of any association.[13-15] The age-distribution of epilepsy follows a bimodal distribution, with higher rates in the extremities of age. During recent years, a decreasing incidence in children has been observed, along with a simultaneous increase in the elderly population.[16, 17] In children, this trend is not fully

understood, although there is speculation that healthier lifestyles by expectant mothers, improved perinatal care, and enhanced immunisation programmes are responsible.[18] The increasing incidence in the elderly is a likely consequence of improved longevity, with independent risk factors for epilepsy, such as cerebrovascular disease, more prevalent with advancing age.[19]

1.1.2. Prevalence

Epilepsy is a common neurological disorder that affects in excess of 50-million people worldwide.[20] Estimates suggest that 5% of the population will experience non-febrile seizures at some point in life.[21] The global prevalence of active epilepsy (a seizure within the last five years) is approximately 0.5-1%.[21] The prevalence of epilepsy is highest in the elderly population, affecting up to 1.5% of those over 75-years of age.[22]

1.1.3. Diagnosis

The diagnosis of epilepsy in adults should be established by a specialist medical practitioner with training and expertise in epilepsy; similarly, epilepsy in children should be diagnosed by a paediatrician with specialist epilepsy training. Diagnosis is usually clinical following a detailed history from the individual and eye-witnesses (where possible), and requires at least two unprovoked seizures separated by a period greater than 24-hours.[23]

Neurological imaging investigations are increasingly used when making a diagnosis of epilepsy to help identify the aetiology. It is recommended that imaging is undertaken in all patients presenting with a first seizure aged 25 years and over, in addition to all patients presenting with apparent partial-onset seizures to identify the aetiology and to rule out treatable or progressive cerebral lesions.[24]

1.1.4. Seizure Classification

Seizures can be classified according to the origin of the abnormal electrical activity in the brain.

Partial (also termed focal or localised) seizures result from abnormal electrical activity originating from a focal brain region. When a partial seizure affects the consciousness of an individual, it is termed a ‘complex partial seizure’; when consciousness is unaffected, the seizure is termed a ‘simple partial seizure’. Both simple and complex partial seizures arise from only one hemisphere of the brain. If the electrical activity spreads to involve the opposing hemisphere, it is termed a ‘partial seizure with secondary generalisation’.

Generalised seizures are the result of widespread abnormal electrical activity affecting both hemispheres of the brain simultaneously. They always distort consciousness, and are classified into six main categories according to the resulting behavioural effects.

The International League Against Epilepsy (ILAE) criteria of 1981 provides a widely-accepted means of seizure classification based upon clinical manifestation[25] (*Table 1.1*).

Table 1.1: ILAE seizure classification criteria, 1981

Seizure Classification	Type of Seizure
Partial (focal) seizures	Simple partial seizures Complex partial seizures Partial seizures with secondary generalised seizures
Generalised seizures	Tonic-clonic (formerly 'grand mal') Absence seizures (formerly 'petit mal') Tonic Clonic Atonic Myoclonic
Unclassifiable seizures	

There have been revisions of the ILAE seizure classification criteria to include various epilepsies and syndromes that do not fit into the 1981 framework. Not only do the 1981 criteria classify according to clinical signs and symptoms, the majority of clinical trials into drug efficacy identified seizure-type according to the same criteria. As a consequence, clinicians have a useful classification tool that allows evidence-based medicine to be implemented, hence its widespread use throughout the UK.

In addition to the above, seizures can be further classified according to whether there is an established causal factor. If there is a known underlying cause, this is denoted by the term 'symptomatic', whereas 'idiopathic' refers to cases where a causal factor cannot be established. The term 'cryptogenic' is used when there is a presumed underlying cause that has failed to be identified during evaluation.

1.1.5. Seizure Pathophysiology

A seizure results when there is an imbalance between inhibitory and excitatory forces within a network of neurones in the brain such that there is overall excessive excitability.[26] The clinical manifestation of a seizure is dependent upon the affected brain regions, either directly at the seizure focus, or indirectly through localised spreading.

Epilepsy is the predisposition for recurrent seizures, yet there are many known aetiologies that lead to its development. Perhaps unsurprisingly, the pathophysiological mechanisms that result in partial-onset seizures differ considerably from those that underlie generalised-onset seizures.[26]

1.1.5.1. Partial-Onset Seizures

The pathophysiology underlying the hyper-excitable state that results in partial-onset seizures can be split into two broad categories; mechanisms leading to decreased inhibition, and mechanisms leading to increased excitation. The main mechanisms leading to decreased inhibition primarily involve γ-aminobutyric acid (GABA) receptors, whilst ion channels and glutamate are related to increased excitation.

GABA is the principal inhibitory neurotransmitter in the human brain.[27] It binds to two major classes of receptor; $GABA_A$ and $GABA_B$.[28] Upon synaptic excitation, $GABA_A$ receptors act to oppose the action potential through decreasing membrane resistance and promoting membrane hyperpolarisation back towards the resting potential.[26] Overall, this makes it more difficult for the membrane to reach the threshold potential required to generate an action potential, thereby acting as an inhibitory force. $GABA_B$ receptors act to suppress synaptic neurotransmitter release from the presynaptic terminal, thereby increasing the initial stimulus required for neuronal firing to occur.[26] Overall, $GABA_A$ and $GABA_B$ receptor stimulation results in inhibition of neuronal action potentials. A lack of this inhibition is believed to be one of the mechanisms underpinning development of partial-onset seizures.[29]

Voltage-gated sodium channels (VGSCs) cycle through three states; open, closed, and inactivated.[30] When open, sodium ions pass through the membrane, creating an electrical current that leads to depolarisation. The VGSC then enters an inactivated state for a period of time before returning to a functioning state; this period of inactivation is termed the 'refractory period', during which re-opening of a VGSC is not possible.[26] The refractory period serves to prevent excessive action potential propagation, with neuronal hyper-excitability avoided as a result. Genetic variants have been implicated in seizure pathogenesis by interfering with this mechanism, and extension of the VGSC refractory period is believed to be the mechanism of action of several antiepileptic drugs (AEDs).[30]

Glutamate is the major excitatory neurotransmitter in the human brain.[27] There are three types of ionotropic glutamate receptors named after agonists that bind to them with high specificity; NMDA-, kainate- and AMPA-type receptors. Fast neurotransmission is achieved through activation of AMPA- and kainate-type receptors, whilst slow excitation involves NMDA receptors. AMPA and kainate

receptors allow passage of monovalent cations such as sodium and potassium, whereas NMDA receptors additionally allow passage of divalent cations such as calcium.[26] Although complex and not fully understood, it is possible that subtle alterations in receptor properties could ultimately result in elevated sensitivity to glutamate, with neuronal hyper-excitability and seizure genesis a likely consequence.

1.1.5.2. Generalised-Onset Seizures

The pathogenesis of generalised-onset seizures remains poorly understood, perhaps due to the variety of seizure types encompassed by this umbrella term. The best understood example of generalised-onset seizure pathogenesis is observed in typical absence seizures.[26] The thalamocortical circuit has normal oscillatory rhythms that govern the sleep-wake cycle; alterations in this rhythm have been implicated in generalised-onset seizures, specifically involving T-type calcium channels and the T-calcium current they generate in the nucleus reticularis of the thalamus.[26] Although not fully understood, this serves to illustrate that generalised-onset seizures may be the result of deep-lying and often genetically determined neuronal idiosyncrasies, whereas partial-onset seizures are more likely the result of cortical-based neuronal pathology.

1.1.6. Pharmacological Management

AEDs form the mainstay of treatment for people with epilepsy. The goal of AED therapy is to achieve seizure freedom whilst minimising adverse effects associated with treatment.[31] An estimated 70% of sufferers respond well to AEDs, while the remaining 30% respond poorly and continue to experience seizures.[32, 33]

In the past twenty years there has been a dramatic increase in the number of licensed AEDs, with currently in excess of 20 agents in the clinicians' arsenal. Although all AEDs aim to offer seizure freedom, the mechanisms through which individual agents achieve this are variable, which in turn reflects the complexities of seizure pathogenesis.

1.1.6.1. Choice of AED

The National Institute for Health and Clinical Excellence (NICE) is an independent special health authority of the National Health Service (NHS) responsible for providing national guidance on promoting good health and preventing and treating ill health.[34, 35] The most recent NICE guidelines pertaining to the pharmacological management of epilepsy were published in October 2004 (*Table 1.2*).

Initiation of AED therapy is recommended once a diagnosis of epilepsy has been made by a specialist clinician.[36] As shown in *Table 1.2* there are often several recommended first-line agents for management of a given seizure type. The final choice of AED is made by the prescribing clinician, whose decision is guided by factors such as patient age, comorbidities and concurrent medications, as well as their knowledge of agent-specific adverse-effects and mechanism of action.

It is estimated that 50-65% of patients respond well to first-line AED monotherapy.[33] If the first-line monotherapy proves ineffective, it is recommended that an alternative first-line monotherapy agent be trialled before considering add-on therapy. Approximately 13% of individuals respond to this alternative first-line AED monotherapy,[33] with the remainder requiring polytherapy with one or more AEDs in an attempt to achieve adequate control. Polytherapy patients suffer higher rates of adverse effects associated with treatment, and evidence suggests this subgroup has a poorer prognosis in comparison to monotherapy responders.[37, 38]

Table 1.2: NICE recommendations – drug options by seizure type.[36]

Seizure Type	First-line agents	Second-line agents	Other agents	Contra-indicated agents
Generalised tonic-clonic	Carbamazepine [a] Lamotrigine Sodium Valproate Topiramate [a]	Clobazam Levetiracetam Oxcarbazepine [a]	Acetazolamide Clonazepam Phenobarbital [a] Phenytoin [a] Primidone [a]	Tiagabine Vigabatrin
Absence	Ethosuximide Lamotrigine Sodium Valproate	Clobazam Clonazepam Topiramate [a]		Carbamazepine [a] Gabapentin Oxcarbazepine [a] Tiagabine Vigabatrin
Myoclonic	Sodium Valproate	Clobazam Clonazepam Lamotrigine Levetiracetam Piracetam Topiramate [a]		Carbamazepine [a] Gabapentin Oxcarbazepine [a] Tiagabine Vigabatrin
Tonic	Lamotrigine Sodium Valproate	Clobazam Clonazepam Levetiracetam Topiramate [a]	Acetazolamide Phenobarbital [a] Phenytoin [a] Primidone [a]	Carbamazepine [a] Oxcarbazepine [a]
Atonic	Lamotrigine Sodium Valproate	Clobazam Clonazepam Levetiracetam Topiramate [a]	Acetazolamide Phenobarbital [a] Primidone [a]	Carbamazepine [a] Oxcarbazepine [a] Phenytoin [a]
Focal with/without secondary generalisation	Carbamazepine [a] Lamotrigine Oxcarbazepine [a] Sodium Valproate Topiramate [a]	Clobazam Gabapentin Levetiracetam Phenytoin [a] Tiagabine	Acetazolamide Clonazepam Phenobarbital [a] Primidone [a]	
a – hepatic enzyme-inducing agents				

1.2. Lamotrigine

Lamotrigine (LTG) was first developed as an AED in the 1970s as a result of its weak inhibitory activity on dihydrofolate reductase, following the hypothesis that folate was epileptogenic.[39] This mechanism of action has since been dismissed following a lack of supporting evidence,[40] however LTG remains one of the most commonly employed AEDs in the developed world.

LTG was first licensed in the UK for the treatment of refractory partial epilepsy in 1991. At present, it is employed as first-line treatment for partial-onset and generalised seizures, may be used for atonic, tonic, and atypical absence seizure types, and also has licensing for the treatment of Lennox-Gastaut syndrome (*Table 1.2*). In addition to its antiepileptic activity, LTG is indicated for the maintenance treatment of bipolar disorder; however its effectiveness in the treatment of acute mood disorders remains to be established.[41] LTG also appears to have some efficacy in the treatment of migraine aura[42] and neuropathic pain.[43]

1.2.1. Chemistry

LTG [3,5-diamino-6-(2,3-dichlorophenyl)-1,2,4-triazine] is a phenyltriazine derivative that is chemically unrelated to other AEDs.[44, 45] It is a white powder with a molecular weight of 259.1 and is poorly soluble in water or ethanol.[46]

In the UK, LTG is available in both tablet and chewable / dispersible forms. Tablets are available in 25, 50, 100 and 200mg strengths, while chewable / dispersible forms come in 2, 5, 25 and 100mg preparations.[47] The active compound is combined with several inactive ingredients to form each preparation. The utilised ingredients have no known effect on LTG pharmacokinetics, and include lactose, magnesium stearate, microcrystalline cellulose, povidone, sodium starch glycolate, denatured alcohol and purified water.[44]

1.2.2. Pharmacodynamics

LTG is thought to exert its antiepileptic effects primarily through its action on VGSCs.[48] Studies suggest this occurs as a result of binding to sodium channels during their refractory period; the result is a stabilisation of the pre-synaptic membrane and a decrease in excitatory neurotransmitter release.[49-51]

In addition to the well-established effects on VGSCs, LTG has also been shown to interfere with neuronal calcium currents through blockade of voltage-gated calcium channels,[52-54] and potassium currents.[55]

In terms of blocking calcium channels, LTG has been shown to block N-, P/Q-, T-, and R-type channels. LTG blockade of N- and P/Q-type calcium channels in cortical neuronal cells has been shown to inhibit glutamate release from presynaptic terminals.[52, 53] Through inhibiting glutamate release, a greater stimulus is required in order to achieve excitation in the postsynaptic terminal.[56] T- and R-type voltage-gated calcium channels are predominantly expressed in the thalamocortical circuitry in the brain, and both have been shown to be targets for LTG.[57] This effect, and the anatomical location of these channels, could explain the efficacy of LTG in the treatment of generalised absence seizures.[58]

LTG has also been shown to modulate intracellular neuronal calcium levels in neocortical slices from rats. Transient elevations in calcium have been correlated with epileptiform discharges; because of the antagonistic effect of LTG on intracellular calcium levels, this has been hypothesised as at least a contributing source of its anticonvulsant efficacy.[54]

Finally, LTG has been shown to modulate potassium-mediated hyperpolarisation currents in slices of rat cerebral cortex. This mechanism has also been postulated as a contributing factor to its efficacy as an anticonvulsant, although the extent to which these currents contribute to the generation of seizures in humans is a topic of discussion.[55]

Overall, the pharmacodynamic profile of LTG is complex and not yet fully understood, although it is generally accepted that LTG's antiepileptic efficacy is primarily mediated through blockade of VGSCs. While there is growing evidence that subsidiary mechanisms may contribute to the efficacy of LTG, the extent of this contribution remains to be elucidated.

1.2.3. Pharmacokinetics

Pharmacokinetics is a branch of pharmacology that studies the course of externally administered compounds as they travel through the body. It focuses on four main areas: absorption, distribution, metabolism, and excretion.

Absorption: LTG is available in both tablet and chewable / dispersible forms, and it has been shown that no significant differences exist between them in terms of rate and extent of absorption.[3] LTG does not undergo first-pass metabolism, and has an absolute oral bioavailability in excess of 98%[46] which is not affected by food.[59] Following oral administration, it is rapidly and completely absorbed, achieving a peak plasma concentration within in 1-3 hours.[45] The peak plasma concentration has been shown to follow a linear relationship with dose in several small-scale trials in both healthy volunteers and epilepsy patients.[60-62]

Distribution: LTG is uniformly distributed throughout the body following oral administration.[63] The volume of distribution is independent of dose, and is similar between single- and multiple-dose administration.[44, 59] LTG has been shown to be approximately 55% bound to human plasma proteins, primarily albumin, at plasma LTG concentrations of 1-10μg/ml *in vitro.*[44] Plasma protein binding of LTG has been shown *in vitro* to be unaffected by therapeutic concentrations of other AEDs such as phenytoin (PHT), phenobarbital (PB) or valproic acid (VPA).[45] Furthermore, LTG did not displace the AEDs carbamazepine (CBZ), PHT or PB from their protein-binding sites.[44] LTG has few, if any, clinically significant drug-interactions at the level of distribution because of this relatively-low plasma protein binding, with competition for protein-binding sites unlikely as a consequence.[3]

Metabolism: The 1A4 isoform of the uridine diphosphate glycosyltransferase (UGT) family of enzymes is believed to be responsible for the majority of LTG glucuronidation. Remmel and Sinz were the first to identify the N2-glucuronide product as a metabolite of LTG in 1991 following animal studies.[64, 65] Since then, it has been estimated that 75% of LTG is metabolised to the N2-glucuronide form, and an additional 10% to the N5-glucuronide form, neither of which are metabolically active.[66] It is widely assumed that UGT1A4 is the enzyme responsible for N2-glucuronidation given its broad role in amine glucuronidation, although data to support this notion is scarce. There has been recent speculation that other UGT isoforms may have the capacity to produce LTG N2-glucuronide products.[67] Rowland and colleagues propose that, in addition to UGT1A4, at least one other UGT is involved in LTG N2-glucuronidation, and roles for several additional isoforms cannot be excluded.[67]

Excretion: Both unconjugated LTG and its metabolites are excreted from the body primarily via the urinary system.[63] In a study on six healthy volunteers, 240mg of radio-labelled LTG was orally-administered, with 94% of the radioactivity recovered in the urine, and 2% in the faeces.[44] Of the radioactivity recovered in the urine, 76% was in the N2-glucuronide form, 10% was in the N5-glucuronide form, and 0.14% was a N2-methyl metabolite.[44]

Elimination half-life is a measure of the time taken for the blood concentration of a substance to reduce to half of its steady state. LTG shows considerable inter-individual variation in terms of elimination half-life, but little intra-individual variability.[63] In healthy volunteers, the half-life of LTG has been shown to range from 25.4-32.8 hours in monotherapy, 48.3-70.3 hours when comedicated with VPA, and 12.6-14.4 hours when comedicated with enzyme-inducing AEDs.[58]

1.2.3.1. Interactions

LTG appears to have little effect on the pharmacokinetics of other AEDs with the exception of clonazepam (CNZ) where co-administration has been shown to reduce serum CNZ concentrations by 20-38%.[68] The pharmacokinetics of LTG however appears to be influenced by many other antiepileptic agents. In 2005, Weintraub and colleagues investigated the effects of AED comedication on LTG clearance by measuring serum LTG levels. Sixteen different regimens of LTG comedication were identified; fourteen dual-therapies and two triple-therapies. A total of 570 patients were included in the analysis aged 12 years and older, the results of which are illustrated graphically in *Figure 1*.[69]

Weintraub concluded that both PHT and CBZ increase LTG clearance by approximately 125% and 30-50%, respectively. In addition to this, VPA was found to decrease LTG clearance by approximately 60%.[69] None of the other comedications investigated were found to result in a statistically significant change in LTG clearance.

Alternative studies have identified PB, oxcarbazepine (OXC) and primidone (PRM) as inducers of LTG metabolism, resulting in increased LTG clearance by approximately 34-52%.[70-72] The findings relating to PB, OXC and PRM were not initially identified by Weintraub *et al*, most probably due to small numbers of patients.

There has also been recent investigation of pharmacokinetic interactions between LTG and oral contraceptives. LTG has been found to cause a small (19%) decrease in the progesterone component of the oral contraceptive pill, with no effect on oestrogen (ethinyl estradiol [EE]) levels.[73] Conversely, progesterone has been shown to have no affect on LTG pharmacokinetics, yet EE leads to a reduction in serum LTG levels in excess of 50%.[74, 75] This means that, although LTG is unlikely to affect women's risk of pregnancy (apart from a slight increase in risk for women taking progesterone-only oral contraception), the increased levels of EE associated with oral contraceptives is likely to interfere with plasma LTG concentrations to the extent that efficacy may be compromised.[58]

Auto-induction is the process where a molecule induces the enzymes that are responsible for its own metabolism. LTG has been shown to induce its own metabolism, resulting in a steady state half-life decrease of 25%, and a clearance increase of 37%.[44] This degree of auto-induction is not considered to impact upon the use of LTG in clinical practice as it is overcome by the dose-titration process.[45]

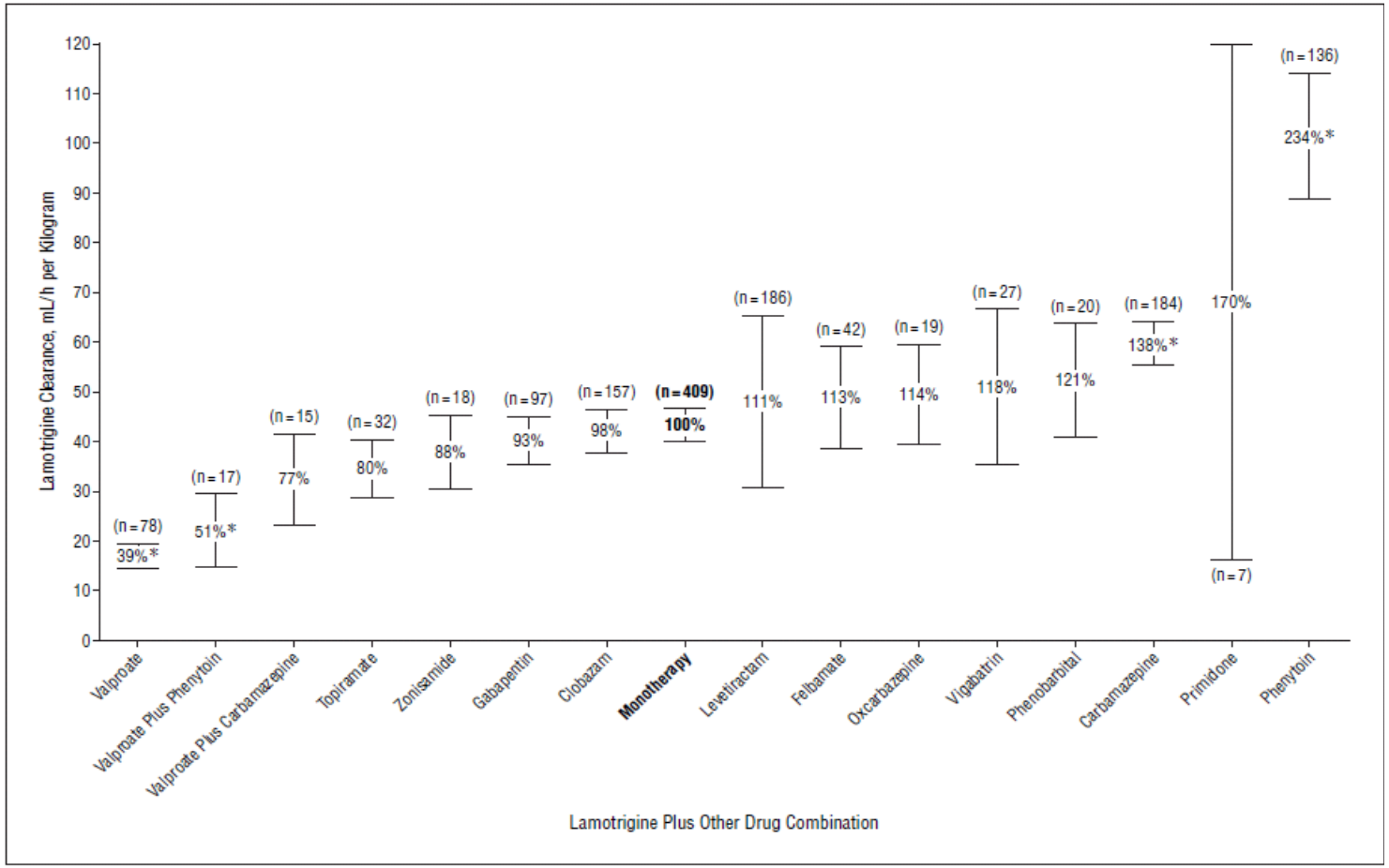


Figure 1 - Effect of drug combinations on lamotrigine clearance.[69]
Asterisk indicates drug combinations with clearance significantly different from monotherapy (P<0.05)
n = number of distinct lamotrigine clearances, including the drug listed and other drugs without significant effects on lamotrigine clearance
Bars = 95% confidence intervals
Percentage = ratio of mean clearance in given combination as a percentage of mean clearance with monotherapy
Monotherapy data are bolded for clarity

1.2.4. Clinical Efficacy and Tolerability

The clinical efficacy and tolerability of LTG in the treatment of partial-seizures and generalised-onset seizures is reviewed below. Although LTG is often used as an adjunctive AED, in the interests of specificity and brevity, LTG monotherapy alone will be considered.

1.2.4.1. Partial Seizures

Studies investigating the efficacy of LTG monotherapy in the treatment of partial seizures are summarised in Appendix I. Various dosing regimens of LTG were administered to people diagnosed with partial-onset seizures, and its efficacy determined through comparison to other AEDs including PHT, CBZ and gabapentin (GBP).

Brodie *et al* compared the efficacy and safety of LTG against CBZ in a cohort of 146 patients with partial-onset seizures and 122 with primary generalised tonic-clonic seizures.[76] Dosing regimens were as follows; LTG 50mg/day, titrated up to 150mg/day over a 4 week period; CBZ 200mg/day, titrated up to 600mg/day over a 4 week period. Dose remained fixed until 24-weeks, at which point it could be adjusted, if clinically indicated, for the remaining 24-weeks. The study showed that for patients with partial-onset seizures, 35% were fully controlled on LTG in comparison to 37% on CBZ, indicating no significant difference in efficacy between the two agents. However, LTG was shown to be better tolerated than CBZ overall; more patients withdrew due to adverse effects on CBZ (21%) than LTG (11.5%), and a comparison of the areas under the Kaplan-Meier curve (for time to withdrawal) supported a significant difference in favour of LTG (hazard ratio 1.57 [95% confidence interval 1.07-2.31]).[76] In addition to time to withdrawal, 15 specific adverse effects were screened for and analysed, with a statistically significant difference shown between groups for levels of 'sleepiness'. The LTG group was shown to have significantly lower levels of withdrawal due to sleepiness than CBZ (12% vs 22% respectively [95% CI -28 to -1]), although this should be interpreted with caution as data was obtained from patient reports, was not standardised, and as such may be susceptible to recall bias.

In a different trial, Reunanen and colleagues also found no statistically significant difference between LTG and CBZ in terms of efficacy, although the authors felt that LTG 200mg/day (LTG-200) was the most effective regimen when compared with LTG 100mg/day (LTG-100) and CBZ 600mg/day (CBZ-600). For LTG-200, 60.4% of participants achieved seizure freedom during the 24-week study period in comparison to 51.3% of LTG-100 and 54.7% of CBZ-600. Although a significant difference in efficacy was not demonstrated, both dosage regimens of LTG were better tolerated than CBZ-600; 53% of patients reported adverse experiences attributable to treatment in the CBZ-600 group in

comparison to 23% (LTG-100) and 28% (LTG-200). Overall, LTG appeared equally effective but better tolerated when compared with CBZ.[77]

Arm A of the Standard and New Antiepileptic Drugs (SANAD) trial concluded that, for time to 12-month remission, CBZ was significantly better than GBP (hazard ratio 0.75 [95% confidence interval 0.63-0.90]) and showed a non-significant advantage for CBZ against LTG (0.91 [0.77-1.09]), topiramate (0.86 [0.72-1.03]) and OXC (0.92 [0.73-1.18]).[2] There was a significant difference between the CBZ and LTG groups for treatment failure owing to unacceptable adverse events throughout the study duration. At six-years, 78% of LTG patients were without failure in comparison to 68% of CBZ patients; a significant difference of 10% [95% CI 3 to 17].[2]

Another study conducted by Brodie and co-workers was designed to assess tolerability differences between LTG and CBZ in 150 elderly patients (mean age 77 years) with newly diagnosed epilepsy.[78] Median daily doses of LTG and CBZ in patients completing the study were 100mg (range: 75-300mg) and 400mg (range: 200-800mg), respectively. The main difference between the groups was the rate of drop-out due to adverse events (LTG 18% versus CBZ 42%). In addition, a statistically significant difference between LTG and CBZ was identified ($p<0.001$), with more patients completing the study on LTG (71%) in comparison to CBZ (42%), which the authors attributed to the superior tolerability of LTG.

A more recent study undertaken by Steiner and colleagues attempted to quantify adverse effects by using the 'Side Effects and Life Satisfaction' (SEALS) inventory;[79] a tool for measuring the effects AEDs have on health-related aspects of quality of life.[80] Comparison of pre-treatment SEALS scores with scores obtained at pre-determined points throughout the 48-week trial was performed, and found that scores decreased (meaning improvement) in the LTG group but increased in a matched PHT group.[79] This is likely to reflect, at least in part, the fewer drug-related adverse effects experienced with LTG monotherapy as compared to PHT.

Although these studies show the efficacy of LTG in the treatment of partial-onset seizures, there are several limitations which mean that results should be interpreted with caution. The randomised, double-blind study by Brodie and colleagues[76] was designed to include patients with 'partial-onset seizures with or without generalisation', as well as patients with 'primary generalised tonic-clonic seizures'. Although the primary outcome analysis was separated according to seizure type, it failed to fully differentiate between them as evident by a number of participants included in both these categories due to mixed seizure types. A greater proportion of these 'mixed' patients were included in the CBZ arm of the study (n=6) in comparison to LTG (n=2) and the responsiveness of these

patients to their respective AEDs has the potential to skew the reported efficacy. This could have been avoided by excluding patients with mixed seizure types, or at least ensuring an even distribution between the LTG and CBZ groups. Nonetheless, these were a minority of the total number of participants (n=260), meaning the majority of participants in the LTG (129/131) and CBZ (123/129) groups were successfully classified to only one of the seizure type sub-classifications. Therefore, the results can reasonably be assumed to be a true reflection of efficacy, and they are arguably more indicative of effectiveness in a clinical setting.

Throughout the studies, measures have been taken to ensure treatment and control groups are as similar as possible to minimise the effects of confounding factors. Although Brodie and colleagues[76] attempted this through stratification according to seizure type, they failed to remove other elements of bias from the study. The investigators acknowledged that disproportionately more patients were assigned to the LTG group with more than 100 seizures before recruitment; this was later identified to be of significant prognostic value (p=0.04) with 'the more seizures reported before randomisation, the greater the likelihood of a seizure after 6-weeks of treatment or premature withdrawal'.[76] As a consequence, the true efficacy of LTG might in fact be higher than reported, although this effect may again be negligible owing to the large sample size (n=260).

Overall, clinical studies suggest that LTG is efficacious in the treatment of partial onset seizures. It has been repeatedly shown that no significant difference exists between LTG and the historical first-line drug CBZ,[76-78] and LTG was most recently shown to be non-inferior to CBZ by Marson and colleagues.[2] In addition to these reports, Steiner and colleagues found PHT to be comparable to LTG in terms of seizure freedom at 48-weeks (48% and 41% respectively),[79] and Brodie *et al* concluded that GBP and LTG were similarly effective at the end of a 24-week maintenance dose period.[81]

1.2.4.2. Generalised-onset seizures

There have been many investigations of the efficacy of LTG in the treatment of generalised-onset seizures, the results of which are arguably inconclusive.

Several studies have suggested the efficacy of LTG in the treatment of childhood absence epilepsy. Frank and colleagues evaluated the efficacy of LTG in a cohort of 45 children (3-15 years) with newly diagnosed absence seizures in a placebo-controlled double-blind trial.[82] Following an initial dose-escalation period where all participants received LTG monotherapy, participants were randomised to receive either LTG or placebo, with treatment success defined as complete freedom from seizures confirmed by 24-hour and hyperventilation electroencephalogram (EEG). A total of 62% of patients receiving LTG achieved successful treatment, significantly more than in the placebo group (21%).

LTG and VPA monotherapies were compared in the treatment of newly-diagnosed absence seizures by Coppola and colleagues in a randomised, un-blinded parallel-group study design.[83] Thirty-eight children (17 boys, 21 girls) aged from 3 to 13 years (mean 7.5 years) were enrolled in the study, and efficacy was determined by seizure freedom at predetermined assessment points of 1, 3 and 12 months. Efficacy was comparable at 12-months, with 13 children taking VPA and 10 taking LTG remaining seizure free. VPA achieved a faster onset of action at the 1 and 3 month assessments, an effect attributed to a shorter titration schedule in the VPA group. In this study, side effects were found to be more common in the LTG group (31.8%, n=6) than VPA (10.6%, n=2), although these were mostly mild and transient.

Most recently, Glauser and colleagues undertook a double-blind, randomised, controlled trial to compare the efficacy, tolerability and neuropsychological effects of ethosuximide (ESM), VPA and LTG in children with newly diagnosed childhood absence epilepsy.[84] The trial recruited a total of 453 children who were randomised to receive either ESM (n=156), LTG (n=149) or VPA (n=148). After 16-weeks of treatment, ESM and VPA were found to be similar in terms of freedom-from-failure rates (53% and 58% respectively, odds ratio VPA vs ESM 1.26 [95% CI 0.80-1.98]), however both achieved significantly higher rates than LTG at 29% (VPA vs LTG; odds-ratio 3.34 [95% CI 2.06-5.42], $p<0.001$), (ESM vs LTG; odds ratio 2.66 [95% CI 1.65 to 4.28], $p<0.001$). In terms of tolerability, there were no significant differences in drop-out rates due to adverse effects between groups, however the incidence of attentional dysfunction in children receiving VPA was found to be significantly higher than those receiving ESM (49% and 33%, respectively) (odds ratio 1.95 [95% CI 1.12 to 3.41], $p=0.03$).[84] Although these results undeniably show LTG to be inferior to ESM and VPA, a longer-term follow-up period will be required to determine whether ESM remains non-inferior to VPA in terms of efficacy.

Prasad and colleagues evaluated whether LTG monotherapy could prove an effective treatment of juvenile myoclonic epilepsy (JME) in a retrospective study.[85] Seizure outcome was found not to differ between patients receiving VPA monotherapy (n=36) and LTG monotherapy (n=14). The withdrawal rate per patient-year of treatment with VPA was lower than that with LTG ($p=0.12$), which, although not statistically significant, suggests VPA may be superior in terms of adverse effects. The results of this trial should be interpreted with caution due to its small patient cohort; VPA may have been shown to be significantly better than LTG had the study been better powered.

In addition, there have been reports of worsening of myoclonus with the administration of LTG. In a study by Guerrini and co-workers, a total of 21 patients diagnosed with severe myoclonic epilepsy of

infancy were treated with LTG, with 80% of participants experiencing a deterioration in seizure control.[86]

In the aforementioned study by Brodie and colleagues, 122 newly-diagnosed patients with primary generalised tonic-clonic seizures were assigned to either LTG or CBZ over a 48-week study period. The study failed to identify any difference in efficacy between LTG and CBZ for the treatment of primary generalised tonic clonic seizures, with the percentage of patients reporting seizure-freedom during the last 24-weeks of the trial almost identical LTG (39%) and CBZ (38%). As mentioned previously, fewer patients on LTG withdrew from the study due to adverse effects than on CBZ (11.5 vs 21%, $p<0.05$), and patients were more likely to complete the study if randomised to LTG rather than CBZ (65 vs 51%, hazard ratio 1.57 [95% CI 1.07-2.31], p=0.018]).

Most recently, Arm B of the SANAD trial recruited 716 patients with newly diagnosed generalised or unclassifiable epilepsy in a randomised, unblinded, pragmatic study design. Patients were randomly assigned on a 1:1:1 basis to receive LTG (n=239), TPM (n=239) or VPA (n=238), and efficacy was measured using time to 12-month remission. VPA was found to be significantly better than LTG overall (hazard ratio 0.76 [95% CI 0.62-0.94]), as well for the subgroup with an idiopathic generalised epilepsy (0.68 [0.53-0.89]). In terms of failure due to adverse effects at 6-years, LTG showed a favourable 7% difference in comparison to VPA (95% CI -1 to 14), whilst TPM showed an additional 14% decrease in numbers in comparison to VPA (95% CI -30 to 1).[1] Although the 95% confidence intervals failed to show statistically significant differences, the data suggests LTG is a favourable AED in terms of adverse events in comparison to other AEDs.

Overall, there is conflicting evidence surrounding the use of LTG in the treatment of generalised-onset seizures. LTG appears to have some efficacy in the treatment of generalised-onset seizures in general; however it appears inferior to other AEDs when certain seizure types are assessed in isolation. In addition, it is not clear whether LTG is as well tolerated by patients with generalised-onset seizures as other AEDs. Overall, the evidence suggests VPA should remain first-line treatment for patients with generalised-onset seizures, although there may be some instances where LTG should be considered a viable alternative.

1.2.4.3. Cost Effectiveness

The recent SANAD trial concluded that LTG is a cost-effective alternative to CBZ for patients diagnosed with partial-onset seizures.[2] The quality adjusted life year (QALY) is a broad measure of benefit that measures health-related quality of life. It is affected not only by clinical outcome measure, but also other factors such as the consequences of drug side-effects on patients' health.

The SANAD trial issued quality-of-life questionnaires (EQ-5D) at baseline and 2-year follow up, and used that information to calculate the cost-effectiveness of each of the AEDs in comparison to CBZ in terms of cost-per-QALY (*Table 1.3*). Two separate analyses were undertaken; one excluding OXC (n=636) and one including OXC (n=414). Each analysis was restricted to adult patients since the EQ-5D questionnaire was not validated for use in children, and only included patients who provided complete EQ-5D responses.

Table 1.3: Incremental cost-effectiveness ratios versus carbamazepine- cost per QALY (SANAD)

	Cost (£)	**QALYs**	**Incremental Cost (£)**	**Incremental QALYs**	**Incremental cost-effectiveness ratio (£/QALY)**
Comparison Excluding OXC					
CBZ	1226	1.477	-	-	-
TPM	2009	1.501	783	0.024	Extended Dominance
LTG	2257	1.564	248	0.063	11851
GBP	2561	1.491	304	-0.073	Dominated
Comparison Including OXC					
CBZ	1095	1.491	-	-	-
OXC	1839	1.611	744	0.12	6200
TPM	1930	1.541	91	-0.07	Dominated
LTG	2078	1.563	148	0.022	Extended Dominance
GBP	2573	1.480	495	-0.083	Dominated
QALYs = quality-adjusted life years, OXC = oxcarbazepine, CBZ = carbamazepine, TPM = topiramate, LTG = lamotrigine, GBP = gabapentin.					

For the analysis that excluded OXC, GBP has a positive incremental cost and a negative incremental QALY gain and is therefore dominated by LTG. Because LTG has a lower incremental cost-effective ratio than TPM, TPM is ruled out on the grounds of extended dominance. The incremental cost-effectiveness ratio for LTG relative to CBZ is £11,851.[2]

The analysis that included OXC showed TPM and GBP have positive incremental costs and negative incremental QALY gains and are therefore dominated by OXC and LTG respectively.[2]

Overall, the results of the SANAD analyses show LTG is a cost-effective alternative to CBZ in the treatment of partial-onset seizures. Further analyses showed the probabilities that LTG is cost-effective at ceiling ratios of £10,000, £30,000, and £50,000 per QALY were 0.42, 0.82, and 0.89 respectively when OXC was excluded from the analysis, and 0.36, 0.66, and 0.73 respectively when OXC was included.[2]

1.2.4.4. Interpretation

Overall, no significant difference has been shown between LTG and other AEDs such as PHT, TPM and the current first-line drug CBZ for the treatment of partial-onset seizures. In addition to this, it has been shown to have some efficacy in terms of seizure control in generalised-onset seizures, although it is not as efficacious as the current first-line agent VPA. Regardless of efficacy in the treatment of specific seizure-types, LTG has been shown to be better tolerated than many other first-line AEDs including VPA, CBZ, PHT and TPM. In situations where efficacy in terms of seizure control has been shown to be at least non-inferior, LTG should be considered first-line therapy owing to its evidently more desirable adverse events profile.

However, treatment with LTG is not without challenges. As with all AEDs, the effective dose of LTG varies widely from patient to patient, and finding the optimum dose can be difficult.[3] When used as monotherapy, some individuals require only 50mg/day to achieve seizure freedom, whereas others need as much as 800mg/day. Current practice involves titration of LTG over a six to eight week period until a pre-determined target dose is achieved; dose titration is required to minimise the risk of recognised side effects such as rash (including the life-threatening Stevens-Johnson syndrome) which is more likely to occur with quicker dose escalation. Once the pre-determined target dose is reached, dose adjustments are made thereafter according to the extent of efficacy and occurrence of adverse effects. Although this method has withstood the test of time, it remains sub-optimal for a large number of patients. If we consider those patients with a low dose requirement, the standard titration process may result in intolerable adverse effects necessitating drug withdrawal due to the excessively high target dose. Alternatively, patients with an unusually high dose requirement may be required to undergo successive increases in LTG dose over many months before the optimum dose is found, during which seizure activity persists. This unpredictability in dose requirement, coupled with the lengthy titration period, makes LTG a less attractive choice than some rival AEDs.

1.3. Pharmacogenetics

The term 'pharmacogenetics' was first used by Vogel in 1958 to describe the impact genetic variation can have on an individual's response to drug therapy.[87] Pharmacogenetic studies aim to determine whether observed differences in drug response within a population can be explained by underlying genetic variation between individuals.

One method of assessing the impact of genetic variation is through a candidate gene approach. Using knowledge of the pharmaokinetic and pharmacodynamic profiles of a given drug, proteins are identified that have the potential to affect the drug as it passes through the body. Once candidate genes encoding these proteins have been identified, the variation across those genes is determined through genotyping members of the study population. Subsequent statistical analyses are performed to identify significant associations between genetic variation and a predetermined clinical outcome measured in the study population.

Early pharmacogenetic trials focussed on rare outcomes such as adverse drug reactions (ADRs) associated with specific drug compounds. An early example of this was the occurrence of acute haemolysis in a minority of recipients of the anti-malarial drug, primaquine.[88] Initial studies concluded that acute haemolysis resulted from the deficiency of an enzyme (glucose-6-phosphate dehydrogenase [G6PD]) found in red blood cells,[89] and this was later attributed to genetic variation in the *G6PD* gene.[90]

The scope of pharmacogenetic research has expanded during the past fifty years to incorporate various aspects of drug response. In addition to ADRs, pharmacogenetic studies now seek to predict drug effectiveness, tolerability, and/or efficacy. Pharmacogenetic studies can also be used to investigate the genetic influences on serum concentrations and effective and maximally-tolerated doses. Patient characteristics such as body mass index (BMI), gender and age are recognised to influence dosing requirements of most drugs; combining patient characteristics and pharmacogenetic variation helps to build up a picture of the true influences on drug-dose requirements, ultimately striving towards stratified prescribing.

Genetic variants that are known to influence drug response have the potential to form an important clinical tool. If it is possible to predict drug response according to an individual's genetic make-up, then screening for genetic variants prior to drug initialisation may allow an individually-tailored therapeutic approach to be employed.

Warfarin is the most commonly prescribed vitamin-K antagonist worldwide because of its anticoagulant properties.[91] Serum concentrations of warfarin can be difficult to maintain within therapeutic limits, requiring serial monitoring of coagulation time via INR. Polymorphisms in two genes, *VKORC1* and *CYP2C9*, have been associated with inter-individual variability in warfarin metabolism, with over 30% of the variance in warfarin dose explained by polymorphisms in these genes.[92] Recent studies suggest that pharmacogenetics-guided dosing algorithms can accurately predict warfarin dosage,[93-96] and as a result genetic testing is now recommended in the United States to predict warfarin response.

1.3.1. Single Nucleotide Polymorphisms

The human genome is comprised of twenty-three pairs of chromosomes (twenty-two autosomal chromosome pairs, and one sex-determining pair), with each chromosome consisting of a coiled strand of DNA. Four different nucleotides join together to form a strand of DNA; adenine (A), thymine (T), cytosine (C), and guanine (G). The structure of DNA is a double-stranded helix; A and T form complementary base pairs (bp) on opposing strands, as do C and G. The human DNA sequence contains an estimated 3,100,000,000 bp,[97] and consists of approximately 20,000 – 25,000 genes.[98]

A genetic polymorphism is defined as the occurrence, together in the same population, of more than one allele or genetic marker at the same locus with the least frequent allele or marker occurring more frequently than can be accounted for by mutation alone.[99] Single nucleotide polymorphisms (SNPs) are DNA sequence variations that occur when a single nucleotide in the genome is altered. By definition, the frequency of this sequence variant must be ≥ 1% in a given population to be termed a SNP; a lower frequency is considered a mutation.

There are in excess of 2.3 million common SNPs (ie, minor allele frequency ≥ 0.05) across the human genome in persons of Caucasian European descent,[100] the majority of which do not have functionally significant effects. SNPs are classified according to their functional significance on protein transcription. There are five main categories of SNP, listed in descending order of perceived functional significance:

- Coding non-synonymous
    - Nonsense, missense, regulatory
- Splice site
    - 5’ splice site, 3’ splice site
- Messenger-RNA untranslated-region (mRNA UTR)
    - 5’ UTR, 3’ UTR

- Coding synonymous
- Intronic

Nonsense SNPs result in a premature stop-codon in the transcribed mRNA sequence, producing a truncated, incomplete and usually non-functional protein product. A missense SNP occurs when a codon is produced that encodes a different amino acid. Depending on the substituted amino acid, the resultant protein can range from non-functional to normal functioning; the effect of such amino acid substitutions depends upon the position within the translated protein. Regulatory SNPs lie in regions of DNA that directly or indirectly affect protein expression; alterations in regulatory regions may lead to a change in protein amount, structure and/or function.

Splice site SNPs can occur upstream (5') or downstream (3') to regions of transcribed DNA. Splicing occurs after RNA transcription, and is the process whereby introns are removed and exonic sequences are joined together. SNPs in splice sites may result in splicing errors, including loss of function at that site, reduced specificity, or displacement, with effects on translated amino acid sequences a common consequence.

mRNA UTR SNPs can also occur either 5' or 3' to regions of transcribed DNA. DNA regions encoding mRNA UTRs are transcribed into mRNA, and the 5' UTR contains several regulatory sequences including binding sites that may regulate gene expression. 3' UTR also contain binding sites for proteins and elements that lead to stabilisation or location of the mRNA sequence within a cell. For synonymous SNPs, the nucleotide substitution results in a codon-triplet that encodes the same amino acid as before. These were previously assumed to be 'silent' with no effect on protein function, however, because certain codons are translated more efficiently than others, this may not be entirely the case. Due to the complex, intricate nature of protein transcription/translation, subtle changes such as a relative delay in amino acid incorporation at the ribosome may impact upon the final structure of the protein. Intronic SNPs are in the non-coding regions of DNA (i.e. not in exons). They may have a role in gene expression but are not directly implicated in protein structure.

There are sixty-four possible triplet-combinations of nucleotides; with three exceptions, each codon encodes for one of the twenty amino acids used in protein synthesis. A table of the possible DNA codon combinations and the resultant amino acids is shown in Appendix II.

1.3.2. Tagging SNPs and Linkage Disequilibrium

Linkage disequilibrium (LD) is the non-random association of alleles at two or more loci. When a particular allele at one locus is found together on the same chromosome with a specific allele at a

second locus more often than expected if the loci were segregating independently in a population, the loci are in disequilibrium.[101] Numerically, it is defined as the difference between observed and expected allelic distributions (assuming random distributions).[101] The degree of LD between SNPs is determined by the r^2-value. r^2 values range from 0 to 1 and are directly proportional to the percentage of the non-centrality parameter shared between a pair of loci.[101] The non-centrality parameter is a measure of how different the true distribution is from the assumption that SNPs are independent events occurring by chance. Thus, for association tagging, r^2 is a measure of power, with values closer to 1 representing more power.

A common approach to LD gene mapping involves firstly identifying a gene region of interest and the degree of LD between SNPs in that region. Next, a subset of highly informative SNP markers (tagging-SNPs [tSNPs]) is identified at a pre-determined r^2 value (usually $r^2 \geq 0.8$) that will capture the majority of genetic variation across that gene region of interest. The tSNPs are supplemented by significant / putatively functional SNPs that are of particular interest and therefore warrant direct genotyping. Finally, the tSNPs and supplementary SNPs are typed in the clinical material.

When undertaking an association study to identify polymorphisms that influence genetically complex traits such as disease or response to treatment, the use of LD and tSNPs has several advantages. It allows for more efficient genotyping protocols, promotes cost-effectiveness, and the data requires less rigorous correction for multiple comparisons given the smaller number of individual tests which effectively leads to an increase in study power. An $r^2 \geq 0.8$ is usually employed in such studies. When an $r^2 \geq 0.8$ is applied, the majority of SNPs will be in LD at a level of $r^2 = 1$, some will have an r^2 between 0.9 and 1, and the remainder will have an r^2 between 0.8 and 0.9. This means that, although a given pair of SNPs may only be in LD 80% of the time (ie, $r^2 = 0.8$), the amount of genetic variation accurately captured by a set of tSNPs is much higher than 80%. Therefore, although an $r^2 \geq 0.8$ may appear lax on an individual SNP basis, the overall ability of a set of tSNPs to capture genetic variation is much higher, and is likely to provide an overall power in the region of 95%.

1.3.3. Genome-wide association studies

A genome-wide association study (GWAS) is an approach to genotyping that involves scanning markers across the whole genome, usually involving two cohorts of patients in a case-control manner. The principle of LD is used to infer haplotypes that capture the majority of common variation across the human genome, allowing for the maximal acquisition of genetic information through a manageable number of tests. Genome-wide association studies are useful in identifying genes associated with a given disease that were perhaps previously unknown to researchers.

Identifying new genetic associations serves to advance knowledge through two main routes; firstly, current understanding is challenged, forcing researchers to re-evaluate the knowledgebase to incorporate the genetic association. Secondly, it identifies gene-regions of particular interest, necessitating further analyses through candidate-gene approaches. This non-hypothesis driven approach to research is not without drawbacks. The massive numbers of tests required to cover the entire genome inevitably leads to false positive associations (type 1 errors). When statistical measures are employed to account for type 1 errors, the result may, somewhat paradoxically, produce a false negative association (type 2 error), which is again a consequence of the vast numbers of individual tests employed. In addition to this, GWAS is expensive to perform; money that could be spent performing hypothesis-driven candidate-gene analyses that are arguably more likely to yield significant results. Nonetheless, GWAS is a useful tool.

1.3.4. AED pharmacogenetics

Epilepsy is an ideal area for systematic pharmacogenetic research. The pharmacokinetic pathways in particular are well described, it is a common disorder, and its management is complicated by unpredictable efficacy, ADRs, and variability in inter-individual dose requirements. The optimal doses of some AEDs may differ more than ten-fold between individuals,[102] and even when factors known to influence dose requirement such as age, sex and weight have been accounted for, there is still a considerable unexplained variability in dose.

Previous investigations have demonstrated clear genetic influences on AED dosing. A splice site variant (rs3812718) in the *SCN1A* gene encoding the α_1-subunit of the neuronal VGSC has been shown to correlate with maximal doses of PHT and CBZ.[103] Dose and DNA information for 425 patients treated with CBZ was analysed; the allelic distribution for these participants was AA = 112, AG = 220, and GG = 93. The mean maximal dose of CBZ for each of these genotypes was 1313, 1225, and 1083mg/day, respectively. Considering arbitrary effects for each genotype, a significant difference between the means was established with an uncorrected p-value of 0.0051. When modelled as an additive effect, the significance increased (uncorrected p-value 0.0014), both of which remained significant after conservative Bonferroni correction for multiple tests.[103]

In the same study, a similar association was established between the *SCN1A* splice site variant and PHT maximal dose. Dose and DNA information for 281 participants was analysed; the distribution of genotypes among these was AA = 73, AG = 109, and GG = 60, which corresponded to mean maximal doses of 373, 340, and 326mg/day, respectively. An unrestricted model revealed a significant association (p = 0.014, uncorrected), and an additive model again revealed a stronger association (p =

0.0045, uncorrected). After Bonferroni correction for seven independent tests, the additive model remained significant, whereas the unrestricted model showed only a trend.[103]

Another study by Makmor-Bakry and colleagues found that maintenance dose of CBZ is associated with the genotype of two known functional variants in the *EPHX1* gene which encodes microsomal epoxide hydrolase.[104] This study analysed 70 patients who were deemed to be well controlled (≥50% reduction in seizure frequency) on CBZ monotherapy. None of the seven genotyped variables or demographic variables was of sufficient strength to influence CBZ dosing in isolation. However, when *EPHX1* c.337T>C (rs1051740) and *EPHX1* c.416A>G (rs2234922) were entered into a multivariate model with age as a covariate, a significant association with the maintenance dose of CBZ was revealed ($r^2 = 0.362$, $p = 0.002$).[104]

These studies show that a pharmacogenetic approach to AED dosing can explain at least some of the observed variation. By taking a similar approach, this project aims to determine whether LTG dosing in epilepsy is influenced by genetic variation in five candidate genes; *SLC22A1, UGT1A4, SCN1A, SCN2A* and *SCN3A*, each of which is discussed in more detail below.

1.3.4.1. *SLC22A1*

The solute carrier (SLC) family of membrane transport proteins can be divided into 47 subfamilies according to their solute-specificity. Each of these subfamilies can be further separated into various isoforms, giving rise to more than 300 SLC family members in total.[105]

SLC22 is one subfamily of the SLC family, the corresponding genes for which are found on the long arm of chromosome 6.[106] The first transporter of the SLC22 subfamily was initially reported by Grundemann and colleagues in 1994; the rat organic cation transporter OCT1, encoded for by the *SLC22A1* gene.[107] In rats, OCT1 is expressed in the liver, kidney and intestine, whereas in humans it is expressed primarily in the liver where it is responsible for the entry of organic cations into hepatocytes.[108] Since the discovery of OCT1, further investigation has revealed more SLC22 isoforms. These include two additional organic cation transporters, OCT2 and OCT3 (encoded by *SLC22A2* and *SLC22A3,* respectively), four organic anion transporters, OAT1 to OAT4 (encoded by *SLC22A6*, *7*, *8*, *11,* respectively), three transporters for carnitine and/or cations, OCTN1, OCTN2 and hCT2 (encoded by *SLC22A4*, *SLC22A5*, *SLC22A16,* respectively), and one urate transporter, URAT1 (encoded by *SLC22A12*).[109-114]

The OCTs have been extensively studied in order to clarify their physiological functions. OCT1 to OCT3 proteins are known to transport a multitude of endogenous compounds *in vitro*, with varying substrate specificity between each isoform. In humans, OCT1 has been shown to transport the

neurotransmitter acetylcholine,[115, 116] as well as the metabolite guanidine.[117-122] In comparison, OCT3 has been shown to transport a wider array of neurotransmitters including epinephrine,[115] histamine[115, 121-123] and norepinephrine,[115, 123, 124] but lacks the ability to transport metabolites. OCT2 has a wider spectrum of substrates; the neurotransmitters acetylcholine,[115, 116] dopamine,[115, 119, 123] epinephrine,[115] histamine,[115, 121-123] norepinephrine,[115, 123, 124] and serotonin,[115, 119, 121, 125] and the metabolites choline,[120, 121, 125-129] creatinine,[118] and guanidine[117-122] All three OCTs have also been shown to transport the amino acid metabolite and putative neurotransmitter agmatine.[130]

In addition to their aforementioned endogenous substrate specificity, the OCTs are also known to transport a wide array of xenobiotics and pharmaceutical drugs. Both OCT1 and OCT2 have been shown to transport the histamine H_2 receptor antagonist cimetidine,[118, 120, 121, 127, 128, 131-133] the antimalarial quinine,[118, 120, 131, 132, 134] and the antidiabetic compound metformin.[118, 121, 135, 136] As with the endogenous compounds, there are inter-isoform substrate specificity variations, with OCT1 involved in the transport of the sodium channel blocker quinidine,[118, 120, 125, 127, 129, 131, 132] in addition to the antiviral agents aciclovir and ganciclovir.[137] OCT2 is believed to transport a minimum of eight additional drugs, and OCT3 is known to transport at least two further pharmaceutical compounds.[138]

There has been substantial research into the impact of variation within the *SLC22A1* gene on the function of OCT1. Kerb and colleagues identified five polymorphisms that result in altered amino acid sequence across *SLC22A1* in a population of 57 Caucasians; Arg61Cys (rs12208357), Cys88Arg (rs55918055), Phe160Leu (rs683369), Gly401Ser (rs34130495) and Met420del (rs35167514), with reported minor allele frequencies (MAFs) of 0.091, 0.006, 0.22, 0.032 and 0.16 respectively.[139] They showed that the variants of Arg61Cys, Cys88Arg and Gly401Ser reduced uptake of 0.1μM [^{3}H] 1-methyl-4-phenylpyridinium ([^{3}H]MPP), a known substrate for OCT1, to 30, 1.4 and 0.9% respectively in comparison to wild-type. Furthermore, the Cys88Arg and Gly401Ser mutants were also found to exhibit altered substrate selectivity.[139]

Shu and colleagues[140] have further investigated human variants of OCT1. They expanded upon the findings of an earlier study on 247 ethnically diverse participants, in which 15 protein variants in OCT1 were detected, with observed changes in both loop (nine variants) and transmembrane domains (six variants).[141] Shu and co-workers determined the impact of these variants on protein function, and observed that five exhibited decreased function (Arg61Cys (rs12208357), Gly220Val [rs36103319], Pro341Leu [rs2282143], Gly401Ser (rs34130495), and Gly465Arg [rs34059508]), and one had increased function (Ser14Phe [rs34447885]), through quantifying transport of [^{3}H]MPP.[140] These findings were in keeping with those of Kerb and colleagues, aside from the failure to replicate

reduced function of Cys88Arg (rs55918055) due to its absence from the study population; perhaps unsurprising given its relative scarcity (MAF = 0.006).

The high variability of OCT1 may have implications for drug response. The transporter interacts with a variety of structurally diverse compounds, and controls access to drug metabolising enzymes in the liver.[114, 126, 132, 142, 143] It is possible that genetic variation in the *SLC22A1* region may influence either expression or function of the resultant OCT1 protein. Variation in the *SLC22A1* gene has been shown to influence protein function *in vitro*; six amino acid alterations have been shown to result in reduced transport activity, and one produces a phenotype with increased action.[139, 140] In addition to this, altered substrate selectivity has been observed in two of these variants.[139]

These *in vitro* observations may provide an explanation for clinically observed differences in response to drugs. Variability in the *SLC22A1* region could impact upon the kinetics of substrates for OCT1, altering the duration of exposure or volume of distribution of substrate drugs, with implications for drug efficacy, safety, and tolerability. OCT1 has very recently been implicated in the transport of LTG both into hepatocytes and at the blood-brain-barrier (BBB) (unpublished data, University of Liverpool), and so poses an ideal target for candidate gene analysis in this project.

1.3.4.2. *UGT1A4*

UGTs are reticulum-bound enzymes found throughout the human hepato-gastrointestinal tract, and are considered major determinants in phase II drug metabolism.[144] There are many known substrates for UGT1A4, including amitriptyline,[145] clozapine,[146] tamoxifen,[147] and LTG.[59]

The UGTs are a superfamily of enzymes, and can be classified into two families; UGT1 and UGT2. Eight UGT1 proteins have been identified in humans, to date; UGT 1A1, 1A3, 1A4, 1A6, 1A7, 1A8, 1A9, 1A10. UGT2 can be subdivided into UGT2A and UGT2B according to evolutionary divergence and homology; UGT 2A1, 2B4, 2B7, 2B10, 2B11, 2B15, 2B17, 2B28.[148, 149] Together, these 16 proteins act to catalyse the glucuronidation process in humans.

Inherited deficiencies of the UGTs are known to result in genetic disorders associated with hyperbilirubinaemia. Gilbert's syndrome is characterised by mild unconjugated non-haemolytic hyperbilirubinaemia that affects approximately 5% of the population.[150] The genetic basis for the disease is a TA insertion polymorphism into the promoter region of the *UGT1A1* gene, which has a MAF of 0.4 in the Caucasian population.[150] Gilbert's syndrome exists on a spectrum of severity ranging from mild hyperbilirubinaemia to life-threatening jaundice, and although it fails to become clinically apparent in a majority of patients, it illustrates how genetic variation in the UGTs can affect phenotype.

The UGT1 locus, which is located on chromosome 2q37, encodes for all UGT1 family members. There are nine functional members of the UGT1A subfamily,[151] with varying tissue-specific expression throughout the hepato-gastrointestinal tract. Strassburg and colleagues have investigated tissue-expression of UGTs extensively, and have concluded that UGT1A1, UGT1A3, UGT1A4, UGT1A6, and UGT1A9 are present in human liver, whereas UGT1A7 (oesophagus and stomach), UGT1A8 (oesophagus and colon), and UGT1A10 (oesophagus, bile ducts, stomach and colon) are found predominantly in extra-hepatic epithelial tissues.[152-157]

Human *UGT1A4* was first cloned in the early 1990s.[158] It is expressed predominantly in the liver, but also in the bile ducts, colon and small intestine.[152, 153] *UGT1A4* has become of increasing interest in recent years, with various polymorphisms resulting in altered enzyme activity. Ehmer and co-workers identified two variants that changed both the catalytic activity and specificity of UGT1A4 towards substrates.[159] Both were in the first exon of *UGT1A4*; at codon 24 a C>A transversion at position 1 (rs6755571) was observed leading to a proline to threonine amino acid change (P24T), and at codon 48 a T>G transversion at position 1 (rs2011425) produced a leucine to valine amino acid change (L48V). P24T has a reported MAF of 0.05-0.09 in Caucasian populations, and L48V has a MAF of 0.14 in Caucasians. In addition, there have been reports of variants in the 5'-upstream region of *UGT1A4* in Japanese patients, although their functional relevance has yet to be established.[160] The presence of SNPs within the coding regions of UGT genes may lead to alteration of specific catalytic activities. UGT1A4 is believed to be responsible for 75% of LTG glucuronidation;[66] polymorphisms within the *UGT1A4* gene may result in alteration of enzyme activity, and as a result explain at least part of the observed inter-individual variation in response to LTG.

1.3.4.3. *SCN1A*, *SCN2A*, *SCN3A*

In the brain, VGSCs are complexes of a 260-kDa α-subunit in association with auxiliary β-subunits (β1-β4) of 33 to 36kDa.[161] The α-subunit contains the voltage sensor and the ion-conducting pore in four internally repeated domains (I-IV), each consisting of six α-helical transmembrane segments (S1-S6) and a pore loop connecting S5 and S6.[161] The β-subunits are known to modify the kinetics and voltage dependence of gating, serving as cell-adhesion-molecules that interact with the extracellular matrix, other cell-adhesion molecules, and the cytoskeleton.[162, 163]

There are nine functional VGSC α-subunits in the human genome.[161] The proteins of these channels are named $Na_v1.1$ through to $Na_v1.9$, each of which is encoded by a corresponding *SCN*- gene. $Na_v1.1$, $Na_v1.2$ and $Na_v1.3$ are encoded by *SCN1A*, *SCN2A* and *SCN3A,* respectively. Together with $Na_v1.6$ (*SCN8A*), these three α-subunit proteins ($Na_v1.1$, $Na_v1.2$ and $Na_v1.3$) represent the primary sodium channels in the mammalian central nervous system.[161, 164-166]

Genetic variation in *SCN1A* has been implicated in the pathogenesis of severe myoclonic epilepsy of infancy (SMEI) as well as generalised epilepsy with febrile seizures plus (GEFS+).[167] More than 600 mutations in *SCN1A* have been identified, accounting for in excess of 70% of all cases of SMEI.[168] Approximately half the SMEI-causing mutations in *SCN1A* are nonsense or frameshift mutations that result in protein truncation, and are found randomly distributed throughout the gene.[168] The remaining SMEI-causing *SCN1A* mutations are predominantly missense mutations located in transmembrane segments of the protein where they may prevent channel expression or severely impair channel function.[168] The 30% of SMEI cases for which a genetic basis has yet to be established could, in theory, result from mutations in regulatory regions that govern *SCN1A* transcription.[167] More than 20 different *SCN1A* mutations have been identified in people with GEFS+, accounting for approximately 10% of cases.[167] The majority of these are missense mutations located primarily within transmembrane segments of $Na_v1.1$ channels.[168] As in SMEI, sodium channel β-subunit mutations have been shown to cause GEFS+ through indirect impairment of the expression and function of $Na_v1.1$ α-subunit containing channels.[169]

As discussed previously, genetic variation in *SCN1A* has been associated with maximum doses of CBZ and PHT.[103] A splice-site variant (rs3812718) showed significant associations with doses of both agents, in particular when analysed in additive models which remained statistically significant after Bonferroni correction for multiple comparisons. There is evidence to suggest that variation in *SCN2A* may affect response to AED therapy. Recently, in a study by Kwan and colleagues, an association was suggested between an *SCN2A* polymorphism (*SCN2A* IVS7-32A>G, rs2304016) and AED responsiveness. A-alleles were found to be associated with drug resistance (odds ratio = 2.1, 95% CI 1.2-3.7, p=0.007) in the 471 Chinese epilepsy patients analysed. In addition to this, a weak association was reported by Sills and colleagues in abstract form between the R19K polymorphism of *SCN2A* and resistance to AED treatment.[170] Finally, genetic variations in *SCN2A* have been reported in benign familial neonatal infantile spasms (BFNIS), with at least eight families now reported with missense mutations causative of BFNIS.[171]

Evidence of genetic variation in *SCN3A* resulting in epilepsy or having an impact on AED dosing is scarce. A single patient with partial epilepsy and a mutation in *SCN3A* has been described, resulting from a heterozygous K354Q missense mutation.[172] Despite this, *SCN3A* expression appears widespread throughout in the human adult brain,[173] and as such it proves a candidate for genotype analysis.

Overall, *SCN1A*, *SCN2A*, and *SCN3A* provide ideal targets for genotyping in this project. Genetic variation across these genes has been shown to result in various epilepsy syndromes, and

considering that the VGSC is its primary pharmacodynamic target, variation in these genes could theoretically influence dose requirement of and response to LTG. Moreover, variation in *SCN1A* has been shown to influence dosing of CBZ and PHT, both of which exert their anticonvulsant effects via sodium channels;[103] plausibly, LTG dosing could be affected in a similar manner.

1.4. Aims of Study

The aim of this project was to determine whether genetic variation is responsible for the observed variability in effective doses of LTG in a well-characterised, prospectively followed cohort of patients with newly-treated epilepsy. Five genes were selected for analysis as a result of their involvement in the pharmacokinetic and pharmacodynamic profiles of LTG; *SLC22A1*, *UGT1A4*, *SCN1A*, *SCN2A*, and *SCN3A*. The study involved patients enrolled on the SANAD trial which randomised two-thousand, four-hundred and thirty-seven patients with newly diagnosed epilepsy from Jan 12, 1999 to Aug 31, 2004. Participants that achieved seizure freedom for a period of at least twelve-months on a stable dose of LTG monotherapy were selected for analysis. Maintenance dose was used as the outcome measure as it reflects a combined measure of efficacy and tolerability, both of which are required in order for an AED to be clinically acceptable. Genotype data from a recent GWAS on the SANAD cohort of patients was supplemented with additional tSNP genotype data to capture all putatively functional variants in the five genes of interest. It was anticipated that a set of reliable genetic markers would be identified that explained a large percentage of inter-individual dose requirement for LTG. Ultimately, these markers may be employed prospectively to determine optimum doses for effective seizure control in individual patients.

Chapter 2: – Methods

2.1. Patient Selection

2.1.1. Patient Cohort

Participants were from the SANAD trial[174] which randomised two-thousand, four-hundred and thirty-seven newly-diagnosed epilepsy patients between January 12th 1999 and August 31st 2004, and collected follow up data where possible until January 13th 2006. The decision to collect DNA samples from SANAD participants was taken after commencement of the trial, resulting in 985 DNA samples being successfully obtained.

2.1.2. Eligible Participants

During the SANAD trial, the clinician recorded information regarding patient demographics and seizure history at recruitment, with details of treatment regimen and effectiveness documented at subsequent follow-up appointments. Patients were to be seen for follow-up at three, six, and twelve months and at successive yearly intervals from the date of randomisation. More frequent appointments were made if clinically indicated. Where patients failed to attend clinic appointments, follow up was achieved through directly contacting patients by telephone, or indirectly via the general practitioner.

Following completion of the trial in 2006, an electronic database of the dosing data for all participants was assembled. This database was manually filtered to include only those participants with available DNA samples, and a shortlist compiled of those who appeared to have taken a stable dose of LTG monotherapy for a period of at least 12 consecutive months.

These shortlisted participants subsequently had their case-notes reviewed to determine whether they met the following inclusion criteria:

- Seizure free for a period of 12 consecutive months
- Stable dose of LTG throughout this period
- LTG monotherapy

A total of 104 participants met the above inclusion criteria, having achieved seizure freedom for a period of at least 12 consecutive months on a stable dose of LTG monotherapy, and were therefore selected for genotype analysis.

2.1.3. Data Collection

For each eligible participant, details of the following factors were collected from the patient notes and assembled into an electronic database:

- Maintenance dose of LTG
- Epilepsy type
- Sex
- Number of seizures before randomisation
- Age at which maintenance dose was initiated
- Previous AED use

Maintenance dose: Maintenance dose is a reflection of drug effectiveness. In turn, drug effectiveness is a combined measure of efficacy and tolerability. The efficacy of an AED is judged on its ability to prevent seizures, whilst tolerability is a reflection of how acceptable a drug is in terms of side-effects. Maintenance dose was chosen to reflect drug dosing in a clinical setting. For each participant, this was the stable daily dose administered during the period of seizure-freedom. During the SANAD trial, AEDs were prescribed as per everyday clinical practice; for LTG, standard practice is to initiate treatment at a low starting dose and to titrate up over a period of 6-8 weeks until a pre-determined target dose is reached. Dose adjustments occur thereafter according to persistence of seizures or emergence of adverse effects.

Epilepsy Type: Participants were classified into one of three categories according to the clinicians' notes on epilepsy type:

- Idiopathic generalised epilepsy (IGE)
- Localisation-related epilepsy (LRE)
- Unclassifiable epilepsy (UNC)

Sex: Participant sex had been documented in the notes and was subsequently extracted.

Number of seizures before randomisation: Prior to randomisation into the SANAD trial, details of previous seizures had been documented, including number of episodes of each type of seizure. The total number of previous seizures was calculated from these data.

Age at which maintenance dose initiated*:* Documented within the patient notes were details of the dosing regimen at each follow-up appointment. Once a 12 month seizure-free period on a stable dose of LTG had been identified, the date at which this maintenance dose had been initiated was

recorded. Subsequently, using each patient's date of birth, the age at maintenance dose initiation was calculated.

Previous AED use: The SANAD trial did not exclude participants if they had received a single AED on a monotherapy dosing regimen before the trial commenced. This information had been recorded in a 'yes/no' format by the clinician at randomisation. In addition to this, participants may have been randomised to receive a different AED from LTG at the start of the SANAD trial but were ultimately controlled on LTG after failure of the originally assigned drug. Regardless of when exposure occurred, participants who had received an alternative AED prior to achieving seizure-freedom with LTG were identified and "previous treatment" noted in the study database.

2.2. SNP Identification

2.2.1. HapMap

The International HapMap project was launched in 2002 with the aim of providing a public resource to accelerate medical genetic research.[175] The objective was to genotype at least one common SNP every 5 kilobases (kb) across the euchromatic portion of the genome in 270 individuals from four geographically diverse populations.[176, 177] Since completion, this information has been made publically available, and was utilised to identify SNPs in the gene regions of interest.

Each individual gene was searched using the HapMap search tool using the Hapmap Genome Browser (Phase 1 & 2, full dataset) accessible via http://hapmap.ncbi.nlm.nih.gov/. The chromosome positions were noted, and then the search criteria for each gene expanded manually by 10kB both upstream and downstream to capture the 5' and 3' flanking regions for each gene (*Table 2.1*).

Table 2.1: Position and expanded search criteria for each gene

Gene Search	Gene Position	Expanded Search Criteria
SLC22A1	Chr6:160,462,853..160,499,740	Chr6:160,452,853..160,509,740
UGT1A4	Chr2:234,292,177..234,346,684	Chr2:234,282,177..234,356,684
SCN1A	Chr2:166,553,917..166,638,395	Chr2:166,543,917..166,648,395
SCN2A	Chr2:165,804,158..165,957,066	Chr2:165,794,158..165,967,066
SCN3A	Chr2:165,652,286..165,768,799	Chr2:165,642,276..165,778,823

Once the relevant gene regions had been identified, the SNP genotype data for these regions was downloaded using the 'Reports and Analysis' utility, ensuring the following parameters were selected:

- Population – CEU (Caucasian-European)

- Strand – rs (SNP reference number)
- Output format – save to disk

This resulted in the download of all SNPs with a known frequency in a Caucasian European population in the gene regions of interest.

2.2.2. Haploview 4.2

Haploview 4.2 is a bioinformatics software tool that can analyse patterns of LD in a given set of genetic data. Once a set of genetic data has been analysed, Haploview is able to provide a set of tSNPs that will capture the genetic variation across the specified region(s).

2.2.2.1. Tagging SNPs

The principle of LD was used to generate a list of tSNPs that would allow for inference of all identified SNPs across the gene regions of interest. Haploview was instructed to analyse the downloaded SNP genotype data files from HapMap and generate a list of tSNPs for each. Under 'HapMap Format' of the Haploview start-up menu, each SNP genotype data file was opened and following parameters entered on the 'Check Markers' tab:

- HW (Hardy-Weinberg) p-value = 0.001
- Minimum genotype = 75%
- Max# Mendelian errors = 1
- Minimum MAF = 0.05

The following limits were selected under the 'tagger' tab:

- Pairwise tagging only
- r^2 threshold = 0.8
- LOD threshold for multi-marker tests = 3.0
- Min distance between tags = 0bp
- Max distance between tags = blank

The Haploview programme was instructed to execute the 'Run Tagger' function. This grouped SNPs that were in LD together, and generated a list of tSNPs with a minimum MAF of 5% which captured approximately 95% of the genetic variation across each gene region at the $r^2 = 0.8$ level of certainty. The tSNP data for each of the five gene regions of interest has been included in Appendix III and is summarised in *Table 2.2*.

Overall, 587 SNPs in the five gene-regions of interest were identified from the HapMap database. Following LD grouping, 118 tSNPs were identified that would allow inference of all 587 SNP genotypes.

Table 2.2: Summary of tSNP data for each gene

Gene	Number of tSNPs	Total number of SNPs captured
SLC22A1	27	54
UGT1A4	20	141
SCN1A	16	116
SCN2A	35	193
SCN3A	20	83
Total	118	587

2.2.3. GWAS data

Genotype data was available from a recent GWAS analysis on the SANAD cohort of patients. Genotyping was performed at the Wellcome Trust Sanger Centre (Hinxton, UK) on an Illumina 660k SNP chip with the aim of identifying genome-wide associations with response to AED treatment in the SANAD cohort as a whole. At this time, that analysis is ongoing, but the availability of GWAS-derived genotypes made a significant contribution to this project and limited the number of novel genotypings that were required to complete the tSNP analysis of the five candidate genes of interest.

The GWAS genotyped 47,993 and 58,599 SNPs across chromosomes 2 and 6, respectively. A list of these GWAS SNPs was available for chromosomes 2 and 6, along with details of SNP position, and the minor and major alleles at each SNP locus (Appendix IV).

The GWAS SNPs were sorted by ascending chromosome position, and the same search criteria used for SNP identification on HapMap applied. Any SNPs that lay between the upper and lower chromosome position boundaries were selected, the results of which are shown in *Table 2.3*. Overall, this resulted in the identification of 166 GWAS SNPs within the gene-regions of interest.

Table 2.3: GWAS SNP identification

Gene	Gene Location	Lower Search Boundary	Upper Search Boundary	Number of GWAS SNPs
SLC22A1	Chr6:160,462,853..160,499,740	160,452,853	160,509,740	37
UGT1A4	Chr2:234,292,177..234,346,684	234,282,177	234,356,684	38
SCN1A	Chr2:166,553,917..166,638,395	166,543,917	166,648,395	27
SCN2A	Chr2:165,804,158..165,957,066	165,794,158	165,967,066	45
SCN3A	Chr2:165,652,286..165,768,799	165,642,286	165,778,799	19
Total				166

2.2.4. Identification of Outstanding tSNPs

Genotype data were available for the identified 166 GWAS SNPs. The next stage was to determine whether these GWAS SNPs were in LD with any SNPs from the tSNP identification process. The tSNP identification process had discovered that 118 tSNPs could capture the majority of the genetic variation across the five gene-regions of interest at an r^2=0.8. If a GWAS SNP was found to be in LD with one of the tSNPs, genotyping of that tSNP was no longer required.

For each gene region, the identified GWAS SNPs were checked against the corresponding output from Haploview. Overall, the GWAS SNPs were in LD with 81 of the 118 tSNPs previously identified. This left 37 outstanding tSNPs that required genotyping in order to capture approximately 95% of the genetic variation across the five gene-regions of interest (*Table 2.4*).

Table 2.4: Summary of outstanding tSNPs

Gene	Number of identified tSNPs	Number of tSNPs in LD with GWAS SNPs	Number of outstanding tSNPs
SLC22A1	27	20	7
UGT1A4	20	13	7
SCN1A	16	10	6
SCN2A	35	24	11
SCN3A	20	14	6
Total	118	81	37

2.2.5. Supplementary SNPs

In addition to the 37 outstanding tSNPs, supplementary SNPs were selected for novel genotyping following a literature search. Putatively functional variants were selected for direct genotyping if they had not yet been captured by the GWAS and were absent from the devised list of outstanding tSNPs. SNPs were manually selected for genotyping if they met all the inclusion criteria shown in *Table 2.5* and excluded if they met any of the exclusion criteria. This process resulted in the inclusion of an additional nine SNPs, bringing the total number of SNPs for novel genotyping to 46.

Table 2.5: Inclusion / exclusion criteria for supplementary SNPs

Inclusion Criteria	Exclusion Criteria
Located within a gene-region of interest	In LD with a GWAS SNP
Known MAF > 0.01 in a Caucasian population	In LD with an outstanding tSNP
Known functional SNP or significant association mentioned in literature	

Abbreviations: MAF = minor allele frequency; SNP = single nucleotide polymorphism; tSNP = tagging SNP; GWAS = genome-wide association study; LD = linkage disequilibrium

2.3. Experimental Design

2.3.1. MySequenom®

Once a list of SNPs had been compiled, an experiment was designed to genotype them using a Sequenom iPLEX® matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF MS) platform. This was achieved through use of the MySequenom® 'human genotyping tool' utility accessible via http://www.mysequenom.com and consisted of five steps:

rs-Sequence Retriever: The list of SNPs to be genotyped was entered into the first stage of the tool. This electronically retrieved the DNA sequence for each SNP from NCBI dbSNP (http://www.ncbi.nlm.nih.gov/projects/SNP), retrieving approximately 200kB both upstream and downstream from each SNP.

ProxSNP®: The data obtained from the rs-sequence retriever was inputted into the ProxSNP® application. This re-formatted the SNP sequences to ensure any proximal SNPs near to the assay SNP were marked and accounted for accordingly. Any sequence with one or more proximal SNPs too close to the assay SNP was rejected at this stage. The output from ProxSNP® consisted of a list of rs-numbers with their corresponding 100kB flanking regions; any unsuitable SNPs were reported separately and excluded from future stages.

PreEXTEND®: The output of successful SNPs from ProxSNP® was entered into the PreEXTEND® application; this modelled each SNP sequence as a potential primer, validating each sequence to ensure that only the region of the SNP of interest would be amplified. SNPs were rejected if their primers were found to anneal to multiple sites.

Assay Design: Primers were designed for the SNPs which remained after the PreEXTEND® process. For each SNP, the assay design application created a forward PCR primer, a reverse PCR primer, and an extension primer.

PleXTEND®: After assay design, the assays were checked by the PleEXTEND® application to ensure that no unintended amplification products would be produced.

2.3.2. Finalised Experimental Design

A total of 46 SNPs were identified as candidates for novel genotyping, however 7 SNPs were excluded at various stages during the experimental design process as shown in *Table 2.6*. This resulted in a

final experimental design consisting of 39 SNPs spread across two plexes; 21 SNPs in plex 1, and 18 SNPs in plex 2 (*Table 2.7*).

Table 2.6: Excluded SNPs

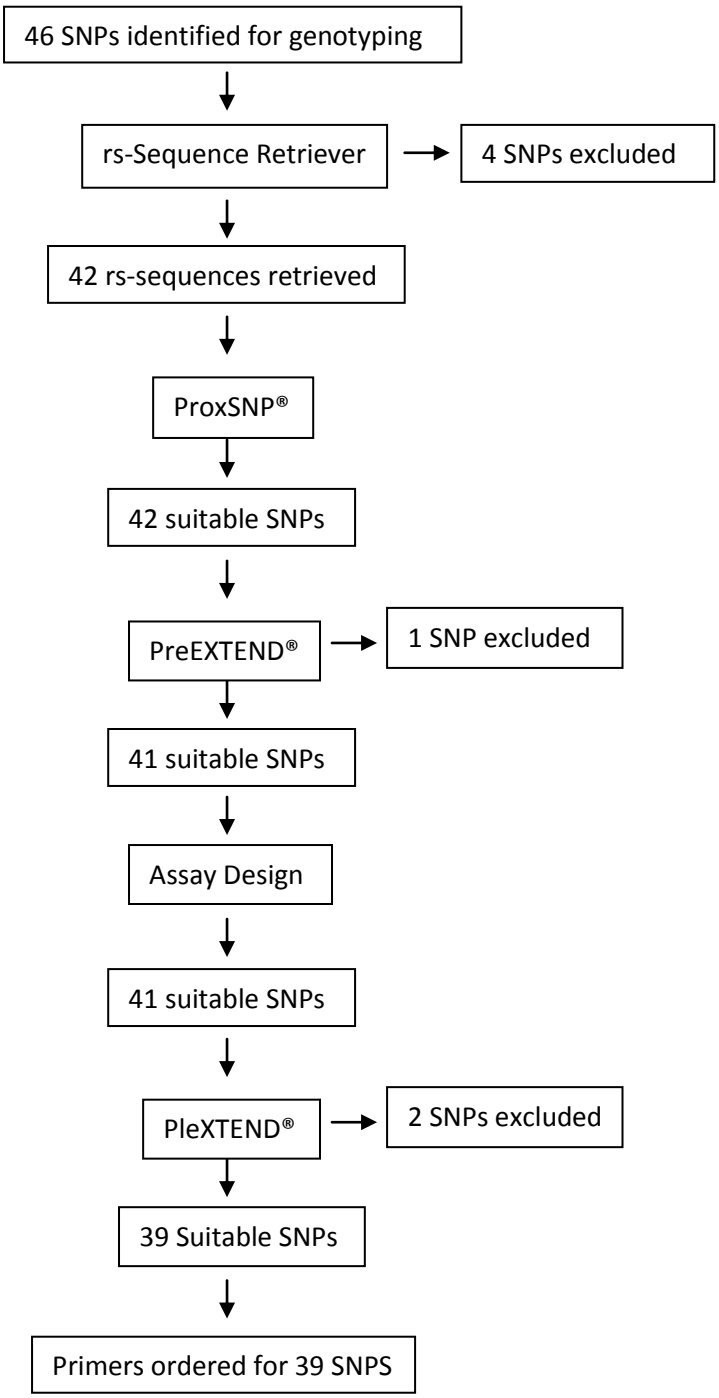
SNP	Gene Region	Stage of Exclusion	Reason for Exclusion
rs76921794	*SCN1A*	rs-Sequence Retriever	Cannot get document summary
rs78391141	*SCN1A*	rs-Sequence Retriever	Cannot get document summary
rs77216276	*SCN1A*	rs-Sequence Retriever	Cannot get document summary
rs78675912	*SCN3A*	rs-Sequence Retriever	Cannot get document summary
rs2011404	*UGT1A4*	PreEXTEND®	Multiple eXTEND hits for scanned primer triplets
rs28946889	*UGT1A4*	PleXTEND®	High dimer potential (0.67) for forward extend primer
rs17744737	*SCN1A*	PleXTEND®	Could not multiplex SNP to minimum multiplexing level

Table 2.7: Final plex design

Plex 1			Plex 2		
rs-number:	**Gene Region**	**Identification Method:**	**rs-number:**	**Gene Region**	**Identification Method:**
rs12151636	*SCN1A*	HapMap	rs13406905	*SCN1A*	HapMap
rs478389	*SCN1A*	HapMap	rs1461195	*SCN1A*	HapMap
rs17791817	*SCN1A*	HapMap	rs6751613	*SCN2A*	HapMap
rs7566636	*SCN2A*	HapMap	rs6705474	*SCN2A*	HapMap
rs2060199	*SCN2A*	HapMap	rs16850317	*SCN2A*	HapMap
rs17183814	*SCN2A*	HapMap	rs353116	*SCN2A*	HapMap
rs6741147	*SCN2A*	HapMap	rs2228988	*SCN2A*	Literature Search
rs1368234	*SCN2A*	HapMap	rs3213904	*SCN3A*	HapMap
rs7573433	*SCN2A*	HapMap	rs2028364	*SCN3A*	HapMap
rs7600082	*SCN2A*	HapMap	rs6719780	*SCN3A*	HapMap
rs1946892	*SCN3A*	HapMap	rs7596422	*SCN3A*	HapMap
rs11677254	*SCN3A*	HapMap	rs3771342	*UGT1A4*	HapMap
rs3806591	*UGT1A4*	HapMap	rs45621441	*UGT1A4*	Literature Search
rs6431633	*UGT1A4*	HapMap	rs10203853	*UGT1A4*	HapMap
rs35956182	*SLC22A1*	Literature Search	rs929596	*UGT1A4*	HapMap
rs619598	*SLC22A1*	HapMap	rs12208357	*SLC22A1*	Literature Search
rs34059508	*SLC22A1*	Literature Search	rs644992	*SLC22A1*	HapMap
rs9456505	*SLC22A1*	HapMap	rs7773429	*SLC22A1*	HapMap
rs10455864	*SLC22A1*	HapMap			
rs10455868	*SLC22A1*	HapMap			
rs461473	*SLC22A1*	HapMap			

Primers for the finalised experimental design across two plexes were ordered from Metabion® (Martinsried, Germany); forward and reverse primers were both at 50µM concentrations, and extension primers were supplied at a concentration of 300µM.

2.3.3. Experimental Design Overview



2.4. Genotyping

2.4.1. Primers

A primer is a strand of nucleic acid that serves as a starting point for DNA synthesis. Typically they are approximately twenty bases in length, and are designed to be complementary to the sequence that is upstream to the DNA region of interest. PCR primers are designed to anneal upstream of a gene region of interest to allow amplification of the gene region containing a SNP. In contrast, extension primers are designed to anneal to the template DNA directly 5' to a variant loci so that a mutant DNA polymerase can carry out single-base extension (SBE).

2.4.1.1. PCR Primer Mix Preparation

Primers were required at a working concentration of 0.5µM for optimal activity. For each plex, the following quantities of reagents were combined in 1000µl eppendorfs:

Plex 1:

- 290µl Nanopure water
- 5µl of each forward primer (21 forward primers in total)
- 5µl of each reverse primer (21 reverse primers in total)

Plex 2:

- 320µl Nanopure water
- 5µl of each forward primer (18 forward primers in total)
- 5µl of each reverse primer (18 reverse primers in total)

For each plex, this resulted in a total PCR primer mix volume of 500µl, with each primer at a concentration of 0.5µM.

2.4.1.2. Extension Primer Mix Preparation

For each plex, the extension primers were arranged in ascending order of weight and assigned a number accordingly; there were 21 extension primers for plex 1 (numbered 1-21) and 18 extension primers for plex 2 (numbered 1-18).

For each plex, the primers were split into four groups depending on their weights as per the four-step adjustment protocol. For plex 1, primers 1-6 were assigned to group 1, primers 7-11 to group 2,

primers 12-16 to group 3, and primers 17-21 to group 4. A 11.67µl volume of each of the group 1 primers were mixed in an eppendorf, along with 15.50µl of each of the group 2 primers, 19.43µl of each of the group 3 primers, and 23.33µl of each of the group 4 primers. This resulted in primer concentrations of the group 1, 2, 3, and 4 primers of 7.00µM, 9.30µM, 11.67µM, and 14.00µM respectively.

The assembly of the plex 1 primer mix is summarised in *Table 2.8* along with the quantities required for assembly of plex 2. According to which of the four groups they had been assigned, pre-determined volumes of each primer were used to assemble each primer mix to achieve desired concentrations; this is the essence of the four-step adjustment protocol.

Table 2.8: Extension primer mix preparation summary

Extension Primer Group	**Final Concentration/ Primer**	**Volume/ Primer**	**Number of Primers per Group (primer numbers)**		**Total volume of primer group**	
			Plex 1	**Plex 2**	**Plex 1**	**Plex 2**
1	7.00µM	11.67µl	6 (1-6)	5 (1-5)	70.02µl	58.35µl
2	9.30µM	15.50µl	5 (7-11)	5 (6-10)	77.50µl	77.50µl
3	11.66µM	19.43µl	5 (12-16)	4 (11-14)	97.15µl	77.72µl
4	14.00µM	23.33µl	5 (17-21)	4 (15-18)	116.65µl	93.32µl
Nanopure Water					138.68µl	193.11µl
Total volume					500µl	500µl

2.4.2. Plating out the DNA

The SANAD DNA stock-aliquot was available at a concentration of 20ng/µl, which is the standard concentration used on the Sequenom MassARRAY® platform. DNA samples for 27 of the identified participants were unavailable due to inadequate volume of existing DNA holding, resulting in 77 DNA samples being available for genotyping.

A working-aliquot of these 77 DNA samples was assembled containing a maximum of 3µl of each sample. The aliquot was stored in the freezer at -20°C, and when required was allowed to thaw at room temperature, centrifuged at 2000 RPM for 2 minutes, and then stored on ice whilst samples were transferred to the final plate design.

A 1µl volume of each DNA sample at the 20ng/µl concentration was transferred into a 384-well PCR plate according to the plate design in Appendix V. The 384-well PCR plate was then covered with tissue and allowed to evaporate.

2.4.3. Polymerase Chain Reaction

2.4.3.1. Polymerase Chain Reaction Mix Preparation

Using the quantities of reagents outline in *Table 2.9*, a polymerase chain reaction (PCR) mix was prepared for each plex.

Table 2.9: PCR reagents

Reagents	Storage (temperature)	Volume Required
Nanopure water	Sterile sealed container (room temperature)	342µl
Plex 1 primer mix (0.5µM)	Fridge (5°C)	120µl
PCR Buffer (10x)	Freezer (-20°C)	75µl
$MgCl_2$ (25mM)	Freezer (-20° C)	39µl
dNTP mix (25mM)	Freezer (-20° C)	12µl
Hot Star Taq (5U/µl)	Freezer (-20° C)	12µl

The PCR buffer, $MgCl_2$ and deoxyribonucleotide triphosphate (dNTP) mix were removed from the freezer and allowed to thaw at room temperature. Once defrosted, they were mixed using a Whirlimixer® for 10 seconds each, and then centrifuged for two minutes at 2000 RPM.

The PCR enzyme was removed from the freezer, mixed for 10 seconds using a Whirlimixer®, centrifuged for two minutes at 2000 RPM, and stored on ice throughout the PCR mix process.

The primer mix was removed from the fridge, mixed for 10 seconds using a Whirlimixer®, and centrifuged for two minutes at 2000 RPM.

The reagents were combined in a 1000µl eppendorf, sealed and centrifuged for 2 minutes at 2000 RPM, and the reagents returned to their respective storage locations after use.

The above process was repeated, substituting the plex 2 primer mix instead of plex 1, resulting in two eppendorfs containing PCR mix; one for plex 1 and the other for plex 2.

2.4.3.2. Transfer of PCR Mix to 384-well PCR plate

A 60µl volume of the plex 1 PCR mix was transferred into each of the wells in the first column of a 96V microwell plate, and 60µl of the plex 2 PCR mix was transferred into each of the wells of the second column. The plate was then sealed and centrifuged for 2 minutes at 2000 RPM to remove any air bubbles. A 96V well plate was used ahead of a standard 96-well plate because the 'V' shaped wells minimised the volume errors often encountered when using a multi-channel pipette.

An 8-tip multi-channel pipette was used to transfer 5µl of the PCR mix from the first column of the 96V microwell plate into each of the wells of columns 1-6 of the 384-well PCR plate containing evaporated DNA. Similarly, 5µl of PCR mix from the second column of the 96V microwell plate was transferred into each of the wells of columns 7-12 of the 384-well PCR plate.

2.4.3.3. PCR Reaction

The 384-well PCR plate was sealed using an Adhesive PCR Seal (AB-0558), centrifuged for 2 minutes at 2000 RPM, and then transferred to the G-Storm® Thermocycler GS4 apparatus (Gene Technologies Ltd, UK). A thermal cycler compression pad was placed on top of the 384-well PCR plate, and the thermocycler instructed to run the 'PCR iPLEX' programme for a reaction volume of 5µl. The PCR process selectively amplified the sequences of DNA containing the identified SNPs, and can be summarised as follows:

Initialisation: The PCR mix was heated to 94.0°C and held for a period of 15 minutes.

Denaturation: The PCR mix was heated to 94°C for a period of 20 seconds, disrupting the hydrogen bonds between complementary bases and separating the double-stranded DNA-complex to yield two single DNA strands.

Annealing: Once the strands had been separated, the reaction temperature was lowered to 56°C for 30 seconds to allow the primers to anneal to their respective single-stranded DNA targets. Simultaneously, the Taq polymerase enzyme bound to the primer-template hybrid in preparation for the elongation phase.

Elongation: The reaction was then heated to 72°C for a period of 60 seconds, providing optimum conditions for the activity of the Taq polymerase enzyme. Taq polymerase synthesised complementary strands of DNA in the 5' to 3' direction through addition of complementary dNTPs with hydrogen bonds. A complementary DNA strand was formed in the 5' to 3' direction from each point at which a primer had annealed to the template DNA strand.

Repetition: The denaturation, annealing and elongation steps were repeated 45 times, with the products of each cycle becoming the template for subsequent cycles, theoretically producing an exponential increase in the quantity of relevant DNA per cycle.

Final Elongation: After 45 cycles, the reaction temperature was held at 72°C for three minutes to ensure that any remaining single-stranded DNA was fully extended.

Final hold: The temperature of the reaction was reduced to 5°C, which provided suitable storage conditions for up to a 12-hour period.

The PCR iPLEX® required approximately 2.5 hours to reach the final hold stage at which point the 384-well PCR plate was removed, centrifuged at 2000RPM for 2 minutes, and stored in the cold room at 0°C.

2.4.4. Shrimp Alkaline Phosphatase

The unincorporated dNTPs from the PCR amplification process were removed to avoid interference with reagents added at subsequent stages. This was achieved through the addition of shrimp alkaline phosphatase (SAP), an enzyme which resulted in the dephosphorylation of any excess dNTPs, thereby rendering them inactive in future reactions.

2.4.4.1. SAP Preparation

The reagents and volumes listed in *Table 2.10* were required for SAP-mix preparation.

Table 2.10: SAP reagents

SAP Reagents	Storage	Volume Required
Nanopure water	Sterile sealed container (room temperature)	405.39µl
hME buffer (10x)	Freezer (-20°C)	45.04µl
Shrimp Alkaline Phosphotase (SAP)	Freezer (-20°C)	79.49µl

The hME buffer was removed from the freezer and allowed to thaw at room temperature. The SAP and hME buffer were then mixed for 10 seconds each using the Whirlimixer®, and centrifuged for 2 minutes at 2000 RPM, with the SAP stored on ice thereafter.

The SAP reagents were mixed in a 1000µl eppendorf, centrifuged for 2 minutes at 2000 RPM, and then 66µl of the SAP solution was transferred into each well of the first column of a new 96V microwell plate. An 8-tip multichannel pipette was then used to transfer 10µl of the SAP solution from column 1 into each of columns 2-6 of the 96V microwell plate. The plate was then sealed and centrifuged for 2 minutes at 2000 RPM to remove any bubbles.

2.4.4.2. Addition of SAP to 384-well PCR plate

The Matrix Liquid Handler (MLH) was used to add 2µl of SAP solution to each of the wells containing DNA in the 384-well PCR plate. Prior to using the MLH, the 384-well PCR plate and the 96V microwell plate were sealed and centrifuged at 2000 RPM for 2 minutes, and their seals removed upon completion.

Once the weekly maintenance protocols had been adhered to, as per the manufacturer's guidelines, a new magazine of tips was inserted into the MLH and the "tip wash" programme initiated using the ControlMate software.

Once the "tip wash" programme had completed, the "SAP addition (96 to 384).cms" programme was selected. This prompted onscreen instructions for positioning of 384-well PCR plate and 96V microwell plate. Upon completion, the 384-well PCR plate was sealed using an adhesive PCR seal (AB-0558) and centrifuged for 2 minutes at 2000 RPM.

2.4.4.3. SAP Reaction

The 384-well PCR plate was placed in the G-Storm® Thermocycler GS4 apparatus with a thermal cycler compression pad on top, and the programme 'SAP iPLEX' selected for a reaction volume of 7µl. The programme consisted of the following steps:

Temperature increased to 37.0°C for 40 minutes. This provided optimum temperature conditions for the SAP to undertake its activity.

Temperature increased to 85.0°C for 5 minutes. This step served to irreversibly denature the SAP, removing any chance of interaction with subsequent reagents.

Storage at 10°C. The reaction products were held at this temperature indefinitely until removed from the thermocycler.

The SAP-iPLEX reaction required approximately 50 minutes to complete, after which the 384-well PCR plate was removed from the apparatus, sealed using AB-0558, and centrifuged for 2 minutes at 2000 RPM.

2.4.5. iPLEX® Gold

After PCR cleanup, the iPLEX® Gold reaction was performed. The reagents in *Table 2.11* and *Table 2.12* were required to assemble the iPLEX® Gold reaction cocktail solutions for plexes 1 and 2, respectively.

Table 2.11: Plex 1 iPLEX reagents

Plex 1 Reagents	Storage (Temperature)	Volume
Nanopure Water	Sterile sealed container (room temperature)	100.02µl
iPLEX®-Buffer (10x)	Freezer (-20°C)	26.50µl
iPLEX®-Termination mix	Freezer (-20°C)	26.50µl
Plex 1 Primer mix	Fridge (5°C)	106.51µl
iPLEX®-Enzyme	Freezer (-20°C)	5.30µl
Total		264.83µl

Table 2.12: Plex 2 iPLEX reagents

Plex 2 Reagents	Storage (Temperature)	Volume
Nanopure Water	Sterile sealed container (room temperature)	100.02µl
iPLEX®-Buffer (10x)	Freezer (-20°C)	26.50µl
iPLEX®-Termination mix	Freezer (-20°C)	26.50µl
Plex 2 Primer mix	Fridge (5°C)	106.51µl
iPLEX®-Enzyme	Freezer (-20°C)	5.30µl
Total		264.83µl

2.4.5.1. iPLEX® Reaction Cocktail Solution

The iPLEX®-buffer and iPLEX®-termination mix were allowed to thaw at room temperature. Each reagent apart from the Nanopure water was mixed using a Whirlimixer® for 10 seconds each, and then centrifuged for 2 minutes at 2000 RPM. The iPLEX®-enzyme was stored on ice throughout preparation, while the other reagents were stored at room temperature. The reagents for plex 1 and plex 2 were combined separately in 1000µl eppendorfs, mixed using a Whirlimixer® for 10 seconds each, and then centrifuged for 2 minutes at 2000 RPM. A 2µl volume of primer extension reaction cocktail solution was added to each well of the 384-PCR reaction plate containing PCR amplification products. This was then placed in the G-Storm® GS4 thermocycler and exposed to the necessary conditions (section 2.4.5.4.) for SBE to occur.

2.4.5.2. Addition of iPLEX® Reaction Cocktail Solution to 384-well PCR Plate

The MLH was used to add 2µl of iPLEX® Reaction Cocktail solution to each of the wells containing DNA in the 384-well PCR plate. As per SAP addition, weekly maintenance and tip wash protocols were followed prior to use of the MLH for iPLEX® reaction cocktail solution addition. Using a new 96V microwell plate, 33µl of the plex 1 iPLEX® reaction cocktail solution was transferred into each well of the first column, and 33µl of the plex 2 iPLEX® reaction cocktail solution into each well of the fourth column. Using an 8-tip multichannel pipette, 10µl of plex 1 cocktail solution was transferred from column 1 into columns 2 and 3, and similarly 10µl of plex 2 cocktail solution was transferred from column 4 into columns 5 and 6. The 384-well PCR plate and the 96V microwell plate were sealed and centrifuged at 2000 RPM for 2 minutes, and their seals removed upon completion. The programme “cocktail addition (96 to 384).cms” was initiated, and the onscreen instructions followed for positioning of the 384-well PCR plate and the 96V microwell plate. Upon completion, the 384-well PCR plate was sealed with an adhesive seal (AB-0558) and centrifuged for two minutes at 2000 RPM.

2.4.5.3. iPLEX® Reaction

The 384-well PCR plate was placed in the G-Storm® Thermocycler GS4 apparatus with a thermal cycler compression pad on top, and the programme 'extend iPlex.scr' initiated for a reaction volume of 9µl per well. This programme consisted of the following stages, required approximately 4 hours, and upon completion the plate was centrifuged for 2 minutes at 2000 RPM and stored in the cold room at 0°C:

Temperature 94°C for 30 seconds.
50 cycles of the following:
Temperature 94°C for 5 seconds
Temperature 52°C for 5 seconds
Temperature 80°C for 5 seconds
Temperature 52°C for 5 seconds
Temperature 80°C for 5 seconds
Temperature 52°C for 5 seconds
Temperature 80°C for 5 seconds
Temperature 52°C for 5 seconds
Temperature 80°C for 5 seconds
Temperature 72°C for 3 minutes
Temperature 4°C for 5 minutes
Stored at 5°C indefinitely.

During the SBE process outlined above, each extension primer anneals to its target directly 5' to the SNP locus. A mutant DNA polymerase (iPLEX®-enzyme) then incorporates a single terminator nucleotide (iPLEX®-termination mix) that is complementary to the base at the variant locus. This results in elongation of the extension primer by a single base unit, with a corresponding mass increase dependent upon which terminator nucleotide has been incorporated. The mass increase can be measured using MALDI-TOF MS, with SNP genotype interpreted accordingly.

2.4.6. Sample Conditioning

Upon completion of the Extend-iPLEX reaction, unincorporated products were desalted via cation exchange through addition of SpectroCLEAN® resin. This step was performed in order to maximise the MALDI-TOF MS resolution, in turn minimising false-positive results.

2.4.6.1. Preparation of SpectroCLEAN® Resin

Resin was spread over the first 12 columns of a 384-well dimple plate using a bevelled Perspex applicator, and allowed to dry for approximately 10 minutes. During this time the 384-well PCR plate was centrifuged for 2 minutes at 2000 RPM and the seal removed.

2.4.6.2. Addition of Nanopure Water

The MLH was used to add 16µl of Nanopure water to each well of the 384-well PCR plate. Using the ControlMate® software, the programme '16µl water addition.cms' was selected and the onscreen instructions followed for placement of the 384-well PCR plate and water bath filled with Nanopure water. Once complete the 384-well PCR plate was sealed, centrifuged for 2 minutes at 2000 RPM, and the seal removed again.

2.4.6.3. Addition of SpectroCLEAN® Resin

The 384-well PCR plate was inverted, placed on top of the resin dimple plate, then both were inverted to transfer the resin into the 384-well PCR plate. The plate was sealed again using Adhesive PCR Seal (AB-0558), secured between two polystyrene blocks in the Heidolph®-Reax 2 rotator, and set to rotate for 10 minutes on the lowest setting.

Once completed, the 384-well PCR plate was removed and centrifuged for 5 minutes at 3000 RPM. The extended centrifuge duration and speed was necessary to settle down all CleanRESIN® so that it was not transferred onto the matrix during the next step; resin particles embedded in the matrix can be detrimental to the accuracy and success of the MALDI-TOF analysis.

2.4.7. Dispensing

A MassARRAY® nanodispenser was used to dispense 15-25nl of the samples onto a 384 SpectroCHIP®. The samples were embedded into the 3-hydroxypicolinic acid matrix of the SpectroCHIP® where they underwent co-crystallisation in preparation for MALDI-TOF MS.

2.4.7.1. Transfer of iPLEX® Reaction Products onto a SpectroCHIP®

Per protocol steps were taken to prepare the nanodispenser for operation, including pin conditioning of the main head and single head, cleaning the main head and single head, and running the volume check programme to avoid poor dispensing quality. Once the nanodispenser had been prepared for use, the Spectropoint® programme was used to initiate the '384-384.tmf' command, with the appropriate dispensing speed values entered in order to achieve a spotting volume of 15-25nl. This programme transferred 15-25nl of each sample onto the SpectroCHIP®. MassARRAY Calibrant® was

removed from the freezer and allowed to thaw at room temperature before being centrifuged for 30 seconds at 2000 RPM. A 70µl volume was transferred into the calibrant reservoir of the nanodispenser, and the programme 'calibrant.tmf' initiated. This transferred calibrant onto the SpectroCHIP® to optimise performance of the MALDI-TOF apparatus.

2.4.8. Mass Spectrometry

The general principle of MS is to produce, separate, and detect gas phase ions (Figure 2). This was achieved through the following stages:

Irradiation of the matrix crystals. The matrix crystals were exposed to an ultraviolet laser of wavelength 337nm for 1 nanosecond. The matrix acted as a buffer, absorbing much of the energy supplied by the laser to prevent degradation of the single-base extended primers whilst allowing enough energy for them to become ionised. The result was structural decomposition of the irradiated crystal and the generation of a particle cloud.

Acceleration through an electrical field. An electrical field accelerated the charged ions to separate them from the particle cloud. Following separation, they were allowed to drift through a vacuum towards the detector apparatus.

Detection of ions and TOF measurement. The accelerated ions were detected by the secondary electron multiplier (SEM) apparatus, which measured the TOF for each particle and allowed for subsequent calculation of ionic mass.

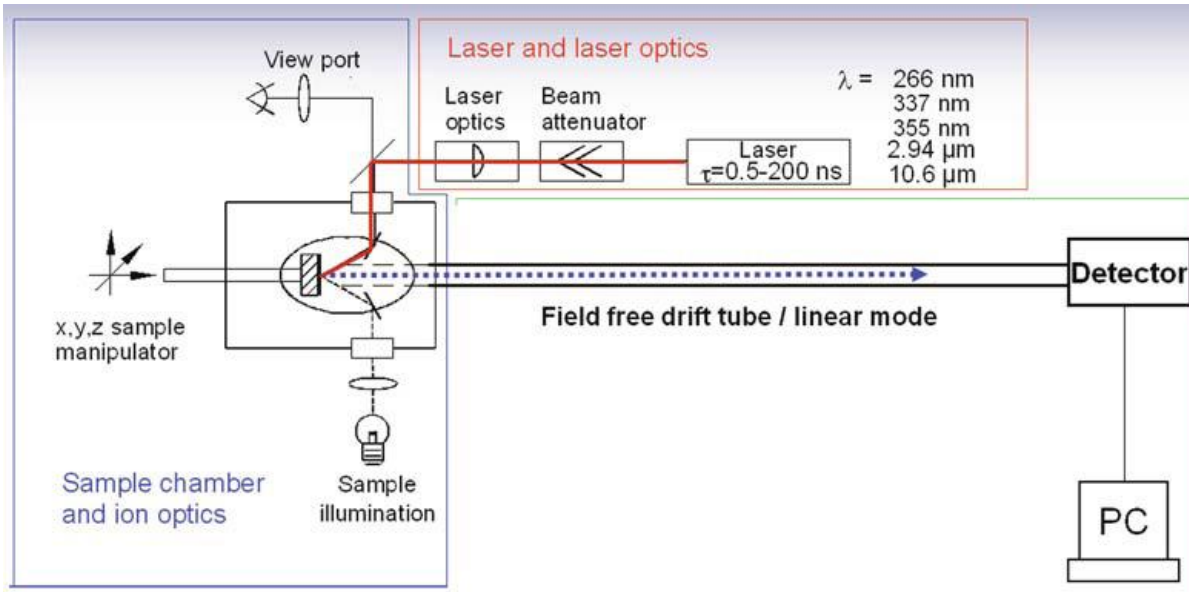


Figure 2: General schematic of a linear MALDI-TOF MS system[178]

2.4.8.1. Reading the SpectroCHIP®

Once the calibrant and samples had been spotted onto the SpectroCHIP®, it was transferred to the MassARRAY® MALDI-TOF MS instrument using handling tongs to avoid contamination. Once inserted, the vacuum was activated and laser-desorption/ionisation process initiated. The SEM collected TOF data for all ions released from each individual sample. The spectrum of detected particles were processed and smoothed with a digitiser (Typer® software program) and the interpreted genotypes assigned on the basis of predetermined values uploaded into a relational database (Typer® database).

2.4.8.2. Genotype Quality Control

The 'MassARRAY® Caller' software package was used to display the genotype data determined by the MALDI-TOF MS process. The following steps were taken to ensure accuracy of data:

Contamination: The plate design incorporated multiple wells that did not contain a DNA sample yet were still analysed for genetic content; any contamination of these blank wells would result in false-positive calls.

Consistency between repeats: Where sufficient volumes of DNA were available, repeats of samples were used to ensure consistency of genotype calls. Twenty-two of the seventy-seven available DNA samples (29%) were repeated across either plex 1 or plex 2, the calls of which were compared, and any anomalies corrected / excluded accordingly.

Manual calls: Where the software was unable to assign a genotype, it was labelled as a 'no call'. The spectra for these were reviewed, and genotypes manually assigned, where possible.

2.5. Data Assembly

The GWAS dataset consisted of genotype data for 166 SNPs across 96 patients. A total of 104 participants were originally selected for inclusion, however the GWAS-genotyping process was deemed to have failed for 8 of them, decreasing the dataset to 96 patients overall.

A further 39 tSNPS were selected for genotyping to capture the outstanding genetic variation. This was undertaken on only 77 of the identified 104 eligible participants due to DNA stock depletions.

All of the available genotype data was assembled into a master database, representing 100 DNA samples and up to 205 SNP genotypes per sample.

2.5.1. Data Quality Control

Before any statistical analyses could be carried out on the data, the following quality control procedures were performed:

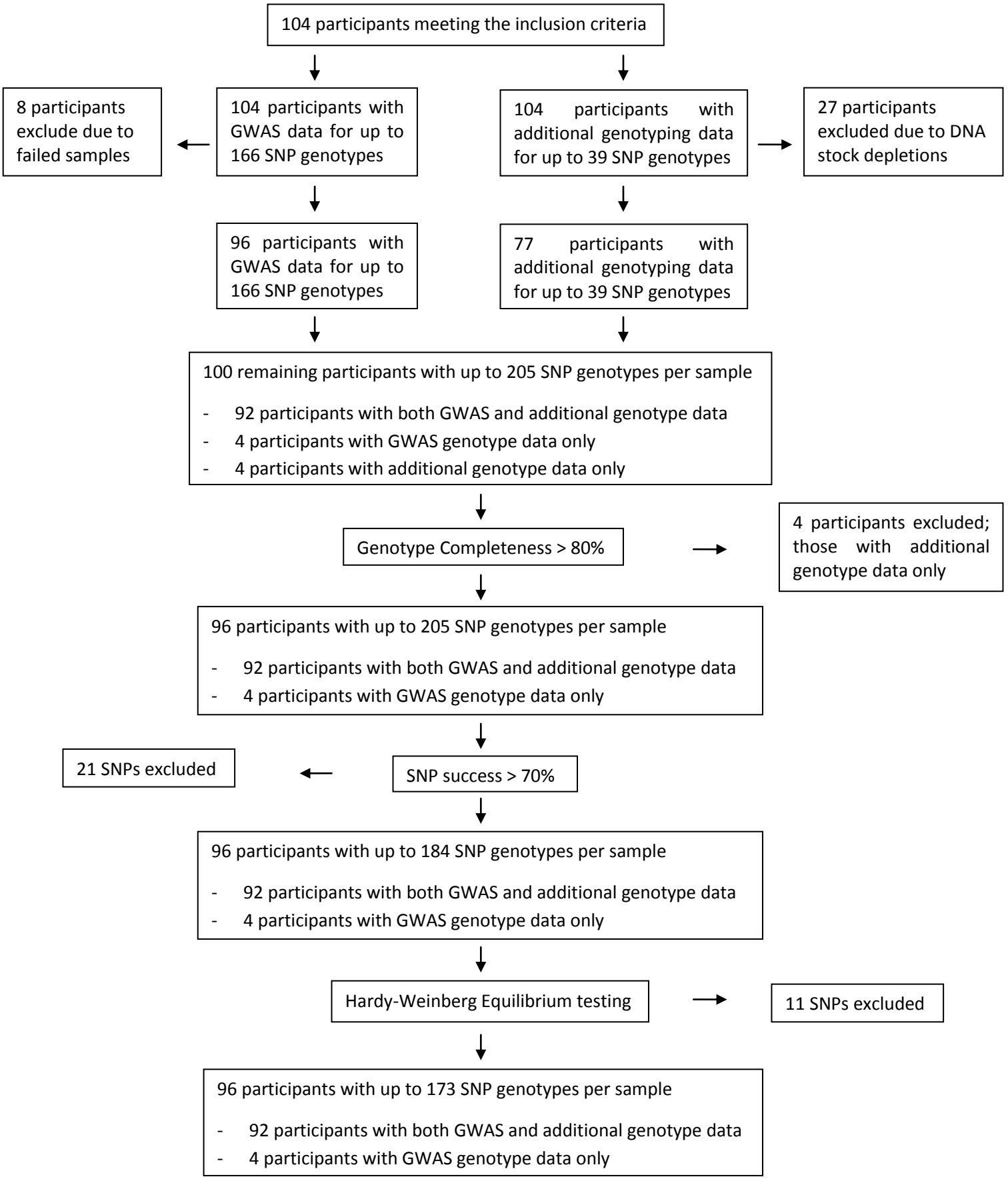
Genotyping Completeness: For each DNA sample, at least 80% of all SNPs must have been successfully genotyped for that DNA sample to be included in that analysis. This resulted in exclusion of 4 DNA samples, reducing the total number of samples to 96.

SNP Success: For each SNP, a genotype must have been established in at least 70% of the DNA samples for that SNP to be included the analysis. This resulted in exclusion of 21 SNPs, reducing the total number of SNPs to 184.

Hardy-Weinberg Equilibrium (HWE): SNPs that failed to meet HWE (p-value = 0.001) were excluded from further analysis. This resulted in exclusion of a further 11 SNPs, reducing the total number of SNPs to 173.

Overall, following data quality control measures, 96 DNA samples with genotype data for 173 SNPs were put forward for statistical analyses.

2.5.2. Data Assembly Overview:



2.6. Statistics

SPSS (version 17) was used to perform various analyses on the collated data.

2.6.1. Data Coding

Following collection, data were coded according to the format in *Table 2.13*.

Table 2.13: Summary of data coding prior to analysis

Variable	Coding References
Maintenance Dose	Daily dose in mg
Epilepsy Type	(0, UNC) (1, LRE) (2, IGE)
Sex	(0, Male) (1, Female)
Number of Seizures before Randomisation	Seizure number
Age at which Maintenance Dose Initiated	Age in years
Previous AED use?	(0, Yes) (1, No)
Genotype	(0, homozygous major allele) (1, heterozygous) (2, homozygous minor allele)

IGE = idiopathic generalised epilepsy; LRE = localisation-related epilepsy; UNC = unclassifiable epilepsy; AED = antiepileptic drug

2.6.2. Descriptive Analyses

Descriptive analyses of each variable were performed. A mean, standard deviation, and range were calculated for normally-distributed continuous variables, and a median and range for skewed continuous variables. The distribution of participants among groups was compared.

2.6.3. Univariate Analyses

Following descriptive analyses, univariate analyses were undertaken for each variable in relation to maintenance dose according to the tests outlined in *Table 2.14*. Since the data for maintenance dose was found to have a skewed distribution, the natural log of maintenance dose (Ln Dose) was modelled as the dependent variable. Following completion of the univariate analyses, correction for multiple comparisons using false-discovery rate was performed.

Table 2.14: Statistical tests performed on each variable

Variable	Statistics Test
Epilepsy Type	One-Way Analysis of Variance (ANOVA)
Sex	Student's t-test
Number of Seizures before Randomisation	Linear regression
Age at which Maintenance Dose Initiated	Linear regression
Previous AED use?	Student's t-test
Genotype	One-Way ANOVA

2.6.4. Multivariate Analyses

Multivariate analysis was carried out using a stepwise linear regression model to identify any association with maintenance dose. Seven separate multivariate analyses were performed to include SNPs from the following gene(s); for each analysis, non-genetic factors were included as covariates and Ln dose was modelled as the dependent variable:

- *SLC22A1* alone
- *UGT1A4* alone
- *SCN1A* alone
- *SCN2A* alone
- *SCN3A* alone
- *SCN1A*, *SCN2A*, *SCN3A* combined
- *SLC22A1*, *UGT1A4*, *SCN1A*, *SCN2A*, *SCN3A* combined

Participants with missing data were excluded in a list-wise fashion to avoid skewing the model. The impact of each variable on the model was indicated by the β-coefficient, with the corresponding p-value denoting the probability that the factor influenced LTG maintenance dose. The r^2 value was used to indicate the goodness of fit of the overall model, and represented the degree of variation in LTG dosing that could be accounted for by that model. The p-residual value indicated the significance of the model overall; a threshold of 0.05 was set to identify statistical significance.

Following successful generation of a multivariate model, a dose-predictive equation for determining the maintenance dose of LTG was constructed using the generated correlation coefficients. This equation was then used to generate predicted doses for each participant used in the generation of that particular model.

Finally, the data was back-transformed from Ln-dose to maintenance dose, and the relationship between observed and predicted maintenance dose determined using Spearman's correlation test on the non-parametric data.

Chapter 3: Results

3.1. Descriptive Statistics

Ninety-six participants were included in the analysis; 48 male and 48 female. The age of participants ranged from 9 – 83 years, with a mean age of 43 years (standard deviation 18.8 years). A total of 61.5% of participants were classified as having LRE, 11.5% with IGE, and the remaining 27.0% had an unclassifiable epilepsy type (*Table 3.1*).

The median number of seizures before randomisation was 10 and ranged from 2 to approximately 1800. For the majority of participants (61.5%), LTG was the first AED.

Table 3.1: Patient demographics

		Patient Cohort (n=96)
Sex	Male	50%
	Female	50%
Age in years	Mean	43
	Minimum	9
	Maximum	83
	Standard Deviation	18.8
Epilepsy type	IGE	11.5%
	LRE	61.5%
	UNC	27%
Number of seizures	Median	10
	Minimum	2
	Maximum	1801
Maintenance dose in mg/day	Median	200
	Minimum	50
	Maximum	675
LTG first AED?	Yes	61.5%
	No	38.5%

IGE = idiopathic generalised epilepsy; LRE = localisation-related epilepsy; UNC = unclassifiable epilepsy; LTG = lamotrigine, AED = antiepileptic drug

The maintenance dose ranged from 50-675mg/day, with a median dose of 200mg/day. Prior to statistical analyses, the distribution of the dose variable was explored and found to be positively skewed (*Figure 3*). The data was therefore transformed by taking the natural log, bringing the distribution closer to that of the normal to allow parametric tests to be carried out.

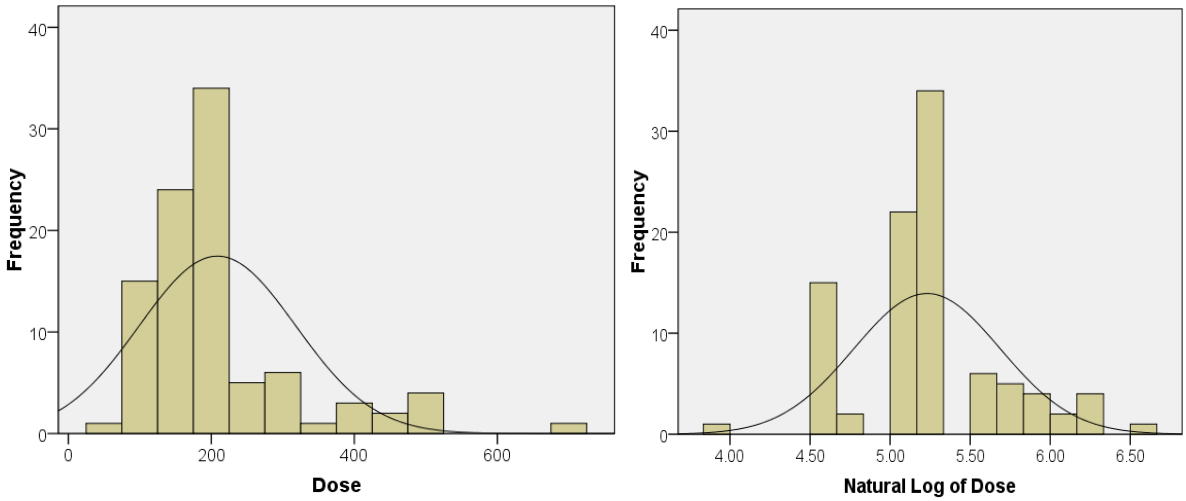


Figure 3: Distribution of LTG maintenance dose before and after log transformation

3.2. Univariate Analyses

Correction for multiple comparisons using false-discovery rate (FDR) was performed, taking into account 178 individual tests (173 genetic, 5 non-genetic); the corrected p-values are displayed in the final column of the relevant tables.

3.2.1. Non-genetic factors

3.2.1.1. Linear Regression

A linear regression model was used to identify associations between maintenance dose and the following factors:

- Number of seizures before randomization (seizure number)
- Age at which maintenance dose was initiated (age)

Ln Dose was modelled as the outcome variable; no significant associations were found between dose and either seizure number or age (*Table 3.2*).

Table 3.2: Univariate Analyses - linear regression

Predictor	Coefficient	r^2 value	p-value	p-value rank	FDR-corrected p-value
Seizure number	0.100	0.010	0.335	72	0.828
Age	-0.038	0.001	0.712	132	0.960

3.2.1.2. Student's t-test

Student's t-test was used to identify associations between Ln Dose and the following independent variables:

- Sex
- Previous AED use

No statistically significant associations were identified (*Table 3.3*).

Table 3.3: Univariate analyses: Student's t-test

Variable	t-statistic	Degrees of freedom	p-value	Confidence interval		p-value rank	FDR-corrected p-value
				Lower	Upper		
Sex	-0.706	94	0.482	-0.25242	0.11993	91	0.943
Previous AED Use	-0.227	94	0.821	-0.21369	0.16976	153	0.955

3.2.1.3. One-way Analysis of Variance

ANOVA was performed to investigate any association between Ln Dose and epilepsy type; none was identified (*Table 3.4*).

Table 3.4: Univariate Analyses: One-way analysis of variance

Variable	F-Statistic	Degrees of Freedom		p-value	p-value rank	FDR-corrected p-value
		Between Groups	Within Groups			
Epilepsy Type	2.305	2	93	0.105	31	0.603

3.2.2. Genetic Factors

ANOVA was performed to investigate associations between Ln Dose and the three possible genotypes at each SNP locus; in situations where there were only two genotypes, Student's t-test was performed. Overall, a total of 147 one-way ANOVAs and 26 Student's t-tests were performed on genetic factors; these data are displayed in Appendix VI. Results with an uncorrected p-value <0.05 are illustrated in *Table 3.5* and *Table 3.6*; those with p<0.02 are explored in more detail thereafter, and the remaining results with a pre-corrected p-value between 0.02 and 0.05 have been included in Appendix VII.

Table 3.5: Summary of one-way ANOVA tests p<0.05 (uncorrected)

SNP (Minor allele)	Gene-Region	F-Stat	Degrees of Freedom		p-value	p-value rank	FDR-corrected p-value
			Between Groups	Within Groups			
rs3798173 (A)	SLC22A1	5.102	2	92	0.008	1	1.000
rs628031 (A)	SLC22A1	4.567	2	93	0.013	4	0.579
rs456598 (A)	SLC22A1	4.264	2	93	0.017	5	0.605
rs806383 (A)	SLC22A1	3.393	2	93	0.038	12	0.564
rs12993173 (C)	SCN2A	4.256	2	92	0.017	5	0.605
rs1439805 (C)	SCN2A	4.184	2	93	0.018	7	0.458
rs9287856 (A)	SCN2A	3.724	2	93	0.028	8	0.623
rs9287857 (G)	SCN2A	3.724	2	93	0.028	8	0.623
rs1866603 (G)	SCN2A	3.629	2	93	0.030	10	0.534
rs17184707 (A)	SCN2A	3.384	2	93	0.039	13	0.534
rs10174400 (A)	SCN2A	3.108	2	91	0.049	14	0.623
rs11903851 (A)	SCN3A	5.018	2	93	0.009	2	0.801

SNP = single-nucleotide polymorphism; F-stat = F-statistic; FDR = False Discovery Rate

Table 3.6: Student's t-test p<0.05 (uncorrected)

SNP (Minor allele)	Gene Region	t-stat	DF	p-value	Mean Diff	95% CI		p-value rank	FDR-corrected p-value
						Lower	Upper		
rs461473 (A)	SLC22A1	-2.575	67	0.012	-0.368	-0.654	-0.083	3	0.712
rs34059508(A)	SLC22A1	-2.151	69	0.035	-0.557	-1.074	-0.040	11	0.566

SNP = single-nucleotide polymorphism; t-stat = t-statistic; DF = degrees of freedom; mean diff = mean difference; CI = confidence interval; FDR = False Discovery Rate.

3.2.2.1. *SLC22A1*

rs3798173

ANOVA identified a significant association between maintenance dose of LTG and genotype at this SNP locus (p=0.008, uncorrected). This intronic SNP in *SLC22A1* tagged an additional seven SNPs (*Appendix Table 3*). The homozygous-minor allele group (AA) appeared to require a higher maintenance dose of LTG than either the heterozygous (AG) or homozygous-major allele group (GG) (*Table 3.7* and *Table 3.8*) although there was no evidence of an allelic-specific effect on dose (*Figure 4*).

Table 3.7: Back-transformed dose data for rs3798173

Genotype	Number of Participants	Mean Dose (mg/day)	Minimum Dose (mg/day)	Maximum Dose (mg/day)
GG	60	190	100	675
GA	32	165	50	300
AA	3	368	200	500
Total	95	185	50	675

Table 3.8: Back-transformed dose data for rs3798173 - between groups comparison

Compared Genotypes		Ratio of Geometric Means	95% Confidence Interval for Ratio of Geometric Means		Significance
			Lower Boundary	Upper Boundary	
GG	GA	1.153	0.917	1.450	0.350
	AA	0.516*	0.278	0.958	0.032
GA	GG	0.867	0.690	1.091	0.350
	AA	0.447*	0.238	0.842	0.008
AA	GG	1.939*	1.044	3.601	0.032
	GA	2.235*	1.188	4.206	0.008

* The mean difference is significant at the 0.05 level.
Number of participants: GG = 60, GA = 32, AA = 3, Total = 95

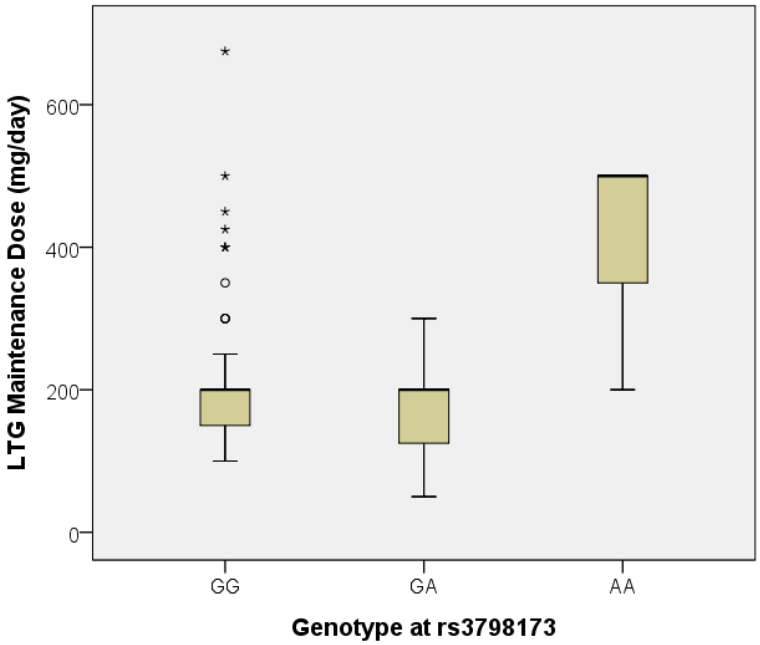


Figure 4: Dose distribution between rs3798173 genotype
Top box boundary = third quartile, lower boundary = first quartile, and bold line = median. Whiskers = values within 1.5 x inter-quartile range (IQR) outside the box. Circles = values 1.5 to 3 x IQR outside the box. Stars = values more than 3 x IQR outside the box.
Number of participants: GG = 60, GA = 32, AA = 3, Total = 95

rs628031

ANOVA identified a significant association between maintenance dose of LTG and genotype at this SNP locus (p=0.013, uncorrected). This is a coding non-synonymous SNP in *SLC22A1* that results in a M408V amino acid substitution in the OCT1 protein. The homozygous-major allele group (GG) appeared to require a higher maintenance dose of LTG than the heterozygous group (GA) and the homozygous-minor allele group (AA) (*Table 3.9* and *Table 3.10*). There was some evidence of an

allele-specific effect on dose, with carriers of the A-allele requiring lower maintenance doses of LTG (*Figure 5*).

Table 3.9: Back-transformed dose data for rs628031

Genotype	Number of Participants	Mean Dose (mg/day)	Minimum Dose (mg/day)	Maximum Dose (mg/day)
GG	33	224	100	675
GA	47	175	50	500
AA	16	157	100	200
Total	96	187	50	675

Table 3.10: Back-transformed dose data for rs628031 - between groups comparison

Compared Genotypes		Ratio of Geometric Means	95% Confidence Interval for Ratio of Geometric Means		Significance
			Lower Boundary	Upper Boundary	
GG	GA	1.280*	1.003	1.633	0.047
	AA	1.430*	1.031	1.984	0.027
GA	GG	0.781*	0.612	0.997	0.047
	AA	1.118	0.819	1.525	0.767
AA	GG	0.699*	0.504	0.970	0.027
	GA	0.895	0.656	1.221	0.767

* The mean difference is significant at the 0.05 level.
Number of participants: GG = 33, GA = 47, AA = 16, Total = 96

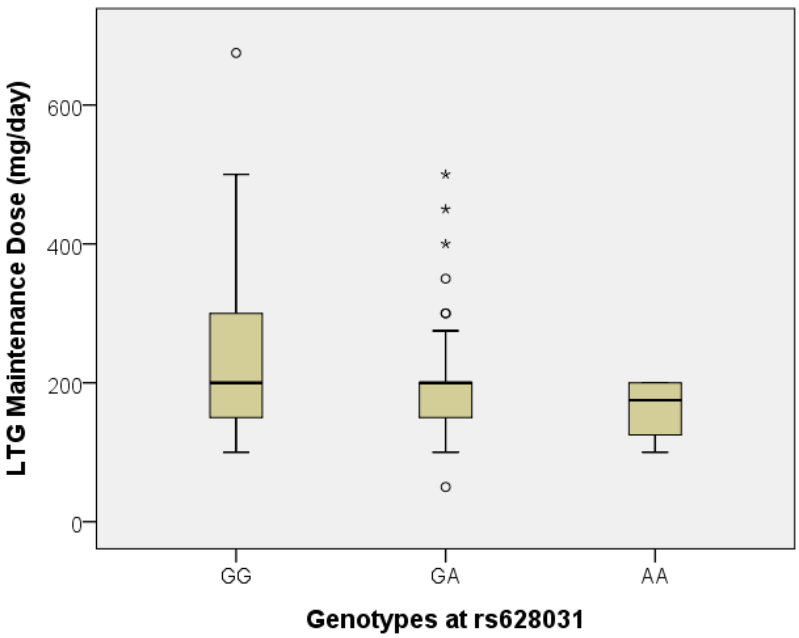


Figure 5: Dose distribution between rs628031 genotypes
Top box boundary = third quartile, lower boundary = first quartile, and bold line = median. Whiskers = values within 1.5 x inter-quartile range (IQR) outside the box. Circles = values 1.5 to 3 x IQR outside the box. Stars = values more than 3 x IQR outside the box.
Number of participants: GG = 33, GA = 47, AA = 16, Total = 96

rs456598

ANOVA identified a significant association between maintenance dose of LTG and genotype at this SNP locus (p=0.017, uncorrected). This SNP in the flanking region of *SLC22A1* did not tag any additional SNPs. Upon scrutiny, the association was found to be based on the genotype of a single participant carrying the heterozygous-minor allele genotype (AA). Post-hoc tests were not performed as a consequence and no efforts were made to explore allele-specific effects on dose (*Table 3.11*).

Table 3.11: Back-transformed dose data for rs456598

Genotype	Number of Participants	Mean Dose (mg/day)	Minimum Dose (mg/day)	Maximum Dose (mg/day)
GG	70	185	50	498
GA	25	181	100	399
AA	1	675	675	675
Total	96	187	50	675

rs461473

Student's t-test identified a significant association between maintenance dose and genotype at this SNP locus (p=0.012, uncorrected). This intronic SNP is found in *SLC22A1* and did not tag any additional SNPs. The heterozygous group (GA, geometric mean = 246mg/day, n=11) appeared to require a higher maintenance dose of LTG than the homozygous-minor allele group (GG, geometric mean = 170mg/day, n=58) (*Table 3.12*). With only two genotypes at this locus, it was not possible to infer allele-specific effects.

Table 3.12: Back-transformed dose data for rs461473 – between groups comparison

Variable	t-statistic	Degrees of freedom	Ratio of Geometric means	95% CI of Ratio of Geometric Means		p-value	p-value rank	FDR-corrected p-value
				Lower	Upper			
rs461473 (A)	-2.575	67	0.692	0.520	0.921	0.012	3	0.712

Number of participants: GG = 58, GA = 11, Total = 69

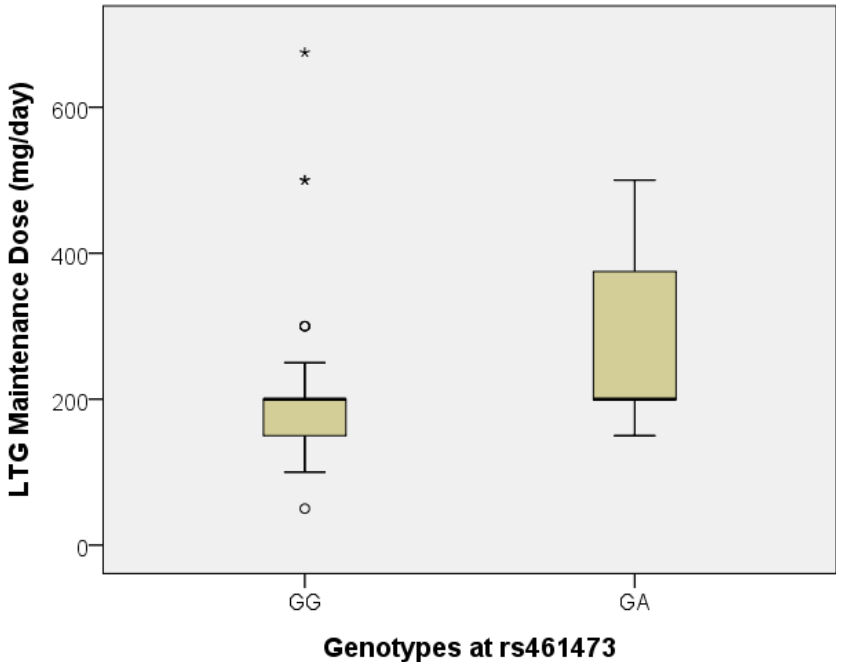


Figure 6: Dose distribution between rs461473 genotypes
Top box boundary = third quartile, lower boundary = first quartile, and bold line = median. Whiskers = values within 1.5 x inter-quartile range (IQR) outside the box. Circles = values 1.5 to 3 x IQR outside the box. Stars = values more than 3 x IQR outside the box.
Number of participants: GG = 58, GA = 11, Total = 69

3.2.2.2. *SCN2A*

rs12993173

ANOVA identified a significant association between maintenance dose of LTG and genotype at this SNP locus (p=0.017, uncorrected). This intronic SNP is found in *SCN2A* and tagged an additional seven SNPs (*Appendix Table 6*). The homozygous-major allele group (AA) appeared to require a higher maintenance dose of LTG than the heterozygous group (AC) (*Table 3.13* and

Table 3.14). Comparison with the homozygous-minor allele group (CC) failed to reach statistical significance and there no was no evidence of an allele-specific effect on dose (*Figure 7*).

Table 3.13: Back-transformed dose data for rs12993173

Genotype	Number of Participants	Mean Dose (mg/day)	Minimum Dose (mg/day)	Maximum Dose (mg/day)
AA	26	229	100	675
AC	50	171	100	500
CC	19	171	50	500
Total	95	185	50	675

Table 3.14: Back-transformed dose data for rs12993173 - between groups comparison

Compared Genotypes		Ratio of Geometric Means	95% Confidence Interval for Ratio of Geometric Means		Significance
			Lower Boundary	Lower Boundary	
AA	AC	1.339*	1.038	1.729	0.020
	CC	1.337	0.972	1.838	0.085
AC	AA	0.747*	0.579	0.964	0.020
	CC	0.998	0.751	1.327	1.000
CC	AA	0.748	0.544	1.029	0.085
	AC	1.002	0.754	1.331	1.000

* The mean difference is significant at the 0.05 level.
Number of participants: AA = 26, AC = 50, CC = 19, Total = 95

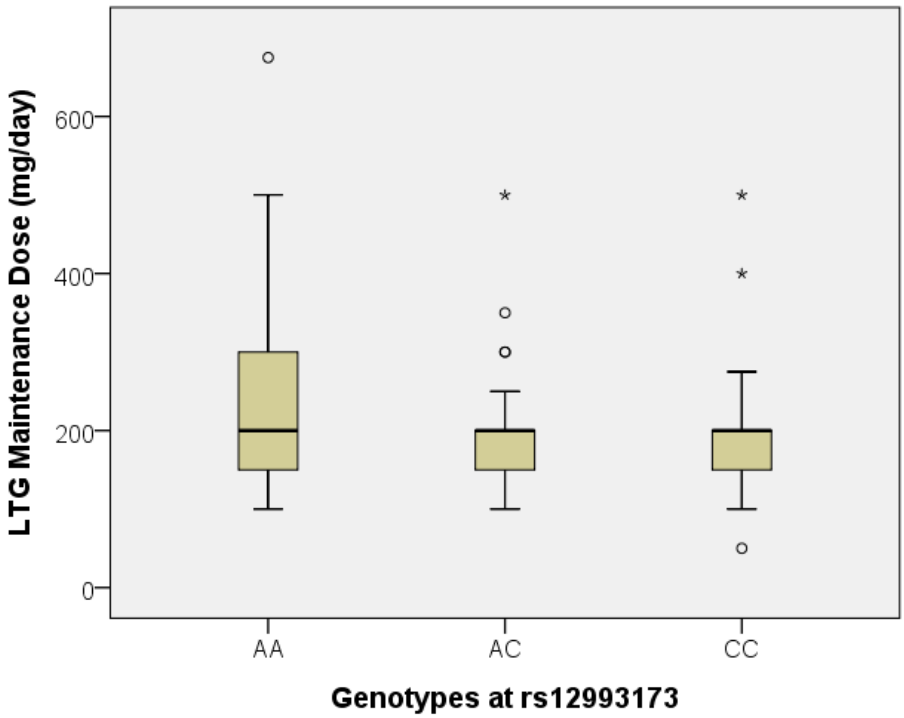


Figure 7: Dose distribution between rs12993173 genotypes
Top box boundary = third quartile, lower boundary = first quartile, and bold line = median. Whiskers = values within 1.5 x inter-quartile range (IQR) outside the box. Circles = values 1.5 to 3 x IQR outside the box. Stars = values more than 3 x IQR outside the box.
Number of participants: AA = 26, AC = 50, CC = 19, Total = 95

rs1439805

ANOVA identified a significant association between maintenance dose of LTG and genotype at this SNP locus (p=0.018, uncorrected). This intronic SNP is found in *SCN2A* and tagged an additional four SNPs (*Appendix Table 6*). The homozygous-major allele group (AA) appeared to require a higher maintenance dose of LTG than the heterozygous group (AC) (*Table 3.15* and *Table 3.17*). Comparison

with the homozygous-minor allele group (CC) failed to reach statistical significance and there no was no evidence of an allele-specific effect on dose (*Figure 8*).

Table 3.15: Back-transformed dose data for rs1439805

Genotype	Number of Participants	Mean Dose (mg/day)	Minimum Dose (mg/day)	Maximum Dose (mg/day)
AA	24	230	100	500
AC	53	168	50	675
CC	19	192	100	500
Total	96	187	50	675

Table 3.16: Back-transformed dose data for rs1439805 - between groups comparison

Compared Genotypes		Ratio of Geometric Means	95% Confidence Interval for Ratio of Geometric Means		Significance
			Lower Boundary	Upper Boundary	
AA	AC	1.369*	1.030	1.820	0.028
	CC	1.200	0.858	1.677	0.392
AC	AA	0.731*	0.550	0.971	0.028
	CC	0.876	0.668	1.151	0.469
CC	AA	0.834	0.596	1.165	0.392
	AC	1.141	0.869	1.498	0.469

* The mean difference is significant at the 0.05 level.
Number of participants: AA = 24, AC = 53, CC = 19, Total = 96

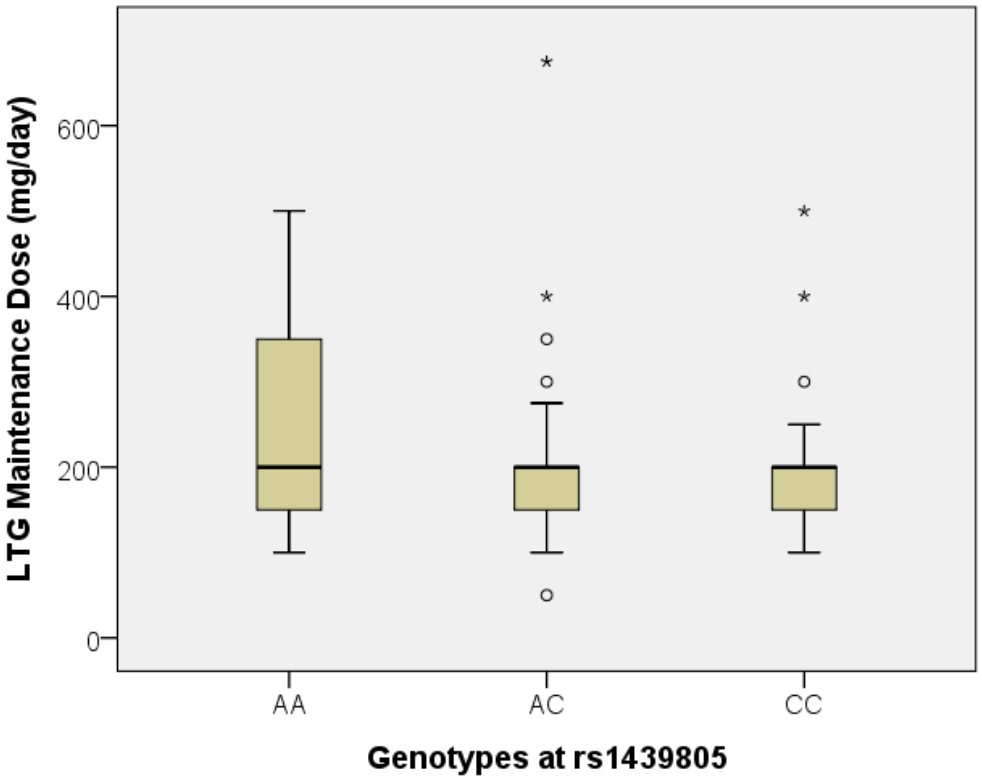


Figure 8: Dose distribution between rs1439805 genotypes
Top box boundary = third quartile, lower boundary = first quartile, and bold line = median. Whiskers = values within 1.5 x inter-quartile range (IQR) outside the box. Circles = values 1.5 to 3 x IQR outside the box. Stars = values more than 3 x IQR outside the box.
Number of participants: AA = 24, AC = 53, CC = 19, Total = 96

3.2.2.3. *SCN3A*

rs11903851

ANOVA identified a significant association between maintenance dose of LTG and genotype at this SNP locus (p=0.009, uncorrected). This intronic SNP in *SCN3A* tagged an additional fifteen SNPs (*Appendix Table 7*). The heterozygous group (GA) appeared to require a higher maintenance dose of LTG than the homozygous-major allele group (GG) (*Table 3.9* and *Table 3.18*). As such, the data does not support an allele-specific effect on dose (*Figure 9*).

Table 3.17: Back-transformed dose data for rs11903851

Genotype	Number of Participants	Mean Dose (mg/day)	Minimum Dose (mg/day)	Maximum Dose (mg/day)
GG	63	171	50	500
GA	25	238	150	675
AA	8	174	100	500
Total	96	187	50	675

Table 3.18: Back-transformed dose data for rs11903851 - between groups comparison

Compared Genotypes		Ratio of Geometric Means	95% Confidence Interval for Ratio of Geometric Means		Significance
			Lower Boundary	Upper Boundary	
GG	GA	0.722*	0.561	0.930	0.007
	AA	-0.985	0.659	1.471	1.000
GA	GG	1.385*	1.076	1.783	0.007
	AA	1.364	0.884	2.106	0.234
AA	GG	1.015	0.680	1.516	1.000
	GA	0.733	0.475	1.132	0.234

* The mean difference is significant at the 0.05 level.
Number of participants: AA = 63, AC = 25, CC = 8, Total = 96

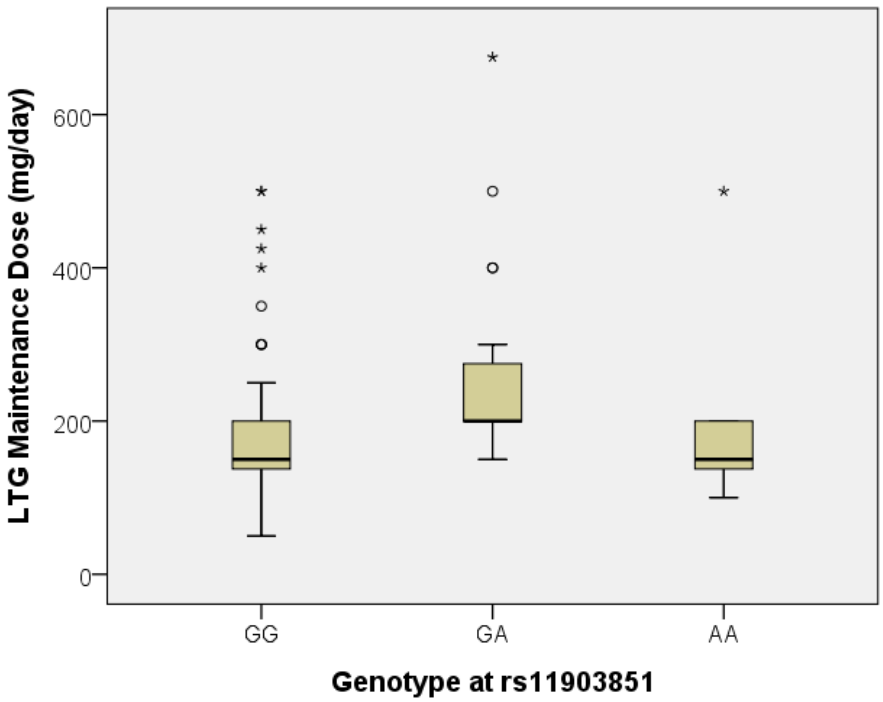


Figure 9: Dose distribution between rs11903851 genotypes
Top box boundary = third quartile, lower boundary = first quartile, and bold line = median. Whiskers = values within 1.5 x inter-quartile range (IQR) outside the box. Circles = values 1.5 to 3 x IQR outside the box. Stars = values more than 3 x IQR outside the box.
Number of participants: AA = 63, AC = 25, CC = 8, Total = 96

3.2.3. Multivariate Analysis

Multivariate analysis was carried out using a stepwise linear regression model with maintenance dose modelled as the dependent variable and the five identified non-genetic variables as covariates. Seven separate analyses were undertaken incorporating SNPs within the following gene groupings:

- *SLC22A1* alone
- *UGT1A4* alone
- *SCN1A* alone
- *SCN2A* alone
- *SCN3A* alone
- *SCN1A, SCN2A* and *SCN3A* combined
- *SCN1A, SCN2A, SCN3A, UGT1A4, SLC22A1* combined

Following successful generation of a multivariate model, a dose-predictive equation for determining the maintenance dose of LTG was constructed using the generated correlation coefficients. This equation was then used to generate predicted doses for each participant used in the generation of that particular model. Finally, the data was back-transformed from Ln-dose to maintenance dose, and the relationship between observed and predicted maintenance dose determined using Spearman's correlation test.

3.2.3.1. *SLC22A1*

Four models were generated following stepwise multivariate regression analysis of the *SLC22A1* gene, all of which had a p-residual value below the pre-determined threshold of 0.05 for a statistically significant result to be inferred (

Table 3.19). The most explanatory model (model 4) associated with 31.0% of the variation in maintenance dose, and was used to generate the following dose-predictive equation:

$LnOD_{LTG} = 5.133 + (0.265 \times SNP\ 1) + (1.258 \times SNP\ 2) - (0.554 \times SNP\ 3) - (0.186 \times SNP\ 4)$

($LnOD_{LTG}$ = Natural logarithm of daily lamotrigine maintenance dose (mg); SNP 1-4 = genotype at SNPs 1-4 scored as homozygous-major allele = 0, heterozygous = 1, homozygous minor allele = 2)

Table 3.19: Multiple linear regression models for SLC22A1

Variables (SNP number in predictive equation)	Model 1		Model 2		Model 3		Model 4	
	β	**p**	**β**	**p**	**β**	**p**	**β**	**p**
Constant	5.060	<0.001	5.018	<0.001	5.054	<0.001	5.133	<0.001
rs806383 (SNP 1)	0.212	0.013	0.232	0.005	0.215	0.007	0.265	0.001
rs34059508 (SNP 2)			0.644	0.010	1.134	0.001	1.258	<0.001
rs6937722 (SNP 3)					-0.520	0.032	-0.554	0.019
rs10455868 (SNP 4)							-0.186	0.033
Model (r^2, p-residual)	0.098, **0.013**		0.192, **0.002**		0.253, **0.001**		0.310, **<0.001**	

β = β coefficient
p = p-value
Models based on 63 participants

Using the dose-predictive equation generated from model 4, predicted LTG maintenance doses were calculated and then correlated with the observed maintenance doses. The data was back-transformed and the degree of correlation between the two variables determined using Spearman's correlation test (correlation coefficient 0.403) which was statistically significant with a corresponding p-value < 0.001.

3.2.3.2. *UGT1A4*

A multiple linear regression model failed to identify any significant associations using SNP genotypes from *UGT1A4.* None of the variables were able to be incorporated into a model with a p-residual value of < 0.05.

3.2.3.3. *SCN1A*

A multiple linear regression model failed to identify any significant associations using SNP genotypes from *SCN1A.* None of the variables were able to be incorporated into a model with a p-residual value of < 0.05.

3.2.3.4. *SCN2A*

A singular model was generated from the *SCN2A* gene, with 24.2% of the variation in LTG maintenance dose explained by genotype at rs2119067 alone. This was found to be statistically significant with a p-residual value of 0.01 (*Table 3.20*), and the following dose-predictive equation was produced using the information:

$LnOD_{LTG}$ = 5.298 - (0.242 x SNP 1)

($LnOD_{LTG}$ = Natural logarithm of daily lamotrigine maintenance dose (mg); SNP 1 = rs2119067; genotypes scored as homozygous-major allele = 0, heterozygous = 1, homozygous minor allele = 2)

Table 3.20: Multiple linear regression model for SCN2A

Variables (SNP number in predictive equation)	**Model 1**	
	B-coefficient	**p-value**
Constant	5.298	<0.001
rs2119067 (SNP 1)	-0.242	0.010
Model (r^2, p-residual)	0.094, **0.010**	

Model based on 69 participants

Using the generated equation, predicted LTG maintenance doses were calculated and then correlated with the observed maintenance doses. The data was back-transformed and the degree of correlation between the two variables determined using Spearman's correlation test (correlation coefficient 0.219) which was statistically significant with a corresponding p-value of 0.032.

3.2.3.5. *SCN3A*

Three models were generated following stepwise multivariate regression analysis of the *SCN3A* gene, all of which had a p-residual value below the pre-determined threshold of 0.05 (*Table 3.21*). The most explanatory model (model 3, r^2 = 23.5%) was used to generate the following dose-predictive equation:

$LnOD_{LTG}$ = 5.046 + (0.918 x SNP 1) – (0.692 x SNP 2) + (0.197 x SNP 3)

($LnOD_{LTG}$ = Natural logarithm of daily lamotrigine maintenance dose (mg); SNP 1-3 = genotype at SNPs 1-3 scored as homozygous-major allele = 0, heterozygous = 1, homozygous minor allele = 2)

The data was back-transformed and Spearman's correlation test was used to determine the degree of correlation between observed and predicted maintenance doses (correlation coefficient 0.246). This was found to be statistically significant with a corresponding p-value of 0.016.

Table 3.21: Multiple linear regression models for SCN3A

Variables (SNP number in predictive equation)	**Model 1**		**Model 2**		**Model 3**	
	B-coefficient	**p-value**	**B-coefficient**	**p-value**	**B-coefficient**	**p-value**
Constant	5.090	<0.001	5.102	<0.001	5.046	<0.001
rs11903851 (SNP 1)	0.242	0.005	0.912	0.002	0.918	0.002
rs1982208 (SNP 2)			-0.698	0.017	-0.692	0.016
rs17829596 (SNP 3)					0.197	0.040
Model (r^2, p-residual)	0.111, **0.005**		0.184, **0.001**		0.235, **<0.001**	

Models based on 70 participants

3.2.3.6. *SCN1A*, *SCN2A* and *SCN3A*

The *SCN1A*, *SCN2A* and *SCN3A* genes were combined into a multiple linear regression model. Six separate models were produced, with r^2 values ranging from 0.109 to 0.450 and significant p-residual values < 0.05 (*Table 3.22*). The most explanatory model (model 6, r^2 = 45.0%) was used to generate the following dose-predictive equation:

$LnOD_{LTG}$ = 5.318 + (1.066 x SNP 1) – (0.361 x SNP 2) – (0.231 x SNP 3) – (0.544 x SNP 4) + (0.355 x SNP 5) – (0.379 x SNP 6)

($LnOD_{LTG}$ = Natural logarithm of daily lamotrigine maintenance dose (mg); SNP 1-6 = genotype at SNPs 1-6 scored as homozygous-major allele = 0, heterozygous = 1, homozygous minor allele = 2)

Spearman's correlation test was used to determine the degree of correlation between observed and predicted maintenance doses (correlation coefficient 0.626) and was found to be statistically significant with a corresponding p-value of <0.001.

Table 3.22: Multiple linear regression models for SCN1A, SCN2A, and SCN3A

Variables (SNP number in predictive equation)	**Model 1**		**Model 2**		**Model 3**		**Model 4**		**Model 5**		**Model 6**	
	β	**p**	**β**	**p**	**β**	**p**	**β**	**p**	**β**	**p**	**β**	**p**
Constant	5.091	<0.001	5.207	<0.001	5.334	<0.001	5.330	<0.001	5.318	<0.001	5.318	<0.001
rs11903851 (SNP 1)	0.245	0.006	0.286	0.001	0.431	<0.001	1.000	<0.001	1.022	<0.001	1.066	<0.001
rs2119067 (SNP 2)			-0.295	0.001	-0.355	<0.001	-0.345	<0.001	-0.344	<0.001	-0.361	<0.001
rs11677254 (SNP 3)					-0.198	0.013	-0.179	0.020	-0.228	0.004	-0.231	0.003
rs1899013 (SNP 4)							-0.609	0.024	-0.587	0.025	-0.544	0.034
rs2165208 (SNP 5)									0.318	0.038	0.355	0.019
rs17242693 (SNP 6)											-0.379	0.047
Model (r^2, p-residual)	0.109, **0.006**		0.243, **<0.001**		0.314, **<0.001**		0.369, **<0.001**		0.412, **<0.001**		0.450, **<0.001**	

β = β coefficient
p = p-value
Models based on 67 participants

3.2.3.7. *SLC22A1*, *UGT1A4*, *SCN1A*, *SCN2A* and *SCN3A*

All 178 variables were combined in a global stepwise multivariate regression analysis. Nine separate models were produced, with r^2 values ranging from 0.117 to 0.606 and significant p-residual values < 0.05 (*Table 3.23*). The most explanatory model (model 9, r^2 = 60.6%) was used to generate the following dose-predictive equation:

$LnOD_{LTG}$ = 5.463 + (0.175 x SNP 1) + (1.021 x SNP 2) – (0.679 x SNP 3) – (0.465 x SNP 4) – (0.453 x SNP 5) + (0.269 x SNP 6) – (0.286 x SNP 7) – (0.258 x SNP 8) + (0.339 x SNP 9)

($LnOD_{LTG}$ = Natural logarithm of daily lamotrigine maintenance dose (mg); SNP 1-9 = genotype at SNPs 1-9 scored as homozygous-major allele = 0, heterozygous = 1, homozygous minor allele = 2)

Spearman's correlation test was used to determine the degree of correlation between observed and predicted maintenance doses (correlation coefficient 0.686) and was found to be statistically significant with a corresponding p-value of <0.001.

Table 3.23: Multiple linear regression model for SLC22A1, UGT1A4, SCN1A, SCN2A, and SCN3A.

Variables (SNP numbers in predictive equation)	**Model 1**		**Model 2**		**Model 3**		**Model 4**		**Model 5**		**Model 6**		**Model 7**		**Model 8**		**Model 9**	
	β	**p**	**β**	**p**	**β**	**p**	**β**	**p**	**β**	**p**	**β**	**p**	**β**	**p**	**β**	**p**	**β**	**p**
Constant	5.069	<0.001	5.027	<0.001	5.065	<0.001	5.182	<0.001	5.288	<0.001	5.243	<0.001	5.333	<0.001	5.452	<0.001	5.463	<0.001
rs806383 (SNP 1)	0.231	0.008	0.248	0.003	0.233	0.004	0.215	0.006	0.220	0.003	0.207	0.004	0.180	0.011	0.172	0.012	0.175	0.008
rs34059508 (SNP 2)			0.629	0.013	1.134	0.001	1.134	0.001	1.134	<0.001	1.134	<0.001	1.134	<0.001	1.134	<0.001	1.021	<0.001
rs6937722 (SNP 3)					-0.537	0.027	-0.649	0.007	-0.685	0.003	-0.618	0.006	-0.691	0.002	-0.713	0.001	-0.679	0.001
rs2119067 (SNP 4)							-0.246	0.016	-0.312	0.003	-0.360	0.001	-0.427	<0.001	-0.452	<0.001	-0.465	<0.001
rs1439806 (SNP 5)									-0.214	0.021	-0.267	0.005	-0.293	0.002	-0.355	<0.001	-0.453	<0.001
rs484926 (SNP 6)											0.237	0.026	0.240	0.020	0.259	0.010	0.269	0.005
rs3798174 (SNP 7)													-0.273	0.037	-0.306	0.016	-0.286	0.019
rs4303727 (SNP 8)															-0.220	0.021	-0.258	0.006
rs2165208 (SNP 9)																	0.339	0.023
Model (r2, p-residual)	0.117, **0.008**		0.210, **0.001**		0.277, **<0.001**		0.352, **<0.001**		0.414, **<0.001**		0.468, **<0.001**		0.512, **<0.001**		0.562, **<0.001**		0.606, **<0.001**	

β = β coefficient
p = p-value
Models based on 59 participants

Chapter 4: Discussion

The ultimate aim of AED therapy is to achieve maximal seizure control in the absence of adverse effects. LTG is a commonly used AED with proven efficacy for various seizure types. It has been shown to be better tolerated than many other AEDs, and was recommended by the SANAD trial as an alternative to CBZ, the current first-line AED of choice in the treatment of partial-onset seizures.[1, 2] As with all AEDs, the effective dose of LTG varies widely from patient to patient, and finding the optimum dose can be difficult.[3] Current practice involves titration of LTG over a six- to eight-week period until a pre-determined target dose is reached, with dose adjustments made thereafter according to the extent of efficacy and occurrence of adverse effects. This 'one size fits all' approach to LTG dosing is sub-optimal for many patients. Persons who are particularly sensitive to LTG may require a low daily dose to achieve seizure control; however, without prior knowledge of this sensitivity, the standard titration schedule may excessively dose these patients to the extent that adverse effects necessitate drug withdrawal. The titration process may also fail to satisfy the needs of patients who require particularly high doses to achieve seizure control. In these patients, LTG may be perceived to be ineffective due to persistent seizures, even at the upper end of the supposed therapeutic dose range. Under these circumstances, where the titration rate is too slow and the initial target dose too low, there is a potential for the patient or clinician to lose confidence in the treatment and withdraw it before the optimal dose is reached. For such patients, the ability to predict the required maintenance dose would prove particularly beneficial, achieving optimal dosing where previously treatment would have been deemed to fail. In addition to this, the treatment of all patients with LTG could become more tailored to individual needs, alleviating the trial and error aspect of current dosing regimens through prediction of dose requirement prior to drug initiation.

The observed variability in individual dose requirement of LTG may, at least in part, be the result of underlying variation in genes encoding proteins that are involved in its pharmacokinetic and pharmacodynamic pathways. LTG is primarily metabolised by the 1A4 isoform of the UGT family of enzymes which is encoded by the *UGT1A4* gene.[59] The organic cation transporter OCT1 has been implicated in the transport of LTG into hepatocytes and at the BBB (unpublished data, University of Liverpool) and is encoded by the gene *SLC22A1*. Finally, the primary mechanism of action of LTG is inhibition of neuronal VGSCs in the brain, the α-subunits of which are encoded by three genes; *SCN1A, SCN2A* and *SCN3A*.[48]

Previous AED pharmacogenetic studies have utilised candidate-gene approaches in their quest to explain inter-individual variation in dose requirement, with PHT and CBZ the most-extensively

studied drugs to date. Despite the current lack of translation into clinical practice, there have been several studies that have successfully identified associations between dose requirement and underlying genetic variation during the past 10 years.[179-181] However, there is little or no information on the source of inter-individual variation in dose requirements of newer AEDs, such as LTG. This current study has attempted to address this gap in the knowledge-base through investigation of variation in five candidate genes (*SLC22A1, UGT1A4, SCN1A, SCN2A,* and *SCN3A*) that might explain, at least in part, the observed inter-individual variation in LTG dosing.

Participants were selected retrospectively from the SANAD cohort of newly-diagnosed epilepsy patients. Analysis was undertaken on a total of 96 individuals with maintenance doses of LTG ranging from 50 to 675mg/day. A total of 178 univariate analyses were performed to investigate associations with LTG maintenance dose. Five demographic variables (age, sex, epilepsy type, previous AED use, and number of seizures before treatment) and genotypes at 173 SNP loci were analysed in isolation using various statistical techniques, resulting in 14 independent variables showing associations with dose ($p<0.05$, uncorrected). However, after correction for multiple comparisons using FDR, none of these variables remained significantly associated with LTG dose. Following univariate analyses, various multivariate models were constructed using stepwise linear regression to determine whether combinations of variables could explain the observed variation in maintenance dose. Significant multivariate models were generated within the *SLC22A1* (

Table 3.19), *SCN2A* (*Table 3.20*), and *SCN3A* (*Table 3.21*) genes. In addition, further models were generated through a combined analysis of the *SCN1A*, *SCN2A*, and *SCN3A* genes (*Table 3.22*) and in a global analysis of all variables examined in this study (*Table 3.23*).

SLC22A1

A total of 38 SNP genotypes across *SLC22A1* were available for analysis; univariate analyses identified six SNPs that were significantly associated with LTG maintenance dose before correction for multiple comparisons (*Table 3.5* and *Table 3.6*). Three of these are intronic SNPs (rs461473, rs806383, rs456598) that do not tag any further genetic variation; the mechanisms by which these might impact upon LTG maintenance dose remains unclear. An additional intronic SNP found to be associated with LTG maintenance dose was rs3798173 which tags a further six SNPs (Appendix Table 3), none of which would be expected to have functional significance. The remaining two polymorphisms that achieved pre-correction significance in *SLC22A1* were the missense SNPs rs34059508 and rs628031 that result in G465R and M408V amino acid substitutions, respectively.

The G465R variant of OCT1 (rs34059508) has been shown to result in reduced transport of both [^{3}H]MPP[140] (a known substrate for OCT1) and the anti-diabetic agent metformin[182] during *in vitro* studies. The overall effect of reduced LTG transport on maintenance dose requirement is likely to be dependent upon the site at which OCT1 is most functionally relevant. If OCT1 expression is of greatest functional relevance at the hepatic level, then impaired OCT1 function is likely to result in elevated plasma concentrations of LTG due to decreased cellular uptake and consequently reduced hepatic metabolism. Under these circumstances, the elevation in plasma concentrations might be expected to out-weigh the impairment of LTG transport across the BBB, resulting in elevated brain concentrations of LTG per unit dose and a lower maintenance dose requirement. Alternatively, if OCT1 were more functionally important at the BBB level, then reduced OCT1 function as a result of the G465R polymorphism might be expected to increase LTG dose requirements in order to overcome reduced BBB transport. In this case, impaired transport into the brain is anticipated to out-weigh any increase in serum LTG concentrations arising from decreased transport into hepatocytes and, as a result, maintenance dose requirements would be higher.

In the present study, the patients who were heterozygous for rs34059508 (G465R) appeared to require a higher maintenance dose of LTG than those who were homozygous for the major allele (*Appendix Table 19* and *Figure 11*). This observation would suggest that OCT1 is of greater functional importance at the BBB than in hepatocytes. However, without adequate understanding of the anatomical distribution of OCT1 and its relative expression at different cell membranes, it is difficult

to draw any robust conclusions which would explain why this SNP is associated with an elevated dose requirement of LTG. In addition, the heterozygous group was comprised of just three participants and therefore the accuracy of the mean maintenance dose in those carrying variant genotypes might be considered questionable. Nonetheless, given the reduced function of OCT1 associated with rs34059508 and its pre-correction significance in univariate analysis (p=0.035), it is reasonable to suggest that this SNP influences LTG maintenance dose requirement, even if the mechanisms by which this occurs remains to be clarified.

The M408V variant of OCT1 (rs628031) has previously been shown to possess similar activity wild-type OCT1 in the transport of [^{3}H]MPP *in vitro.*[140] In the present study, patients who were homozygous for the major allele appeared to require a higher maintenance dose of LTG than either heterozygotes or those who were homozygous for the minor allele (*Table 3.10*) and there was some evidence for a gene-dose effect (*Figure 5*). Again, it is difficult to put the observed dosing trend into context given our lack of understanding of LTG transport by OCT1. For rs628031, it appears LTG maintenance dose requirement decreases when the variant allele is expressed. If OCT1 primarily transports LTG into the brain, this might indicate that the M408V variant results in increased activity of OCT1-mediated LTG transport. Conversely, if LTG is preferentially transported into hepatocytes for metabolism by UGT1A4, the observed trend would indicate that the M408V variant results in the decreased transport of LTG. Any change in LTG transport (either gain or loss of function) would, however, contradict previous data which suggests that this variant is without functional effect. The only possible explanation under those circumstances would be a substrate-specific effect of the polymorphism on the transport of LTG but not on that of MPP. This speculative hypothesis requires further detailed investigation in appropriately-designed drug transport assays.

A stepwise linear regression model was constructed using the 38 SNP genotypes in the *SLC22A1* gene in addition to the five demographic variables (*Table 3.1*). Four significant models were generated from these variables, the last of which comprised four SNP genotypes that collectively explained 31.0% of the observed variation in LTG maintenance dose (p<0.001,

Table 3.19). Of the SNPs included in the model, two were intronic (rs806383, rs6937722), one was in the upstream flanking sequence (rs10455868) and the remaining SNP was the putatively functional G465R variant (rs34059508) discussed above. None of these polymorphisms were identified as being in LD with other SNPs, suggesting that they were not surrogates for an unidentified causal variant.

It is somewhat surprising that 31.0% of the total variation in LTG maintenance dose can be accounted for by genetic variation in these four SNPs alone. However, the fact that six separate *SLC22A1* polymorphisms were found to be associated with LTG maintenance dose in the univariate analysis (albeit prior to correction for multiple testing), none of which were in LD with one another, would suggest that OCT1 has a significant role in the pharmacokinetics of LTG and in determining its dose requirements. This proposition is arguably supported by a dose-predictive equation based on the multivariate model, which showed a reasonable correlation between observed and predicted maintenance doses of LTG (Spearman's correlation coefficient = 0.403, p=0.001). However, these results must be interpreted with caution owing to the small number of participants in some of the genotype groups. In particular, the reliability of the rs34059508 (G465R) component of the model is questionable, with only three heterozygotes in the study population. All three of these individuals had a relatively high LTG dose requirement which, if coincidental, might have over-estimated the relative importance of this polymorphism. By comparison, rs806383 showed a more even genotype distribution amongst the study population and its contribution to LTG dose requirements can be concluded with more confidence. In summary, genetic variants in *SLC22A1* appear to have a significant influence on LTG dose requirement but it is questionable whether this study has reliably quantified the extent of that influence. Validation of these findings in a separate and substantially larger cohort of patients is required.

UGT1A4

Following univariate and multivariate analyses, no significant associations between LTG maintenance dose and genetic variation in *UGT1A4* were identified. Two polymorphisms in *UGT1A4* have previously been shown to reduce catalytic activity and alter substrate specificity.[159] The first of these is rs6755571, a missense SNP in which a C>A transversion at codon 24 results in a proline to threonine amino acid change (P24T). This SNP was tagged in the current study by rs11568318 (r^2 = 1.0) (*Appendix Table 4*), but univariate analysis failed to detect any significant association with dose (p=0.830, *Appendix Table 10*). The second SNP is another missense polymorphism where a T>G transversion at codon 48 (rs2011425) produces a leucine to valine amino acid change (L48V) [159]. This SNP has recently been shown by Gulcebi and colleagues to be associated with significantly lower plasma LTG levels in a cohort of 35 Turkish epilepsy patients.[183] The mean plasma LTG level was

found to be 29.4% lower in participants who were heterozygous at this locus ($p<0.01$). Mean plasma levels were 2.4mg/L for participants carrying the minor allele (n=11) and 3.4mg/L for the homozygous-major allele group (n=24); none of the participants were homozygous for the minor allele. These results should, however, be viewed with caution; partly because the patient cohort was small (n=35) but, more importantly, because the authors apparently failed to correct their data for either dose or response to treatment. Nonetheless, these results suggest that the L48V variant in *UGT1A4* can influence plasma LTG concentrations, presumably by modifying its metabolism, and would thereby be expected to influence dose requirement. Unfortunately, rs2011425 was not typed in the current study because, at the time of SNP selection, a MAF for this polymorphism could not be established in a Caucasian population. Furthermore, rs2011425 was neither genotyped directly or represented by a tSNP in the GWAS data. Thus, it is not possible to confirm the influence of this polymorphism on LTG dose requirement or on LTG pharmacokinetics in general without further investigation.

SCN1A

Neither univariate nor multivariate analyses identified any significant associations between LTG maintenance dose and genetic variation in *SCN1A*. Tate *et al* reported an association between a splice-site variant (rs3812718) in the *SCN1A* gene and maximal doses of CBZ and PHT,[103] and later described an association of the same variant with PHT serum concentrations at maintenance dose.[181] Like LTG, these drugs exert their anticonvulsant effects by inhibition of the VGSC and thus an association might have been anticipated. The current study did not genotype the rs3812718 SNP directly, but it was tagged by rs922224 ($r^2 = 1.0$) (*Appendix Table 5*), which failed to show a significant association with LTG dose ($p = 0.061$) even before correction for multiple testing. This failure to confirm a previous observation, even indirectly, can be explained by differences in study design and patient populations. Tate *et al* explored maximal rather than maintenance doses and in doing so, they focused more on the limit of tolerability than on drug effectiveness. In addition, a significant proportion of their study population was taking multiple AEDs and they failed to consider the possible influence of drug interactions on doses of PHT and CBZ, compounds which are well-known for their interaction potential. To avoid the possible confounding influence of drug interactions, the current study included patients on LTG monotherapy alone. These differences alone would be sufficient to explain the discrepant findings. However, there also remains the possibility that the original report was a false positive association. Indeed, a study by Zimprich and colleagues did not find a significant difference in the average CBZ doses between genotypes for the *SCN1A* splice-site variant (rs3812718), and failed to even report a trend in the same direction as the original study.[184]

The notion of a false-positive association is more in keeping with the current study, particularly when the multivariate analysis of *SCN1A* SNPs failed to detect a single significant model, suggesting that variation in this gene has no bearing on the dose requirement of LTG.

SCN2A

Univariate analyses identified seven SNPs in *SCN2A* that were significantly associated with LTG maintenance dose before correction for multiple comparisons (*Table 3.5*). These were all intronic SNPs, and cumulatively tagged an additional 23 polymorphisms (*Appendix Table 6*). A stepwise linear regression model was constructed using the 51 available SNP genotypes in the *SCN2A* gene, generating a dose-predictive model (*Table 3.20*) containing a single intronic SNP (rs20119067) that did not tag any additional genetic variants. This model associated with 9.4% of the observed variation in LTG maintenance dose (p=0.010), however the explanatory capacity of the equation derived from observed and expected doses was poor (Spearman's correlation coefficient = 0.219, p=0.032). A report by Sills *et al* previously identified a weak association between a non-synonymous polymorphism in *SCN2A* (rs17183814, R19K) and response to AED treatment.[170] The present study directly genotyped this SNP, but an association with LTG maintenance dose requirement was not identified (p=0.892, *Appendix Table 10*). A more recent study showed an association between genotype at an intronic SNP locus (rs2304016) and resistance to AED treatment in a cohort of 471 Chinese participants.[185] This polymorphism is not present in Caucasian populations and its influence on LTG maintenance dose requirement could not, therefore, be assessed in the present study. Previous studies have demonstrated a modest contribution of variation in *SCN2A* to heterogeneity in the response to AED therapy. The results of the present study indicate that variation in *SCN2A* might also be associated with maintenance dose requirement of sodium channel blocking drugs, although again its relative importance appears to be limited.

SCN3A

Univariate analyses of genetic variation in *SCN3A* identified significant association between an intronic SNP (rs11903851) and maintenance dose requirement of LTG (*Table 3.5*). This association did not, however, remain significant after correction for multiple comparisons. The rs11903851 SNP was found to tag an additional 15 polymorphisms (*Appendix Table 7*); 14 further intronic SNPs and one in the coding region of the gene (rs16850131) but which did not alter the amino acid sequence (L481L). As a result, it is debatable whether this SNP, or indeed any that it tags, has functional relevance. Three stepwise linear regression models were constructed using genetic variation in *SCN3A* (*Table 3.21*). The last of these models comprised three SNP genotypes, (rs11903851, rs1982208 and rs17829596) that collectively associated with 23.5% of the observed variation in LTG maintenance

dose (p<0.01). However, the explanatory capacity of the equation built on *SCN3A* genotypes and derived from observed and expected doses was poor (Spearman's correlation coefficient = 0.246, p=0.016). Overall, the results of these univariate and multivariate analyses suggest that genetic variation in *SCN3A* may contribute towards LTG maintenance dose requirement. However, given that the association with rs11903851 was not significant after correction, that the variant genotypes do not appear to have functional consequences, and that the multivariate model had limited explanatory ability, the extent of this contribution, if genuine, is likely to be modest.

Combined multivariate analyses

SCN1A, SCN2A, SCN3A

A multivariate model was constructed to determine whether genetic variation across all three sodium channel genes investigated in this study could accurately predict LTG maintenance dose requirements when considered together. This was deemed to be a reasonable approach as it isolated genetic influences on LTG pharmacodynamics from those on its pharmacokinetics and also because these three genes are located together on chromosome 2 and, as such, there may be an element of LD across all three. The optimal multivariate model included four SNPs from *SCN3A* and two from *SCN2A* and associated with 45.0% of the observed variability in LTG maintenance dose (*Table 3.22*). All of the SNPs that directly comprised the model were intronic and they tagged an additional 16 SNPs in *SCN3A* (*Appendix Table 7*). The dosing equation derived from this model had a moderate explanatory capability (0.626, p<0.001) indicating that combined genetic variation in *SCN2A* and *SCN3A* may be important factors in determining LTG maintenance dose. This would support the observations made in the analysis of these two genes in isolation and also serves to confirm that the *SCN1A* gene does not have a significant influence on maintenance dose requirements of LTG.

SCN1A, SCN2A, SCN3A, UGT1A4, SLC22A1

Finally, multivariate models were constructed using stepwise linear regression to determine whether genetic variation across all five genes can be combined to accurately predict maintenance dose. All 173 SNP genotypes and the five non-genetic variables were included in the analysis, with the optimal model associating with 60.6% of the observed variation in LTG maintenance dose (Table 3.23). This model comprised four SNPs from *SLC22A1*, two from *SCN2A* (which tagged six additional SNPs, *Appendix Table 6*), two from *SCN3A* (which tagged one further SNP, *Appendix Table 7*), and one SNP from *SCN1A*. All of these SNPs were intronic, with the exception of the known missense polymorphism in *SLC22A1* (rs34059508, G465R) discussed above. The dosing equation derived from

this multivariate model showed moderate explanatory capabilities (Spearman's correlation coefficient = 0.686, $p<0.001$) indicating that, if employed in an alternative cohort, the model may potentially provide reasonably accurate dose predictions based on genotypes of the nine SNP variants included. Of all the multivariate analyses undertaken in this project, this model associated with the greatest amount of observed variation in maintenance dose (60.6%). Overall, this model suggests that genetic variation in *SLC22A1*, *SCN3A*, *SCN2A*, and to a lesser extent *SCN1A,* is responsible for determining LTG maintenance dose requirement, whereas variation in the *UGT1A4* gene does not appear influential.

Summary of results

Overall, five demographic variables and 173 sites of genetic variation across five candidate genes have been analysed to investigate associations with LTG maintenance dose. None of the univariate analyses remained significant after correction for multiple comparisons using FDR, which is perhaps unsurprising given the number of tests undertaken. None of the demographic variables appeared to associate with LTG maintenance dose, suggesting that age, sex, epilepsy-type, etc, do not impact on dosing of LTG. This is somewhat intriguing, as there is evidence for other AEDs that dose requirements are higher in males (due to higher body weight), decrease with age (due to decline of hepatic metabolism) and are lower in generalised epilepsies (mechanism unknown). Of the genetic variants found to be significant before FDR correction, seven were in *SCN2A*, six were in *SLC22A1*, and one was in *SCN3A*. Of these, the putatively functional missense SNP in *SLC22A1* that results in a G465R amino acid substitution is of particular note. Interestingly, polymorphisms in the *UGT1A4* or *SCN1A* genes did not appear to associate with dose, despite the fact that these genes encode proteins which might have been expected *a priori* to play a prominent role in LTG pharmacokinetics and pharmacodynamics, respectively.

Multivariate analyses were undertaken to determine whether combinations of genetic variants and non-genetic factors could explain the observed variation in LTG maintenance dose requirement. Of all the multivariate models that were generated, the one which associated with the greatest amount of variation in LTG maintenance dose incorporated polymorphisms from four of the five candidate genes; *SLC22A1*, *SCN3A, SCN2A,* and *SCN1A*. The model appeared to explain 60.6% of the observed variation in LTG maintenance dose, and the dose equation derived from this model showed stronger explanatory capabilities than the other models (Spearman's correlation coefficient = 0.686, $p<0.01$). Of the four genes which comprise this model, variation in *SLC22A1* appears to influence dose

requirements to the greatest degree, contributing four of nine SNPs to the model. The effect of variation in *SLC22A1* on maintenance dose requirement is further highlighted from the multivariate analysis of *SLC22A1* alone; 31.0% of the variation in maintenance dose was explained by four SNPs, and the model demonstrated reasonable explanatory properties (Spearman's correlation coefficient = 0.403).

Genetic variation in *SCN2A* and *SCN3A* also appeared to contribute to LTG maintenance dose requirements. Multivariate models were constructed for each gene individually and associated with 24.2% and 23.5% of the variation in dose, respectively. However, both of these models showed poor explanatory capabilities when observed and predicted doses were correlated (Spearman's correlation coefficients: 0.219 for *SCN2A*, 0.246 for *SCN3A*). In contrast, when the genetic variation across *SCN1A*, *SCN2A* and *SCN3A* was analysed collectively in a combined stepwise linear regression model, the optimal model appeared to associate with 45.0% of the variation in dose and showed a moderate explanatory capacity (Spearman's correlation coefficient 0.626). This combined sodium channel model incorporated four SNPs from *SCN3A* and two from *SCN2A,* with no contribution from *SCN1A*. Thus, inter-individual differences in LTG dose also appear to be dependent, at least in part, on genetic variation in *SCN2A* and *SCN3A* genes and this association is most evident when variation across the genomic region rather than single genes is considered.

Limitations

Design of study

There are several stages where confounders may have been introduced to this study. The analysis itself was retrospective and employed DNA samples and clinical information that had been collected as part of a previous clinical trial. However, the data in that trial were collected prospectively and, as such, can be assumed to be reliable and unbiased. Nevertheless, the SANAD trial was not designed with a pharmacogenetic analysis in mind and so, in some aspects, the case report forms were less than adequate. One of the principal concerns, in this regard, is the use of concomitant medication and its reliable reporting in the trial notes. Patients were specifically excluded from this analysis if it was known that they were being treated with multiple AEDs, even where those drugs might not be expected to interact with LTG. Some established AEDs are well known for their ability to affect LTG pharmacokinetics; sodium valproate decreases LTG clearance by approximately 60%, and PHT and CBZ increase LTG clearance by 125% and 30-50%, respectively.[69] Co-administration with these agents would clearly have had an impact on LTG dose requirements. Unfortunately, it was not possible to exclude the possible confounding effects of co-medication with other drugs as this was not routinely recorded in the SANAD study notes. For example, recent studies have demonstrated that LTG serum levels are significantly reduced by the ethinyl estradiol component of the oral contraceptive pill.[74, 75] Fifty-percent of participants in this analysis were female and many of those were of child-bearing age. The use of oral contraceptives in this sub-population could, in theory, have affected LTG dose requirement but without knowledge of contraceptive use, this potential confounder could not be excluded nor controlled for.

Further significant shortcomings in the patient notes included the absence of patient weight and/or body mass index (BMI) and information regarding smoking status. Both of these demographic variables have the potential to influence LTG dosing requirements. Persons with higher BMIs are likely to require higher doses of LTG to achieve similar volume-distribution and steady concentrations to those with 'normal' BMIs; the opposite is likely to be true for participants with low BMI. In any future studies of genetic influence on AED dose requirement, it is imperative that body mass is included in the multivariate analysis or controlled for by reporting serum drug concentrations. In the present study, although the relationship between LTG dose and serum concentration is linear, serum-LTG concentration would have provided a more accurate outcome measure and arguably negated the need to account for patient BMI. Smoking status should also be included as a non-genetic variable; a recent preliminary study by Reinsberger and colleagues demonstrated that smokers have a significantly lower serum level-to-dose ratio for LTG than non-smokers ($p=0.0014$).[186]

The mechanism for this association is not yet certain, although it is thought to be reflective of tobacco smoke and nicotine inducing glucuronidation enzymes for which LTG is a substrate. Although this study was relatively small (n=44) and recruited predominantly females, it highlights another possible shortcoming of the current analysis and reflects the limitations of undertaking retrospective pharmacogenetic analysis on clinical data which was not collected for that purpose.

Implementation of study

In addition to unavoidable confounding factors arising from the clinical data, there are several potential limitations to the methodology. Patients were included in the analysis if they achieved seizure freedom for a period of at least 12 consecutive months on a stable dose of LTG monotherapy. The rationale for this was to include only those patients who responded to LTG monotherapy so that a maintenance dose could be established. However, from a pragmatic perspective, this might be considered an overly conservative approach as it effectively excluded all patients who experienced a significant reduction in their seizure frequency and who opted to continue treatment with LTG as a result. Anyone with a high pre-treatment seizure frequency which was reduced to almost zero was excluded and yet, arguably, these individuals were optimally controlled and receiving a maintenance dose of LTG. Their exclusion undoubtedly had an effect on the power of the study.

Another possible confounder is the introduction of selection bias on the basis of seizure type. If, as discussed above, the study favoured patients with relatively low pre-treatment seizure frequency, then it may have preferentially included those with partial-onset and primary generalised tonic-clonic seizures at the expense of patients with more frequent absence or myoclonic events. Patients with primary generalised epilepsies tend to require lower doses of AEDs for effective or optimal seizure control, which could have skewed the overall dataset if they were unknowingly recruited with preference. In the present study the distribution of participants between the IGE, LRE and UNC seizure groups was 11.5%, 61.5% and 27.0%, respectively. Although the proportion of LRE is reasonable given the cohort was largely an adult population, a higher proportion of IGE and a lower proportion of UNC might have been expected in the final analysis, the reasons for which are unclear. There is evidence to suggest that LTG is not efficacious in the treatment of generalised-seizures, which may provide an explanation for the observed distribution of epilepsy types in the present study. Conversely, it could be that patients with UNC epilepsy may respond particularly well to LTG monotherapy, or equally the observed trend may be a combination of both these factors. Aside from seizure type, the participants included in the analysis of this study were a good representation of the SANAD cohort as a whole for other variables such as age and gender distribution. Overall, even though the present trial has analysed a minority of the two-thousand four-hundred and thirty-seven

participants enrolled on the SANAD trial, the analysed participants show characteristics well matched to the original study. The apparent discrepancy between the observed and expected epilepsy-type distribution is likely a reflection of LTGs different efficacy toward epilepsy-types, and, given that epilepsy type failed to show a significant association with Ln dose in both univariate (ANOVA $p = 0.105$) and multivariate analyses, the influence of this selection bias is likely to be slight.

There are some possible shortcomings in the SNP selection process employed in this study, specifically relating to the inclusion of putatively functional polymorphisms. In an effort to minimise the required number of genotypings, tSNPs were only included in the analysis if they had a MAF of greater than 5% in a Caucasian population. For putatively functional SNPs, the MAF cut-off was lowered to 1%. In hindsight, however, it might have been more appropriate to include all known functional SNPs (except those not represented in a Caucasian population) irrespective of their MAF. This would have captured more of the important variation across each candidate gene and permitted the inclusion of potentially informative polymorphisms that do not have widely reported allele frequencies. An example of such a SNP is the non-synonymous rs2011425 in *UGT1A4* that has recently been shown in a Turkish study[181] to be associated with plasma concentrations of LTG but which, at the time of SNP selection, did not have a published MAF. A caveat to the inclusion of rarer SNPs is that, since only 96 participants were included in the final analyses, this study was not sufficiently powered to detect associations at the required level of certainty ($p < 0.05$). In retrospect, given that the number of participants available for inclusion was determined at the outset, it may have been appropriate to undertake a power calculation to determine a suitable MAF threshold for inclusion/exclusion of SNPs. However, this approach would have resulted in the exclusion of many functional SNPs from the analysis, and therefore the inclusion of rarer SNPs and relying upon subsequent validation studies to confirm/reject apparent associations was arguably appropriate in this instance.

As previously discussed, using LTG serum concentration as an outcome measure may have been more appropriate than maintenance dose requirement. Not only would this have gone some way towards accounting for identified confounders such as patient BMI, smoking status, and concomitant medication use, but could also have eliminated the inherent bias from using an outcome measure that combines efficacy and tolerability. The candidate genes selected for this study were based upon the pharmacokinetic and pharmacodynamic profiles of LTG, with polymorphisms in these genes likely to influence both efficacy and tolerability to varying degrees. In contrast, genetic variation in alternative genes is unlikely to influence the efficacy of LTG; however it is quite conceivable that the tolerability aspect of LTG dosing could be affected by polymorphisms in alternative genes such as

those encoding the HLA-proteins, resulting in a phenotype with a predisposition to hypersensitivity reactions. Although such patients would have been excluded from the current analysis, this serves to illustrate that, in an ideal scenario, the efficacy and tolerability aspects of LTG should be considered separately.

Impact of this study

This study has sought to determine the impact of genetic variation on LTG maintenance dose requirements when successfully employed as monotherapy in the treatment of epilepsy. Through a candidate gene approach, it has shown that variation in *SLC22A1*, *SCN2A* and *SCN3A* influences LTG dose requirement. This is a novel finding and therefore represents an important addition to the existing knowledgebase on AED pharmacogenetics. As part of the analysis, stepwise linear regression allowed the development of a dose-predictive equation for LTG maintenance dose which can explain up to 61% of inter-individual variability in LTG dose requirement when used as monotherapy in newly or recently-diagnosed epilepsy. Although validation of these results is required in an independent and substantially larger cohort, the data presented here have the potential to form the basis of a predictive algorithm for LTG dosing in clinical practice.

Future work

The results reported in this study require validation in a large, independent cohort. The validation study should be undertaken in a prospective manner to ensure that the confounding factors acknowledged above, including patient BMI, smoking status and concomitant medication use, are appropriately accounted for. Assuming the results withstand validation, it may then be appropriate to undertake a prospective clinical trial to assess whether LTG maintenance dose prediction can improve upon current clinical practice. In theory, participants could be randomised to receive LTG monotherapy via one of two dosing regimens; a dose-predicted regimen based on genotype or the standard dosing regimen currently employed clinically. Comparisons could be made on the basis of efficacy (time to first seizure, time to X months remission), tolerability (time to withdrawal due to adverse effects), or overall effectiveness (time to treatment discontinuation) of the two dosing regimens. Such a study would allow an indication of whether *a priori* dose-prediction can improve or optimise the use of LTG in newly-diagnosed epilepsy. If so, it would have the potential to influence prescribing practice with LTG, validate the use genetic-based dose-prediction models for AEDs, and mark the first step towards personalised-prescribing for epilepsy.

Another avenue for future research is the systematic assessment of the relationship between OCT1 and LTG. Findings of the present study were in keeping with recent data from the Department of Pharmacology at the University of Liverpool which suggest that LTG is subject to OCT1-dependent transport. However, the functional significance (if any) of LTG transport at the level of both hepatocytes and the BBB remains to be clarified. Without an adequate understanding of the anatomical distribution of OCT1 and its relative expression at different cell membranes, it is difficult to draw robust conclusions about the influence of genetic variation in *SLC22A1* on LTG maintenance dose requirement. In particular, the impact of two recognised missense polymorphisms in *SLC22A1* (rs34059508 and rs628031) on the transport of LTG by OCT1 needs to be established using site-directed mutagenesis and drug transport assays. The trends reported by this study would suggest that these variants have a potentially significant effect on LTG pharmacokinetics, even though the mechanisms by which this might occur are yet to be elucidated.

There are no previous reports of transport of LTG, or indeed any AED, by OCT1 or other members of the SLC transporter super-family. This is potentially a novel area of investigation for epilepsy therapeutics. It remains unclear whether LTG transport is specific to OCT1 or if it is also a substrate for other OCT isoforms. OCT1 and OCT2 have previously shown overlapping specificity for various substrate drugs, including cimetidine, quinine and metformin.[118,120, 121] It is possible that LTG is also transported by OCT2 and this would require detailed characterisation and assessment of its functional significance. It would be reasonable to extend these studies to the wider SLC-transporter family, which remains relatively under-investigated with respect to its influence on therapeutics in general. If nothing else, it is essential that all clinically relevant pharmacokinetic and pharmacodynamic pathways are explored as we move towards an era of personalised medicine.

Finally, it may be beneficial to return to the SANAD dataset and identify participants who failed treatment with LTG monotherapy. As discussed in the introduction, there is possibility that some of these failures arose as a result of a “one-size-fits-all” dosing strategy, with exquisitely sensitive patients being over-dosed and withdrawing from treatment due to adverse effects and relatively insensitive patients being under-dosed and withdrawing due to perceived lack of efficacy. In these cases, it may be possible to retrospectively calculate the individual LTG dose requirement based on genotype and to compare that to the actual doses employed. This would add validation to the dose predictive model and provide insight into whether LTG treatment failures in the SANAD study were genuine or a consequence of the relatively inflexible dosing and titration schedules that are currently used in the treatment of epilepsy.

Concluding remarks

LTG is one of the most commonly employed AED agents in the developed world with proven efficacy for the treatment of various epilepsy types.[3] The dose of LTG required to achieve seizure freedom varies considerably between patients; this study aimed to determine whether the observed variation in dose requirement could at least partially be explained by inter-individual variation in pharmacokinetic and pharmacodynamic profiles. We have assessed whether LTG maintenance dose requirement is associated with variation in five candidate genes; *SLC22A1, UGT1A4, SCN1A, SCN2A* and *SCN3A.*

Multivariate stepwise linear regression analysis generated a model that appeared to explain 60.6% of the observed variation in LTG maintenance dose through incorporation of polymorphisms in *SLC22A1, SCN1A, SCN2A* and *SCN3A.* The dose equation derived from this model showed a moderate association between observed and predicted doses (Spearman's correlation coefficient = 0.686, $p<0.01$).

The results show that genetic variation in *SLC22A1, SCN3A, SCN2A,* and to a lesser extent *SCN1A,* influence LTG maintenance dose requirement. Although these results require validation in a larger, independent cohort, these findings potentially form the basis for a predictive algorithm for LTG maintenance dose requirement.

References

1. Marson AG, Al-Kharusi AM, Alwaidh M, Appleton R, Baker GA, Chadwick DW, et al. The SANAD study of effectiveness of valproate, lamotrigine, or topiramate for generalised and unclassifiable epilepsy: an unblinded randomised controlled trial. *Lancet* 2007;369(9566):1016-26.
2. Marson AG, Al-Kharusi AM, Alwaidh M, Appleton R, Baker GA, Chadwick DW, et al. The SANAD study of effectiveness of carbamazepine, gabapentin, lamotrigine, oxcarbazepine, or topiramate for treatment of partial epilepsy: an unblinded randomised controlled trial. *Lancet* 2007;369(9566):1000-15.
3. Biton V. Pharmacokinetics, toxicology and safety of lamotrigine in epilepsy. *Expert Opin Drug Metab Toxicol* 2006;2(6):1009-18.
4. Engel J, Jr, Pedley TA. *Epilepsy - A Comprehensive Textbook, Second Edition*: Wolters Kluwer / Lippincott Williams & Wilkins, 2008.
5. Goldensohn ES, Porter RJ, Schwartzkroin PA. The American Epilepsy Society: an historic perspective on 50 years of advances in research. *Epilepsia* 1997;38(1):124-50.
6. Magiorkinis E, Sidiropoulou K, Diamantis A. Hallmarks in the history of epilepsy: epilepsy in antiquity. *Epilepsy Behav* 2010;17(1):103-8.
7. Berger H. Uber das Elektrenkephalogramm des Menschen. *Arch. Psychiatrie Nerv.* 1929;87:527-70.
8. Tomson T. Mortality in epilepsy. *J Neurol* 2000;247(1):15-21.
9. Perucca E, Beghi E, Dulac O, Shorvon S, Tomson T. Assessing risk to benefit ratio in antiepileptic drug therapy. *Epilepsy Res* 2000;41(2):107-39.
10. Zarrelli MM, Beghi E, Rocca WA, Hauser WA. Incidence of epileptic syndromes in Rochester, Minnesota: 1980-1984. *Epilepsia* 1999;40(12):1708-14.
11. MacDonald BK, Cockerell OC, Sander JW, Shorvon SD. The incidence and lifetime prevalence of neurological disorders in a prospective community-based study in the UK. *Brain* 2000;123 (Pt 4):665-76.
12. Kotsopoulos IA, van Merode T, Kessels FG, de Krom MC, Knottnerus JA. Systematic review and meta-analysis of incidence studies of epilepsy and unprovoked seizures. *Epilepsia* 2002;43(11):1402-9.
13. Gaitatzis A, Purcell B, Carroll K, Sander JW, Majeed A. Differences in the use of health services among people with and without epilepsy in the United Kingdom: socio-economic and disease-specific determinants. *Epilepsy Res* 2002;50(3):233-41.
14. Heaney DC, MacDonald BK, Everitt A, Stevenson S, Leonardi GS, Wilkinson P, et al. Socioeconomic variation in incidence of epilepsy: prospective community based study in south east England. *BMJ* 2002;325(7371):1013-6.
15. Lindsten H, Stenlund H, Edlund C, Forsgren L. Socioeconomic prognosis after a newly diagnosed unprovoked epileptic seizure in adults: a population-based case-control study. *Epilepsia* 2002;43(10):1239-50.
16. International League Against Epilepsy Commission Report. The epidemiology of the epilepsies: future directions. International League Against Epilepsy. *Epilepsia* 1997;38(5):614-8.
17. Everitt AD, Sander JW. Incidence of epilepsy is now higher in elderly people than children. *BMJ* 1998;316(7133):780.
18. Sander JW. The epidemiology of epilepsy revisited. *Curr Opin Neurol* 2003;16(2):165-70.
19. Bell GS, Sander JW. The epidemiology of epilepsy: the size of the problem. *Seizure* 2002;11 (Suppl A):306-14.
20. Duncan JS, Sander JW, Sisodiya SM, Walker MC. Adult epilepsy. *Lancet* 2006;367(9516):1087-100.

21. Hauser WA, Annegers JF, Rocca WA. Descriptive epidemiology of epilepsy: contributions of population-based studies from Rochester, Minnesota. *Mayo Clin Proc* 1996;71(6):576-86.
22. de la Court A, Breteler MM, Meinardi H, Hauser WA, Hofman A. Prevalence of epilepsy in the elderly: the Rotterdam Study. *Epilepsia* 1996;37(2):141-7.
23. NICE. *The diagnosis and management of the epilepsies in adults and children in primary and secondary care*. 2004. http://www.nice.org.uk/nicemedia/live/10954/29533/29533.pdf (accessed 10th July 2010).
24. Bone I, Lindsay KW. *Neurology and Neurosurgery Illustrated*. Fourth ed: Churchill Livingstone, 2004.
25. Proposal for revised clinical and electroencephalographic classification of epileptic seizures. From the Commission on Classification and Terminology of the International League Against Epilepsy. *Epilepsia* 1981;22(4):489-501.
26. Rho JM, Sankar R, Cavazos JE. *Epilepsy - Scientific Foundation of Clinical Practice* New York: NY: Marcel Decker, 2004. (Neurological Disease & Therapy).
27. Dale M, Haylett D. *Pharmacology Condensed*: Churchill Livingstone, 2004.
28. Watanabe M, Maemura K, Kanbara K, Tamayama T, Hayasaki H. GABA and GABA receptors in the central nervous system and other organs. *Int Rev Cytol* 2002;213:1-47.
29. Mares P, Kubova H. What is the role of neurotransmitter systems in cortical seizures? *Physiol Res* 2008;57 (Suppl 3):S111-20.
30. Mantegazza M, Curia G, Biagini G, Ragsdale DS, Avoli M. Voltage-gated sodium channels as therapeutic targets in epilepsy and other neurological disorders. *Lancet Neurol* 2010;9(4):413-24.
31. St Louis EK. Minimizing AED adverse effects: improving quality of life in the interictal state in epilepsy care. *Curr Neuropharmacol* 2009;7(2):106-14.
32. Cockerell OC, Johnson AL, Sander JW, Hart YM, Shorvon SD. Remission of epilepsy: results from the National General Practice Study of Epilepsy. *Lancet* 1995;346(8968):140-4.
33. Kwan P, Brodie MJ. Early identification of refractory epilepsy. *N Engl J Med* 2000;342(5):314-9.
34. NICE. *The National Institute for Clinical Excellence (Establishment and Constitution) Order*. 1999. http://www.opsi.gov.uk/si/si1999/uksi_19990220_en.pdf (accessed 30th May 2010).
35. NICE. *The National Institute for Health and Clinical Excellence*. 2009. http://www.nice.org.uk/ (accessed 31st May 2010).
36. NICE. *The epilepsies: diagnosis and management of the epilepsies in adults in primary and secondary care*. 2004. http://www.nice.org.uk/nicemedia/live/10954/29529/29529.pdf (accessed 29th May 2010).
37. Beghi E, Gatti G, Tonini C, Ben-Menachem E, Chadwick DW, Nikanorova M, et al. Adjunctive therapy versus alternative monotherapy in patients with partial epilepsy failing on a single drug: a multicentre, randomised, pragmatic controlled trial. *Epilepsy Res* 2003;57(1):1-13.
38. Peltola J, Peltola M, Raitanen J, Keranen T, Kharazmi E, Auvinen A. Seizure-freedom with combination therapy in localization-related epilepsy. *Seizure* 2008;17(3):276-80.
39. Hommes OR, Obbens EA. The epileptogenic action of Na-folate in the rat. *J Neurol Sci* 1972;16(3):271-81.
40. Choi H, Morrell MJ. Review of lamotrigine and its clinical applications in epilepsy. *Expert Opin Pharmacother* 2003;4(2):243-51.
41. Bowden CL, Calabrese JR, Sachs G, Yatham LN, Asghar SA, Hompland M, et al. A placebo-controlled 18-month trial of lamotrigine and lithium maintenance treatment in recently manic or hypomanic patients with bipolar I disorder. *Arch Gen Psychiatry* 2003;60(4):392-400.
42. Lampl C, Katsarava Z, Diener HC, Limmroth V. Lamotrigine reduces migraine aura and migraine attacks in patients with migraine with aura. *J Neurol Neurosurg Psychiatry* 2005;76(12):1730-2.

43. Eisenberg E, Shifrin A, Krivoy N. Lamotrigine for neuropathic pain. *Expert Rev Neurother* 2005;5(6):729-35.
44. LAMICTAL® (lamotrigine) prescribing information. GlaxoSmithKline, Research Triangle Park, NC, USA; 2005.
45. Rambeck B, Wolf P. Lamotrigine clinical pharmacokinetics. *Clin Pharmacokinet* 1993;25(6):433-43.
46. Yuen WC, Peck AW. Lamotrigine pharmacokinetics: oral and IV infusion in man. *Br J Clin Pharmacol* 1988;26:242.
47. Committee JF. *British National Formulary*. 59 ed. London: British Medical Association and Royal Pharmaceutical Society of Great Britain, 2010.
48. Macdonald RL, Greenfield LJ, Jr. Mechanisms of action of new antiepileptic drugs. *Curr Opin Neurol* 1997;10(2):121-8.
49. Castel-Branco MM, Falcao AC, Figueiredo IV, Caramona MM. Lamotrigine pharmacokinetic/pharmacodynamic modelling in rats. *Fundam Clin Pharmacol* 2005;19(6):669-75.
50. Moshe SL. Mechanisms of action of anticonvulsant agents. *Neurology* 2000;55(5 (Suppl 1)):S32-40.
51. Waldmeier PC, Baumann PA, Wicki P, Feldtrauer JJ, Stierlin C, Schmutz M. Similar potency of carbamazepine, oxcarbazepine, and lamotrigine in inhibiting the release of glutamate and other neurotransmitters. *Neurology* 1995;45(10):1907-13.
52. Stefani A, Spadoni F, Siniscalchi A, Bernardi G. Lamotrigine inhibits Ca2+ currents in cortical neurons: functional implications. *Eur J Pharmacol* 1996;307(1):113-6.
53. Wang SJ, Huang CC, Hsu KS, Tsai JJ, Gean PW. Inhibition of N-type calcium currents by lamotrigine in rat amygdalar neurones. *Neuroreport* 1996;7(18):3037-40.
54. Pisani A, Bonsi P, Martella G, De Persis C, Costa C, Pisani F, et al. Intracellular calcium increase in epileptiform activity: modulation by levetiracetam and lamotrigine. *Epilepsia* 2004;45(7):719-28.
55. Zona C, Tancredi V, Longone P, D'Arcangelo G, D'Antuono M, Manfredi M, et al. Neocortical potassium currents are enhanced by the antiepileptic drug lamotrigine. *Epilepsia* 2002;43(7):685-90.
56. Wang SJ, Sihra TS, Gean PW. Lamotrigine inhibition of glutamate release from isolated cerebrocortical nerve terminals (synaptosomes) by suppression of voltage-activated calcium channel activity. *Neuroreport* 2001;12(10):2255-8.
57. Hainsworth AH, McNaughton NC, Pereverzev A, Schneider T, Randall AD. Actions of sipatrigine, 202W92 and lamotrigine on R-type and T-type Ca2+ channel currents. *Eur J Pharmacol* 2003;467(1-3):77-80.
58. Malik S, Arif H, Hirsch LJ. Lamotrigine and its applications in the treatment of epilepsy and other neurological and psychiatric disorders. *Expert Rev Neurother* 2006;6(11):1609-27.
59. Garnett WR. Lamotrigine: pharmacokinetics. *J Child Neurol* 1997;12 (Suppl 1):S10-5.
60. Yau MK, Garnett WR, Wargin WA, al. E. A single-dose, dose-proportionality, and bioequivalence study of lamotrigine in normal volunteers. *Epilepsia* 1991;32 (Suppl 3):S8.
61. Cohen AF, Land GS, Breimer DD, Yuen WC, Winton C, Peck AW. Lamotrigine, a new anticonvulsant: pharmacokinetics in normal humans. *Clin Pharmacol Ther* 1987;42(5):535-41.
62. Ramsay RE, Pellock JM, Garnett WR, Sanchez RM, Valakas AM, Wargin WA, et al. Pharmacokinetics and safety of lamotrigine (Lamictal) in patients with epilepsy. *Epilepsy Res* 1991;10(2-3):191-200.
63. Elwes RD, Binnie CD. Clinical pharmacokinetics of newer antiepileptic drugs. Lamotrigine, vigabatrin, gabapentin and oxcarbazepine. *Clin Pharmacokinet* 1996;30(6):403-15.
64. Remmel RP, Sinz MW. A quaternary ammonium glucuronide is the major metabolite of lamotrigine in guinea pigs. In vitro and in vivo studies. *Drug Metab Dispos* 1991;19(3):630-6.

65. Sinz MW, Remmel RP. Isolation and characterization of a novel quaternary ammonium-linked glucuronide of lamotrigine. *Drug Metab Dispos* 1991;19(1):149-53.
66. Bialer M. The pharmacokinetics and interactions of new antiepileptic drugs: an overview. *Ther Drug Monit* 2005;27(6):722-6.
67. Rowland A, Elliot DJ, Williams JA, Mackenzie PI, Dickinson RG, Miners JO. In vitro characterization of lamotrigine N2-glucuronidation and the lamotrigine-valproic acid interaction. *Drug Metab Dispos* 2006;34(6):1055-62.
68. Johannessen Landmark C, Patsalos PN. Drug interactions involving the new second- and third-generation antiepileptic drugs. *Expert Rev Neurother* 2010;10(1):119-40.
69. Weintraub D, Buchsbaum R, Resor SR, Jr., Hirsch LJ. Effect of antiepileptic drug comedication on lamotrigine clearance. *Arch Neurol* 2005;62(9):1432-6.
70. Armijo JA, Bravo J, Cuadrado A, Herranz JL. Lamotrigine serum concentration-to-dose ratio: influence of age and concomitant antiepileptic drugs and dosage implications. *Ther Drug Monit* 1999;21(2):182-90.
71. May TW, Rambeck B, Jurgens U. Influence of oxcarbazepine and methsuximide on lamotrigine concentrations in epileptic patients with and without valproic acid comedication: results of a retrospective study. *Ther Drug Monit* 1999;21(2):175-81.
72. Anderson GD, Yau MK, Gidal BE, Harris SJ, Levy RH, Lai AA, et al. Bidirectional interaction of valproate and lamotrigine in healthy subjects. *Clin Pharmacol Ther* 1996;60(2):145-56.
73. Sidhu J, Job S, Singh S, Philipson R. The pharmacokinetic and pharmacodynamic consequences of the co-administration of lamotrigine and a combined oral contraceptive in healthy female subjects. *Br J Clin Pharmacol* 2006;61(2):191-9.
74. Christensen J, Petrenaite V, Atterman J, Sidenius P, Ohman I, Tomson T, et al. Oral contraceptives induce lamotrigine metabolism: evidence from a double-blind, placebo-controlled trial. *Epilepsia* 2007;48(3):484-9.
75. Sabers A, Ohman I, Christensen J, Tomson T. Oral contraceptives reduce lamotrigine plasma levels. *Neurology* 2003;61(4):570-1.
76. Brodie MJ, Richens A, Yuen AW. Double-blind comparison of lamotrigine and carbamazepine in newly diagnosed epilepsy. UK Lamotrigine/Carbamazepine Monotherapy Trial Group. *Lancet* 1995;345(8948):476-9.
77. Reunanen M, Dam M, Yuen AW. A randomised open multicentre comparative trial of lamotrigine and carbamazepine as monotherapy in patients with newly diagnosed or recurrent epilepsy. *Epilepsy Res* 1996;23(2):149-55.
78. Brodie MJ, Overstall PW, Giorgi L. Multicentre, double-blind, randomised comparison between lamotrigine and carbamazepine in elderly patients with newly diagnosed epilepsy. The UK Lamotrigine Elderly Study Group. *Epilepsy Res* 1999;37(1):81-7.
79. Steiner TJ, Dellaportas CI, Findley LJ, Gross M, Gibberd FB, Perkin GD, et al. Lamotrigine monotherapy in newly diagnosed untreated epilepsy: a double-blind comparison with phenytoin. *Epilepsia* 1999;40(5):601-7.
80. Brown SW. Anticonvulsant side-effects: a self report questionnaire for use in community surveys. *Br J Clin Pract* 1982;18:147-9.
81. Brodie MJ, Chadwick DW, Anhut H, Otte A, Messmer SL, Maton S, et al. Gabapentin versus lamotrigine monotherapy: a double-blind comparison in newly diagnosed epilepsy. *Epilepsia* 2002;43(9):993-1000.
82. Frank LM, Enlow T, Holmes GL, Manasco P, Concannon S, Chen C, et al. Lamictal (lamotrigine) monotherapy for typical absence seizures in children. *Epilepsia* 1999;40(7):973-9.
83. Coppola G, Auricchio G, Federico R, Carotenuto M, Pascotto A. Lamotrigine versus valproic acid as first-line monotherapy in newly diagnosed typical absence seizures: an open-label, randomized, parallel-group study. *Epilepsia* 2004;45(9):1049-53.
84. Glauser TA, Cnaan A, Shinnar S, Hirtz DG, Dlugos D, Masur D, et al. Ethosuximide, valproic acid, and lamotrigine in childhood absence epilepsy. *N Engl J Med* 2010;362(9):790-9.

85. Prasad A, Kuzniecky RI, Knowlton RC, Welty TE, Martin RC, Mendez M, et al. Evolving antiepileptic drug treatment in juvenile myoclonic epilepsy. *Arch Neurol* 2003;60(8):1100-5.
86. Guerrini R, Dravet C, Genton P, Belmonte A, Kaminska A, Dulac O. Lamotrigine and seizure aggravation in severe myoclonic epilepsy. *Epilepsia* 1998;39(5):508-12.
87. Vogel F. Medizinische probleme der humangenetik. *Med Kinderheilkunde* 1959;12:52-125.
88. Clayman CB, Arnold J, Hockwald RS, Yount EH, Jr., Edgcomb JH, Alving AS. Toxicity of primaquine in Caucasians. *J Am Med Assoc* 1952;149(17):1563-8.
89. Alving AS, Carson PE, Flanagan CL, Ickes CE. Enzymatic deficiency in primaquine-sensitive erythrocytes. *Science* 1956;124(3220):484-5.
90. Hirono A, Beutler E. Molecular cloning and nucleotide sequence of cDNA for human glucose-6-phosphate dehydrogenase variant A(-). *Proc Natl Acad Sci USA* 1988;85(11):3951-4.
91. Hirsh J, Dalen J, Anderson DR, Poller L, Bussey H, Ansell J, et al. Oral anticoagulants: mechanism of action, clinical effectiveness, and optimal therapeutic range. *Chest* 2001;119 (Suppl 1):8S-21S.
92. Tan GM, Wu E, Lam YY, Yan BP. Role of warfarin pharmacogenetic testing in clinical practice. *Pharmacogenomics* 2010;11(3):439-48.
93. Klein TE, Altman RB, Eriksson N, Gage BF, Kimmel SE, Lee MT, et al. Estimation of the warfarin dose with clinical and pharmacogenetic data. *N Engl J Med* 2009;360(8):753-64.
94. Caraco Y, Blotnick S, Muszkat M. CYP2C9 genotype-guided warfarin prescribing enhances the efficacy and safety of anticoagulation: a prospective randomized controlled study. *Clin Pharmacol Ther* 2008;83(3):460-70.
95. Anderson JL, Horne BD, Stevens SM, Grove AS, Barton S, Nicholas ZP, et al. Randomized trial of genotype-guided versus standard warfarin dosing in patients initiating oral anticoagulation. *Circulation* 2007;116(22):2563-70.
96. Hillman MA, Wilke RA, Yale SH, Vidaillet HJ, Caldwell MD, Glurich I, et al. A prospective, randomized pilot trial of model-based warfarin dose initiation using CYP2C9 genotype and clinical data. *Clin Med Res* 2005;3(3):137-45.
97. Little PF. Structure and function of the human genome. *Genome Res* 2005;15(12):1759-66.
98. Finishing the euchromatic sequence of the human genome. *Nature* 2004;431(7011):931-45.
99. Guo YM, Wang Q, Liu YZ, Chen HM, Qi Z, Guo QH. Genetic polymorphisms in cytochrome P4502E1, alcohol and aldehyde dehydrogenases and the risk of esophageal squamous cell carcinoma in Gansu Chinese males. *World J Gastroenterol* 2008;14(9):1444-9.
100. *dbSNP Home Page* http://www.ncbi.nlm.nih.gov/sites/entrez (accessed 23rd August 2010).
101. Al-Chalabi A. *Genome-Wide Association Studies*. 2009. http://cshprotocols.cshlp.org/cgi/content/full/2009/12/pdb.top66?ijkey=d39e3e6aa48cc2d6ad4105b9e6cbd0417ac0db9c&keytype2=tf_ipsecsha&text_only=true (accessed 23rd March 2010).
102. Kwan P, Brodie MJ. Effectiveness of first antiepileptic drug. *Epilepsia* 2001;42(10):1255-60.
103. Tate SK, Depondt C, Sisodiya SM, Cavalleri GL, Schorge S, Soranzo N, et al. Genetic predictors of the maximum doses patients receive during clinical use of the anti-epileptic drugs carbamazepine and phenytoin. *Proc Natl Acad Sci USA* 2005;102(15):5507-12.
104. Makmor-Bakry M, Sills GJ, Hitiris N, Butler E, Wilson EA, Brodie MJ. Genetic variants in microsomal epoxide hydrolase influence carbamazepine dosing. *Clin Neuropharmacol* 2009;32(4):205-12.
105. Hediger MA, Romero MF, Peng JB, Rolfs A, Takanaga H, Bruford EA. The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteinsIntroduction. *Pflugers Arch* 2004;447(5):465-8.
106. Tzvetkov MV, Vormfelde SV, Balen D, Meineke I, Schmidt T, Sehrt D, et al. The effects of genetic polymorphisms in the organic cation transporters OCT1, OCT2, and OCT3 on the renal clearance of metformin. *Clin Pharmacol Ther* 2009;86(3):299-306.

107. Grundemann D, Gorboulev V, Gambaryan S, Veyhl M, Koepsell H. Drug excretion mediated by a new prototype of polyspecific transporter. *Nature* 1994;372(6506):549-52.
108. Zolk O. Current Understanding of the Pharmacogenomics of Metformin. *Clin Pharmacol Ther* 2009;86(6):595-8.
109. Koepsell H, Endou H. The SLC22 drug transporter family. *Pflugers Arch* 2004;447(5):666-76.
110. Koepsell H, Schmitt BM, Gorboulev V. Organic cation transporters. *Rev Physiol Biochem Pharmacol* 2003;150:36-90.
111. Koepsell H. Organic cation transporters in intestine, kidney, liver, and brain. *Annu Rev Physiol* 1998;60:243-66.
112. Koepsell H. Polyspecific organic cation transporters: their functions and interactions with drugs. *Trends Pharmacol Sci* 2004;25(7):375-81.
113. Wright SH, Dantzler WH. Molecular and cellular physiology of renal organic cation and anion transport. *Physiol Rev* 2004;84(3):987-1049.
114. Dresser MJ, Leabman MK, Giacomini KM. Transporters involved in the elimination of drugs in the kidney: organic anion transporters and organic cation transporters. *J Pharm Sci* 2001;90(4):397-421.
115. Amphoux A, Vialou V, Drescher E, Bruss M, Mannoury La Cour C, Rochat C, et al. Differential pharmacological in vitro properties of organic cation transporters and regional distribution in rat brain. *Neuropharmacology* 2006;50(8):941-52.
116. Lips KS, Volk C, Schmitt BM, Pfeil U, Arndt P, Miska D, et al. Polyspecific cation transporters mediate luminal release of acetylcholine from bronchial epithelium. *Am J Respir Cell Mol Biol* 2005;33(1):79-88.
117. Wu X, Huang W, Ganapathy ME, Wang H, Kekuda R, Conway SJ, et al. Structure, function, and regional distribution of the organic cation transporter OCT3 in the kidney. *Am J Physiol Renal Physiol* 2000;279(3):449-58.
118. Masuda S, Terada T, Yonezawa A, Tanihara Y, Kishimoto K, Katsura T, et al. Identification and functional characterization of a new human kidney-specific H+/organic cation antiporter, kidney-specific multidrug and toxin extrusion 2. *J Am Soc Nephrol* 2006;17(8):2127-35.
119. Okabe M, Unno M, Harigae H, Kaku M, Okitsu Y, Sasaki T, et al. Characterization of the organic cation transporter SLC22A16: a doxorubicin importer. *Biochem Biophys Res Commun* 2005;333(3):754-62.
120. Yabuuchi H, Tamai I, Nezu J, Sakamoto K, Oku A, Shimane M, et al. Novel membrane transporter OCTN1 mediates multispecific, bidirectional, and pH-dependent transport of organic cations. *J Pharmacol Exp Ther* 1999;289(2):768-73.
121. Suhre WM, Ekins S, Chang C, Swaan PW, Wright SH. Molecular determinants of substrate/inhibitor binding to the human and rabbit renal organic cation transporters hOCT2 and rbOCT2. *Mol Pharmacol* 2005;67(4):1067-77.
122. Grundemann D, Liebich G, Kiefer N, Koster S, Schomig E. Selective substrates for non-neuronal monoamine transporters. *Mol Pharmacol* 1999;56(1):1-10.
123. Busch AE, Karbach U, Miska D, Gorboulev V, Akhoundova A, Volk C, et al. Human neurons express the polyspecific cation transporter hOCT2, which translocates monoamine neurotransmitters, amantadine, and memantine. *Mol Pharmacol* 1998;54(2):342-52.
124. Grundemann D, Schechinger B, Rappold GA, Schomig E. Molecular identification of the corticosterone-sensitive extraneuronal catecholamine transporter. *Nat Neurosci* 1998;1(5):349-51.
125. Otsuka M, Matsumoto T, Morimoto R, Arioka S, Omote H, Moriyama Y. A human transporter protein that mediates the final excretion step for toxic organic cations. *Proc Natl Acad Sci USA* 2005;102(50):17923-8.
126. Gorboulev V, Ulzheimer JC, Akhoundova A, Ulzheimer-Teuber I, Karbach U, Quester S, et al. Cloning and characterization of two human polyspecific organic cation transporters. *DNA Cell Biol* 1997;16(7):871-81.

127. Peltekova VD, Wintle RF, Rubin LA, Amos CI, Huang Q, Gu X, et al. Functional variants of OCTN cation transporter genes are associated with Crohn disease. *Nat Genet* 2004;36(5):471-5.
128. Wu X, Huang W, Prasad PD, Seth P, Rajan DP, Leibach FH, et al. Functional characteristics and tissue distribution pattern of organic cation transporter 2 (OCTN2), an organic cation/carnitine transporter. *J Pharmacol Exp Ther* 1999;290(3):1482-92.
129. Wagner CA, Lukewille U, Kaltenbach S, Moschen I, Broer A, Risler T, et al. Functional and pharmacological characterization of human Na(+)-carnitine cotransporter hOCTN2. *Am J Physiol Renal Physiol* 2000;279(3):584-91.
130. Grundemann D, Koster S, Kiefer N, Breidert T, Engelhardt M, Spitzenberger F, et al. Transport of monoamine transmitters by the organic cation transporter type 2, OCT2. *J Biol Chem* 1998;273(47):30915-20.
131. Ohashi R, Tamai I, Yabuuchi H, Nezu JI, Oku A, Sai Y, et al. Na(+)-dependent carnitine transport by organic cation transporter (OCTN2): its pharmacological and toxicological relevance. *J Pharmacol Exp Ther* 1999;291(2):778-84.
132. Zhang L, Schaner ME, Giacomini KM. Functional characterization of an organic cation transporter (hOCT1) in a transiently transfected human cell line (HeLa). *J Pharmacol Exp Ther* 1998;286(1):354-61.
133. Tahara H, Kusuhara H, Endou H, Koepsell H, Imaoka T, Fuse E, et al. A species difference in the transport activities of H2 receptor antagonists by rat and human renal organic anion and cation transporters. *J Pharmacol Exp Ther* 2005;315(1):337-45.
134. Sata R, Ohtani H, Tsujimoto M, Murakami H, Koyabu N, Nakamura T, et al. Functional analysis of organic cation transporter 3 expressed in human placenta. *J Pharmacol Exp Ther* 2005;315(2):888-95.
135. Kimura N, Masuda S, Tanihara Y, Ueo H, Okuda M, Katsura T, et al. Metformin is a superior substrate for renal organic cation transporter OCT2 rather than hepatic OCT1. *Drug Metab Pharmacokinet* 2005;20(5):379-86.
136. Dresser MJ, Xiao G, Leabman MK, Gray AT, Giacomini KM. Interactions of n-tetraalkylammonium compounds and biguanides with a human renal organic cation transporter (hOCT2). *Pharm Res* 2002;19(8):1244-7.
137. Takeda M, Khamdang S, Narikawa S, Kimura H, Kobayashi Y, Yamamoto T, et al. Human organic anion transporters and human organic cation transporters mediate renal antiviral transport. *J Pharmacol Exp Ther* 2002;300(3):918-24.
138. Koepsell H, Lips K, Volk C. Polyspecific organic cation transporters: structure, function, physiological roles, and biopharmaceutical implications. *Pharm Res* 2007;24(7):1227-51.
139. Kerb R, Brinkmann U, Chatskaia N, Gorbunov D, Gorboulev V, Mornhinweg E, et al. Identification of genetic variations of the human organic cation transporter hOCT1 and their functional consequences. *Pharmacogenetics* 2002;12(8):591-5.
140. Shu Y, Leabman MK, Feng B, Mangravite LM, Huang CC, Stryke D, et al. Evolutionary conservation predicts function of variants of the human organic cation transporter, OCT1. *Proc Natl Acad Sci USA* 2003;100(10):5902-7.
141. Leabman MK, Huang CC, DeYoung J, Carlson EJ, Taylor TR, de la Cruz M, et al. Natural variation in human membrane transporter genes reveals evolutionary and functional constraints. *Proc Natl Acad Sci USA* 2003;100(10):5896-901.
142. Zhang L, Brett CM, Giacomini KM. Role of organic cation transporters in drug absorption and elimination. *Annu Rev Pharmacol Toxicol* 1998;38:431-60.
143. Zhang L, Dresser MJ, Gray AT, Yost SC, Terashita S, Giacomini KM. Cloning and functional expression of a human liver organic cation transporter. *Mol Pharmacol* 1997;51(6):913-21.
144. Ménard V, Girard H, Harvey M, Pérusse L, Guillemette C. Analysis of inherited genetic variations at the <I>UGT1</I> locus in the French-Canadian population. *Human Mutation* 2009;30(4):677-87.

145. Green MD, Tephly TR. Glucuronidation of amines and hydroxylated xenobiotics and endobiotics catalyzed by expressed human UGT1.4 protein. *Drug Metab Dispos* 1996;24(3):356-63.
146. Mori A, Maruo Y, Iwai M, Sato H, Takeuchi Y. UDP-glucuronosyltransferase 1A4 polymorphisms in a Japanese population and kinetics of clozapine glucuronidation. *Drug Metab Dispos* 2005;33(5):672-5.
147. Sun D, Chen G, Dellinger RW, Duncan K, Fang JL, Lazarus P. Characterization of tamoxifen and 4-hydroxytamoxifen glucuronidation by human UGT1A4 variants. *Breast Cancer Res* 2006;8(4):R50.
148. Mackenzie PI, Bock KW, Burchell B, Guillemette C, Ikushiro S, Iyanagi T, et al. Nomenclature update for the mammalian UDP glycosyltransferase (UGT) gene superfamily. *Pharmacogenet Genomics* 2005;15(10):677-85.
149. Uchaipichat V, Mackenzie PI, Guo XH, Gardner-Stephen D, Galetin A, Houston JB, et al. Human udp-glucuronosyltransferases: isoform selectivity and kinetics of 4-methylumbelliferone and 1-naphthol glucuronidation, effects of organic solvents, and inhibition by diclofenac and probenecid. *Drug Metab Dispos* 2004;32(4):413-23.
150. Strassburg CP. Pharmacogenetics of Gilbert's syndrome. *Pharmacogenomics* 2008;9(6):703-15.
151. Sugatani J, Kojima H, Ueda A, Kakizaki S, Yoshinari K, Gong QH, et al. The phenobarbital response enhancer module in the human bilirubin UDP-glucuronosyltransferase UGT1A1 gene and regulation by the nuclear receptor CAR. *Hepatology* 2001;33(5):1232-8.
152. Strassburg CP, Kneip S, Topp J, Obermayer-Straub P, Barut A, Tukey RH, et al. Polymorphic gene regulation and interindividual variation of UDP-glucuronosyltransferase activity in human small intestine. *J Biol Chem* 2000;275(46):36164-71.
153. Strassburg CP, Manns MP, Tukey RH. Expression of the UDP-glucuronosyltransferase 1A locus in human colon. Identification and characterization of the novel extrahepatic UGT1A8. *J Biol Chem* 1998;273(15):8719-26.
154. Strassburg CP, Nguyen N, Manns MP, Tukey RH. Polymorphic expression of the UDP-glucuronosyltransferase UGT1A gene locus in human gastric epithelium. *Mol Pharmacol* 1998;54(4):647-54.
155. Strassburg CP, Nguyen N, Manns MP, Tukey RH. UDP-glucuronosyltransferase activity in human liver and colon. *Gastroenterology* 1999;116(1):149-60.
156. Strassburg CP, Oldhafer K, Manns MP, Tukey RH. Differential expression of the UGT1A locus in human liver, biliary, and gastric tissue: identification of UGT1A7 and UGT1A10 transcripts in extrahepatic tissue. *Mol Pharmacol* 1997;52(2):212-20.
157. Strassburg CP, Strassburg A, Nguyen N, Li Q, Manns MP, Tukey RH. Regulation and function of family 1 and family 2 UDP-glucuronosyltransferase genes (UGT1A, UGT2B) in human oesophagus. *Biochem J* 1999;338 (Pt 2):489-98.
158. Ritter JK, Owens IS, Negishi M, Nagata K, Sheen YY, Gillette JR, et al. Mouse pulmonary cytochrome P-450 naphthalene hydroxylase: cDNA cloning, sequence, and expression in Saccharomyces cerevisiae. *Biochemistry* 1991;30(48):11430-7.
159. Ehmer U, Vogel A, Schutte JK, Krone B, Manns MP, Strassburg CP. Variation of hepatic glucuronidation: Novel functional polymorphisms of the UDP-glucuronosyltransferase UGT1A4. *Hepatology* 2004;39(4):970-7.
160. Saeki M, Saito Y, Jinno H, Sai K, Hachisuka A, Kaniwa N, et al. Genetic variations and haplotypes of UGT1A4 in a Japanese population. *Drug Metab Pharmacokinet* 2005;20(2):144-51.
161. Catterall WA. From ionic currents to molecular mechanisms: the structure and function of voltage-gated sodium channels. *Neuron* 2000;26(1):13-25.

162. Isom LL, Ragsdale DS, De Jongh KS, Westenbroek RE, Reber BF, Scheuer T, et al. Structure and function of the beta 2 subunit of brain sodium channels, a transmembrane glycoprotein with a CAM motif. *Cell* 1995;83(3):433-42.
163. Isom LL. The role of sodium channels in cell adhesion. *Front Biosci* 2002;7:12-23.
164. Goldin AL, Barchi RL, Caldwell JH, Hofmann F, Howe JR, Hunter JC, et al. Nomenclature of voltage-gated sodium channels. *Neuron* 2000;28(2):365-8.
165. Goldin AL. Resurgence of sodium channel research. *Annu Rev Physiol* 2001;63:871-94.
166. Trimmer JS, Rhodes KJ. Localization of voltage-gated ion channels in mammalian brain. *Annu Rev Physiol* 2004;66:477-519.
167. Catterall W, Kalume F, Oakley JC. NaV1.1 Channels and Epilepsy. *J Physiol* 2010;588 (Pt 11):1849-59.
168. Meisler MH, Kearney JA. Sodium channel mutations in epilepsy and other neurological disorders. *J Clin Invest* 2005;115(8):2010-7.
169. Wallace RH, Wang DW, Singh R, Scheffer IE, George AL, Jr., Phillips HA, et al. Febrile seizures and generalized epilepsy associated with a mutation in the Na+-channel beta1 subunit gene SCN1B. *Nat Genet* 1998;19(4):366-70.
170. Sills GJ, Mohanraj R, Butler E. A single-nucleotide polymorphism in the *SCN2A* gene is associated with uncontrolled epilepsy. *Epilepsia* 2004;44 (Suppl 7):226.
171. Berkovic SF, Heron SE, Giordano L, Marini C, Guerrini R, Kaplan RE, et al. Benign familial neonatal-infantile seizures: characterization of a new sodium channelopathy. *Ann Neurol* 2004;55(4):550-7.
172. Holland KD, Kearney JA, Glauser TA, Buck G, Keddache M, Blankston JR, et al. Mutation of sodium channel SCN3A in a patient with cryptogenic pediatric partial epilepsy. *Neurosci Lett* 2008;433(1):65-70.
173. Meisler MH, O'Brien JE, Sharkey LM. Sodium channel gene family: epilepsy mutations, gene interactions and modifier effects. *J Physiol* 2010;588 (Pt 11):1841-8.
174. Marson AG, Appleton R, Baker GA, Chadwick DW, Doughty J, Eaton B, et al. A randomised controlled trial examining the longer-term outcomes of standard versus new antiepileptic drugs. The SANAD trial. *Health Technol Assess* 2007;11(37):iii-iv, ix-x, 1-134.
175. Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, et al. A second generation human haplotype map of over 3.1 million SNPs. *Nature* 2007;449(7164):851-61.
176. Integrating ethics and science in the International HapMap Project. *Nat Rev Genet* 2004;5(6):467-75.
177. The International HapMap Project. *Nature* 2003;426(6968):789-96.
178. Oeth P, del Mistro G, Marnellos G, Shi T, van den Boom D. Qualitative and quantitative genotyping using single base primer extension coupled with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MassARRAY). *Methods Mol Biol* 2009;578:307-43.
179. van der Weide J, Steijns LS, van Weelden MJ, de Haan K. The effect of genetic polymorphism of cytochrome P450 CYP2C9 on phenytoin dose requirement. *Pharmacogenetics* 2001;11(4):287-91.
180. Hung CC, Lin CJ, Chen CC, Chang CJ, Liou HH. Dosage recommendation of phenytoin for patients with epilepsy with different CYP2C9/CYP2C19 polymorphisms. *Ther Drug Monit* 2004;26(5):534-40.
181. Tate SK, Singh R, Hung CC, Tai JJ, Depondt C, Cavalleri GL, et al. A common polymorphism in the SCN1A gene associates with phenytoin serum levels at maintenance dose. *Pharmacogenet Genomics* 2006;16(10):721-6.
182. Shu Y, Sheardown SA, Brown C, Owen RP, Zhang S, Castro RA, et al. Effect of genetic variation in the organic cation transporter 1 (OCT1) on metformin action. *J Clin Invest* 2007;117(5):1422-31.

183. Gulcebi MI, Ozkaynakci A, Goren MZ, Ozkara C, Gulhan R, Onat FY. *9th European Congress on Epileptology*, 2010, Rhodes, .
184. Zimprich F, Stogmann E, Bonelli S, Baumgartner C, Mueller JC, Meitinger T, et al. A functional polymorphism in the SCN1A gene is not associated with carbamazepine dosages in Austrian patients with epilepsy. *Epilepsia* 2008;49(6):1108-9.
185. Kwan P, Poon WS, Ng HK, Kang DE, Wong V, Ng PW, et al. Multidrug resistance in epilepsy and polymorphisms in the voltage-gated sodium channel genes SCN1A, SCN2A, and SCN3A: correlation among phenotype, genotype, and mRNA expression. *Pharmacogenet Genomics* 2008;18(11):989-98.
186. Reinsberger C, Dorn T, Kramer G. Smoking reduces serum levels of lamotrigine. *Seizure* 2008;17(7):651-3.
187. *Amino acid and codon table* http://www.cbs.dtu.dk/courses/27619/codon.html (accessed 4th January 2010).

Appendix I

Study	Number of participants	LTG Group	Control Group	Dose Regimen, Assessments and Study Duration	Study Type	Inclusion Criteria	Exclusion Criteria	Outcome Measure	Dropouts / Withdrawals (?Accounted For)	Observations	Conclusions
Brodie *et al*[76]	260	LTG 150mg, n=131	CBZ 600mg, n=129	4-week planned, fixed dose escalation. Assessed every 2-weeks until 12-weeks, then every 6-weeks. Dose altered if seizures persist, AE, or blood concentration not adequate. 48-weeks total duration.	Double-blind, randomised, parallel-group comparison. Patients stratified according to seizure type.	> 13-years old. Newly diagnosed epilepsy.	Previous AED use.	TTW, proportion of patients seizure-free at 24 and 40-weeks	151/260 participants completed the study. 27 patients from the LTG group withdrew, and 38 patients from the CBZ group. All withdrawals accounted for.	Higher seizure frequency before treatment predisposed to further seizures after 6-week treatment period. LTG less dropouts than CBZ. Difference in AEs between groups. Similar efficacy	No difference in efficacy between LTG and CBZ in the treatment of partial seizures (with or without secondary generalisation). Bias between groups due to non-matching of seizure frequency. Statistically significant difference in AEs & 'sleepiness' – lower in LTG group.
Reunanen *et al*[77]	343	LTG 100mg, n=115 LTG 200mg, n=111	CBZ 600mg, n=117	4-week escalation period. Dose adjusted if persistent seizures / significant AE. Reassessed at end of weeks 2, 6, 12 and 30. 30 weeks total duration.	Blinded randomisation, open treatment.	> 12-years old. Newly diagnosed or current epilepsy. No concurrent medications	More than 2 doses of AED in prior 6-months. History of status epilepticus. Significant organic / psychiatric disease. Abnormal laboratory values. Pregnancy, lactation, or exposure to risk of pregnancy.	TTW, proportion of patients seizure-free at 30-weeks.	LTG-100, n=23 LTG-200, n=10 CBZ-600, n=29 All accounted for.	Higher proportion of LTG-200 seizure free at 30-weeks (60.4%) than LTG-100 (51.3%) and CBZ-600 (54.7%). More AEs on CBZ-600 (66%) than LTG-100 (53%) & LTG-200 (58%).	No statistically significant difference between groups. LTG appeared equally effective but better tolerated than CBZ.
Brodie *et al*[78]	150	LTG, n=102	CBZ, n=48	6-week planned, fixed dose escalation. Assessed at 2, 4, 6, 12 and 24 weeks; unscheduled visits allowed as necessary. Dosage could be adjusted from week 6 onwards while maintaining the blind. 24-weeks total duration.	Multicentre, randomised, randomised controlled trial	> 65 years old. Newly diagnosed epilepsy. Two or more seizures in the past 12-months, at least one during the past 6-months.		Rate of drop out due to adverse events, time to first seizure, proportion of patients completing the study.	92/150 participants completed the study. 30 patients from the LTG group withdrew, and 28 patients from the CBZ group. All accounted for.	Main difference was rate of drop-out due to adverse effects (LTG 18%, CBZ 42%). No difference between drugs in time to first seizure. More patients continued on treatment with LTG than CBZ (LTG 71%, CBZ 42%, $p<0.001$). HR for withdrawal 2.4 (95% CI 1.4-4.0)	LTG can be regarded as an acceptable choice as initial treatment for elderly patients with newly diagnosed epilepsy.

Steiner *et al*[79]	181	LTG (modal daily dose 150mg, maximal daily dose 400mg), n=81	PHT (modal daily dose 300mg, maximal daily dose 600mg), n=93	Reassessed at clinic at 2, 4, 6, 8, 12, 18, 24, 36, and 48 weeks. 48-weeks total duration	Double-blind, parallel-group study. Assessed at entry, stratified according to seizure type, then randomised within strata to receive either LTG or PHT	14-75 years old. ≥2 seizures in past 6-months, at least one in previous 3-months.	Previous AED use, absence seizures, significantly abnormal laboratory values, other chronic medical disorders, severe mental sub-normality, alcohol/substance abuse, pregnancy / risk of becoming pregnant.	Percentage of patients remaining on treatment and seizure free. Number of seizures within last 24 and 40-weeks	Seven protocol violators; all accounted for.	Percentages of patients remaining on treatment and seizure free differed little between treatments at 24 and 40-weeks. >50% patients discontinued from both groups.	PHT and LTG appear comparable of the primary efficacy index of percentage of patients remaining seizure-free
Brodie *et al*[81]	309	LTG, n=143	GBP, n=143	GBP 600mg/day, titrated up to 1200mg/day over 2-weeks. LTG 25mg/day, titrated up to 100mg/day over 6-weeks. Dose adjusted according to seizures / AE. 30-weeks total duration.	Multicentre, double-blind, randomised parallel-group study. Seizures classified, adverse events recorded, physical and neurological examinations at baseline and final week / at withdrawal	≥ 16-yearsold. Partial seizures with/without secondary generalisation or primary generalised seizures. Newly diagnosed epilepsy. At least 2-seizures in past 12-months. Untreated patients must have had a seizure in past 3-months.	Absence/myoclonic seizure type, history of status epilepticus, progressive CNS disease, or seizures related to alcohol, acute mental illness or head-trauma. Previous treatment with LTG or GBP, pregnancy,	TTW	291/309 patients included in evaluable population. 19-patients in each group had an exit event. All accounted for.	Median TTW 69-days for GBP, 48-days for LTG. 71.6% of GBP and 67.1% of LTG populations completed study. 80/106 (75.5%) of GBP and 73/96 (76%) of LTG patients remained seizure free during final 12-weeks. 8.9% of GBP and 9.9% of LTG withdrew citing AEs.	GBP and LTG were similarly effective and well tolerated in patients with newly diagnosed epilepsy.
Marson *et al*[2]	1721	LTG, n=378	CBZ, n=378. GBP, n=377. OXC, n=210. TPM, n=378	As per clinical judgment, aim to control seizures with minimum effective dose. Followed up at 3, 6, 12, and yearly intervals. More-regular if clinically indicated	Unblinded, randomised, multicentre, controlled trial. Stratified according to centre, sex, treatment history.	History of two or more clinically definite unprovoked epileptic seizures in previous year. CBZ deemed better standard treatment option compared with VPA (Arm B)	Previous failed monotherapy for one of the AEDs in the study arm. Patient / clinician deemed treatment contraindicated. All seizures acute and symptomatic. 4-years old or younger. History of progressive neurological disease.	TTW, time to 12-month remission.	71 deaths, all patient withdrawals accounted for.	For TTF: LTG significantly better than CBZ (HR 0.78 [95% CI 0.63-0.97]), GBP (0.65 [0.52-0.79]) and showed a non-significant advantage for OXC (1.15 [0.86-1.54]). For time to 12-0-month remission, CBZ has a non-significant advantage versus LTG (0.91 [0.77-1.09])	LTG is clinically better than CBZ in terms of TTF. Given its non-inferiority in comparison to CBZ for time to 12-month remission, it is therefore a cost-effective alternative for patients diagnosed with partial onset seizures.

LTG = lamotrigine; CBZ = carbamazepine; AE = adverse effects; AED = antiepileptic drugs; TTW = time-to-withdrawal; PHT = phenytoin; GBP = gabapentin; OXC = oxcarbazepine; TPM = topiramate; HR = hazard ratio.

Appendix Table 1: Summary of LTG monotherapy trials in the treatment of partial-onset seizures

Appendix II

Appendix Table 2: Amino acid and codon table[187]

Amino Acid	Abbreviation	DNA Codons
Isoleucine	I	ATT, ATC, ATA
Leucine	L	CTT, CTC, CTA, CTG, TTA, TTG
Valine	V	GTT, GTC, GTA, GTG
Phenylalanine	F	TTT, TTC
Methionine	M	ATG
Cysteine	C	TGT, TGC
Alanine	A	GCT, GCC, GCA, GCG
Glycine	G	GGT, GGC, GGA, GGG
Proline	P	CCT, CCC, CCA, CCG
Threonine	T	ACT, ACC, ACA, ACG
Serine	S	TCT, TCC, TCA, TCG, AGT, AGC
Tyrosine	Y	TAT, TAC
Tryptophan	W	TGG
Glutamine	Q	CAA, CAG
Asparagine	N	AAT, AAC
Histidine	H	CAT, CAC
Glutamic acid	E	GAA, GAG
Aspartic acid	D	GAT, GAC
Lysine	K	AAA, AAG
Arginine	R	CGT, CGC, CGA, CGG, AGA, AGG
Stop codons	Stop	TAA, TAG, TGA

Appendix III

SLC22A1

Appendix Table 3: SLC22A1 haploview output - tSNPs

Test	Alleles Captured
rs2297374	rs9295124, rs3798164, rs9295125, rs9295123, rs1443844, rs3818678, rs9347386, rs1382785, rs9347388, rs2297374, rs9295122
rs4709400	rs3822841, rs1867351, rs4709399, rs4709400, rs3798173, rs6935207, rs7769472
rs1871389	rs2083867, rs1871389, rs4709403, rs7744238, rs10455780
rs3798168	rs1867350, rs1871388, rs3798168
rs1564348	rs11753995, rs1564348, rs662138
rs609468	rs609468, rs622591
rs622342	rs650284, rs622342
rs594709	rs628031, rs594709
rs3101826	rs3101826
rs9457843	rs9457843
rs2197296	rs2197296
rs6455682	rs6455682
rs9456505	rs9456505
rs3798167	rs3798167
rs6937722	rs6937722
rs7773429	rs7773429
rs3798174	rs3798174
rs461473	rs461473
rs10455864	rs10455864
rs644992	rs644992
rs3777392	rs3777392
rs683369	rs683369
rs456598	rs456598
rs10455868	rs10455868
rs4646283	rs4646283
rs651164	rs651164
rs619598	rs619598

UGT1A4

Appendix Table 4: UGT1A4 haploview output - tSNPs

Test	Alleles Captured
rs3806591	rs2013030, rs3806591
rs3755319	rs871514, rs7597496, rs4663333, rs2221198, rs4663965, rs10179091, rs4663967, rs3755319, rs3806597, rs4663969, rs4148326, rs10929301, rs6431628, rs6741669, rs7556676, rs6715325, rs3806596, rs4663971, rs4663963, rs7574296, rs4294999, rs4399719, rs2008595, rs4124874
rs6714634	rs17864701, rs6744284, rs4148325, rs4148324, rs6742078, rs11695484, rs6714634, rs17862875, rs6722076, rs887829, rs10929302
rs2018985	rs7604115, rs2018985, rs869283, rs11673726, rs11891311, rs1875263, rs10178992, rs11888459, rs7564935, rs1983023
rs28899189	rs28898621, rs17862878, rs12463641, rs3796088, rs10929293, rs904855, rs1018124, rs4663964, rs904856, rs28898615, rs17864705, rs28899186, rs12468543, rs12468356, rs12466997, rs4663968, rs12479045, rs12466779, rs17863798, rs28899187, rs12052787, rs12479208, rs28899189
rs6431632	rs1042640, rs1500482, rs8330, rs1587493, rs10199525, rs6431632, rs4663972, rs6431631, rs10929303
rs11568318	rs11568318, rs17863800, rs17868341, rs17863795, rs6755571, rs28898590, rs17874945, rs17862874
rs10199882	rs4663335, rs9287649, rs7586006, rs4663973, rs11563250, rs10199882
rs6719561	rs6719561, rs6728520, rs6728940, rs6746002, rs10209214
rs6717546	rs4148329, rs6717546, rs10199512
rs11563251	rs11563251, rs6431630, rs11888492
rs929596	rs13009407, rs929596
rs3771342	rs3771342, rs2003569
rs12475068	rs3755321, rs17868334, rs3806594, rs17863791, rs12477216, rs2013021, rs3732221, rs4556969, rs2013018, rs4663327, rs28898605, rs3732218, rs3806593, rs3806595, rs17862870, rs12475068, rs4233633, rs17874943, rs3732220, rs28946885, rs5020121, rs3892170, rs28898596, rs3893334, rs17863792, rs4663945, rs2885295
rs4148328	rs4148328
rs2302538	rs2302538
rs2011404	rs2011404
rs10203853	rs10203853
rs6431633	rs6431633
rs28946889	rs28946889

SCN1A

Appendix Table 5: SCN1A haploview output - tSNPs

Test	Alleles Captured
rs4667867	rs4667867, rs2390322, rs552878, rs545331, rs492299, rs2859817, rs10167228, rs7580482, rs504059, rs1841547, rs7423423, rs490317, rs13421166, rs4667503, rs2114760, rs2126152, rs508585, rs536744, rs545238, rs7601520, rs1834840, rs6432860, rs1461202, rs577306, rs565348, rs13006006, rs7598539, rs7574618, rs4667862, rs7606888, rs5006656, rs568141, rs567652, rs1461193, rs523119, rs3812719, rs2298771, rs4667861, rs496571, rs10930201
rs4667866	rs4667866, rs16851381, rs2020035, rs16822821, rs16851356, rs12614431, rs16851327, rs1427651, rs12613942, rs994398, rs16851400, rs10497276, rs16851332, rs11674130, rs1381108, rs12998913, rs10497275, rs11686142, rs16851478, rs10930202
rs7607629	rs4667869, rs7607455, rs484926, rs1841548, rs2169312, rs10497278, rs13405797, rs1020852, rs13383881, rs7607543, rs1841549, rs2162600, rs6731591, rs7607629
rs1461197	rs1824551, rs967614, rs1461197, rs1381105, rs11691355, rs10202285, rs2217198, rs6722462, rs1381109, rs1824549, rs991716
rs1841550	rs3812718, rs1542484, rs2195143, rs7609055, rs1841550, rs1972445, rs922224, rs1841546, rs10192608, rs13398150
rs10168027	rs10188577, rs10197430, rs11692675, rs557222, rs11691603, rs10168027
rs17744737	rs12999293, rs17744737, rs11884723
rs10182473	rs6749076, rs1020853, rs10182473
rs498631	rs498631, rs478389
rs10176603	rs10176603
rs1461195	rs1461195
rs16851382	rs16851382
rs17791817	rs17791817
rs13406905	rs13406905
rs13397210	rs13397210
rs12151636	rs12151636

SCN2A

Appendix Table 6: SCN2A haploview output - tSNPs

Test	Alleles Captured
rs4667484	rs12467383, rs1991774, rs12464010, rs12468128, rs4667484, rs2390210, rs2304013, rs3769948, rs2043255, rs12619626, rs12614399, rs6744911, rs16850430, rs16850467, rs3943809, rs2075703, rs3769941, rs12469667, rs12477385, rs6731083, rs3754963, rs4667481, rs3769932, rs16850454, rs12468669, rs1864885, rs11889342, rs1947114, rs3769928, rs16850433, rs12619604, rs12618468, rs12612104, rs4667812, rs6718242, rs4667808, rs16850426
rs12692768	rs4566378, rs7580734, rs7589423, rs10167223, rs2390258, rs764660, rs10204322, rs2116658, rs2163708, rs12692767, rs17245688, rs7581427, rs1368236, rs3769943, rs12692768, rs10199969, rs1579865, rs10203729
rs7593568	rs353138, rs6432820, rs7596560, rs13023748, rs2075704, rs7593568, rs3769938, rs777138, rs13012293, rs1965757, rs1838847, rs11885321, rs2121371, rs4667805, rs10497259, rs1816918, rs7592445
rs2304010	rs10930160, rs2043254, rs16850532, rs2304012, rs4667485, rs2304010, rs1432545, rs2892961, rs4667810, rs1007722, rs1821223, rs1368237, rs10181853, rs4296442, rs1469649, rs1821225
rs4667807	rs10191771, rs17185905, rs2043256, rs6716702, rs7589614, rs10203424, rs4667807, rs4667809, rs7578237, rs4480996, rs935403, rs3769950, rs7596027, rs2060198, rs10930162, rs2116659
rs4667802	rs10174400, rs10197716, rs10182570, rs7589211, rs13432006, rs16850290, rs10192208, rs4667802, rs1470089, rs6709306, rs1866604
rs168478	rs353120, rs353115, rs353118, rs353119, rs353129, rs353121, rs353111, rs168478, rs353122
rs12993173	rs12993173, rs7581811, rs1867864, rs353139, rs997508, rs1371466, rs7607897, rs6432817
rs1966633	rs4303727, rs2390163, rs3769944, rs1966633, rs6738837, rs3769947, rs2304014
rs17182982	rs17182645, rs6755708, rs17182714, rs1439804, rs2390162, rs17182982, rs6736704
rs3769931	rs3769934, rs4387807, rs7567616, rs3769931, rs7586412, rs997507
rs9287858	rs10209034, rs9287857, rs13413719, rs9287856, rs9287858
rs2028892	rs1866603, rs1446579, rs2028892, rs1439805, rs1898970
rs6740895	rs6740895, rs16850311, rs7561826
rs6721613	rs6721613, rs12997787
rs353112	rs353112, rs6741147
rs13387970	rs13387970, rs3769949
rs17182784	rs17182784, rs6718960
rs17183814	rs1529668, rs17183814
rs10207911	rs2119068, rs10207911
rs353116	rs353128, rs353116
rs7573433	rs7573433
rs6705474	rs6705474

rs13025009	rs13025009
rs7600082	rs7600082
rs2119067	rs2119067
rs353123	rs353123
rs10497258	rs10497258
rs10184275	rs10184275
rs2060199	rs2060199
rs16850317	rs16850317
rs1368234	rs1368234
rs17242693	rs17242693
rs7566636	rs7566636
rs17184707	rs17184707

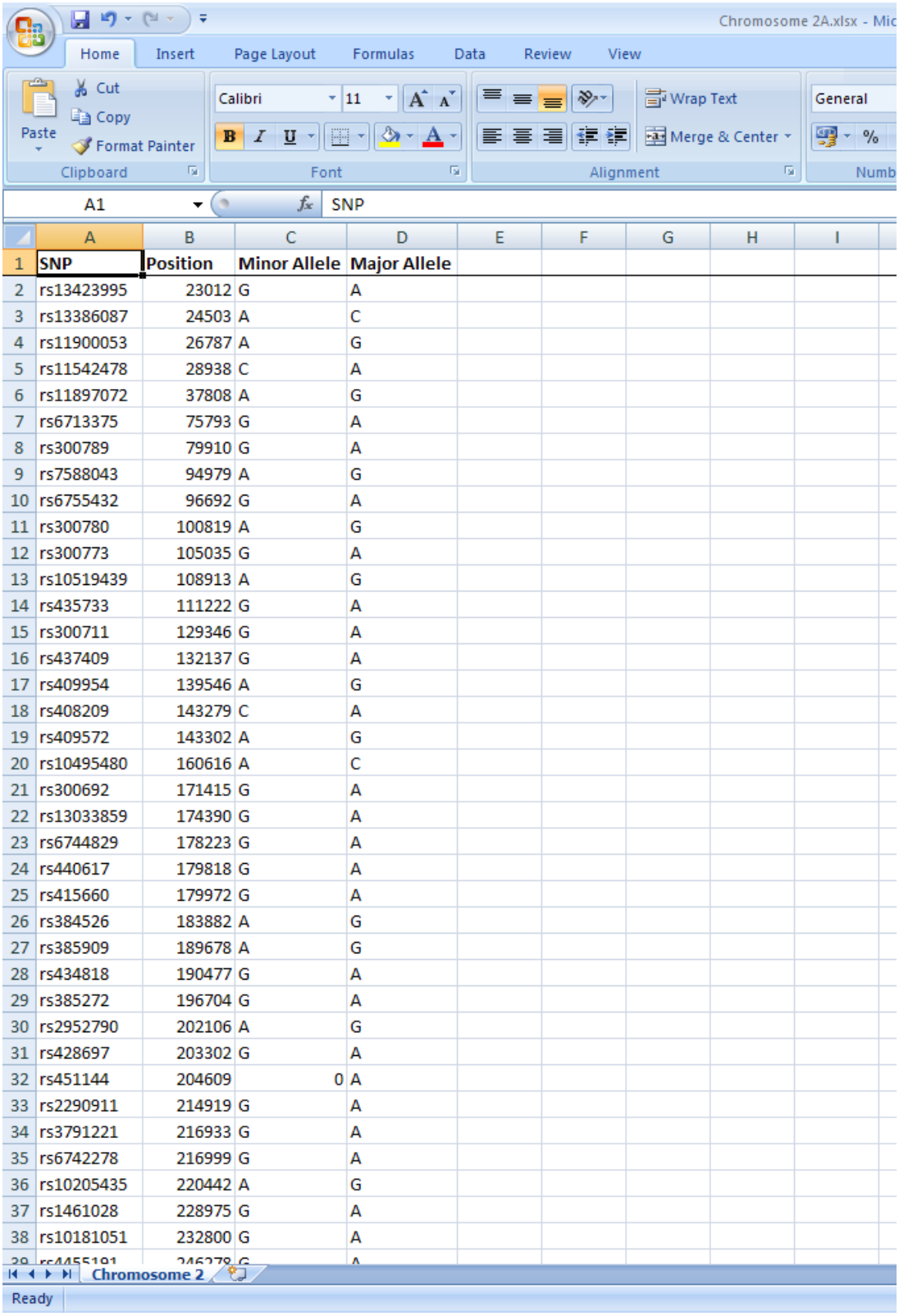
SCN3A

Appendix Table 7: SCN3A haploview output - tSNPs

Test	Alleles Captured
rs16850191	rs4667793, rs3754962, rs7556825, rs11885920, rs16850188, rs4667789, rs7573017, rs11903523, rs16822785, rs4667795, rs16850191, rs4667787, rs3754961, rs16850192, rs1446576, rs6711595, rs16850186, rs4145346, rs6717810, rs7586796, rs7579000, rs2304710, rs4667790
rs17829650	rs1982208, rs11903851, rs2289401, rs1439807, rs17829626, rs1982212, rs16850136, rs16850131, rs1439808, rs1899013, rs17829759, rs1982213, rs16850134, rs10930152, rs17829650, rs2028363
rs11686777	rs10930151, rs10930149, rs1439993, rs1347992, rs11686777, rs1550385, rs3816195, rs10930148
rs3731760	rs4667786, rs3731760, rs6432812, rs1158135, rs7598098, rs10048748
rs17829560	rs17829596, rs17829560, rs10930150, rs13015737
rs6755352	rs7572969, rs12464762, rs6755352, rs10497257, rs11898238, rs12615864, rs1583762
rs11677254	rs11677254, rs11894144
rs6719780	rs6719780, rs16850230
rs2028891	rs2028891, rs6727857
rs7571512	rs7571512, rs7574918
rs6756406	rs6756406, rs1439806
rs7596422	rs7596422
rs1946892	rs1946892
rs3213904	rs3213904
rs4667792	rs4667792
rs2028364	rs2028364
rs2165208	rs2165208
rs13011371	rs13011371
rs2043932	rs2043932
rs2390165	rs2390165

Appendix IV

Screen-shot of GWAS database detailing the SNPs genotyped on Chromosome 2, position of each SNP, and the minor and major allele possibilities at each locus.



Chromosome 2A.xlsx - Mic

Home | Insert | Page Layout | Formulas | Data | Review | View

Cut | Copy | Format Painter | Paste | Clipboard | Calibri | 11 | Font | Wrap Text | Merge & Center | Alignment | General | Numb

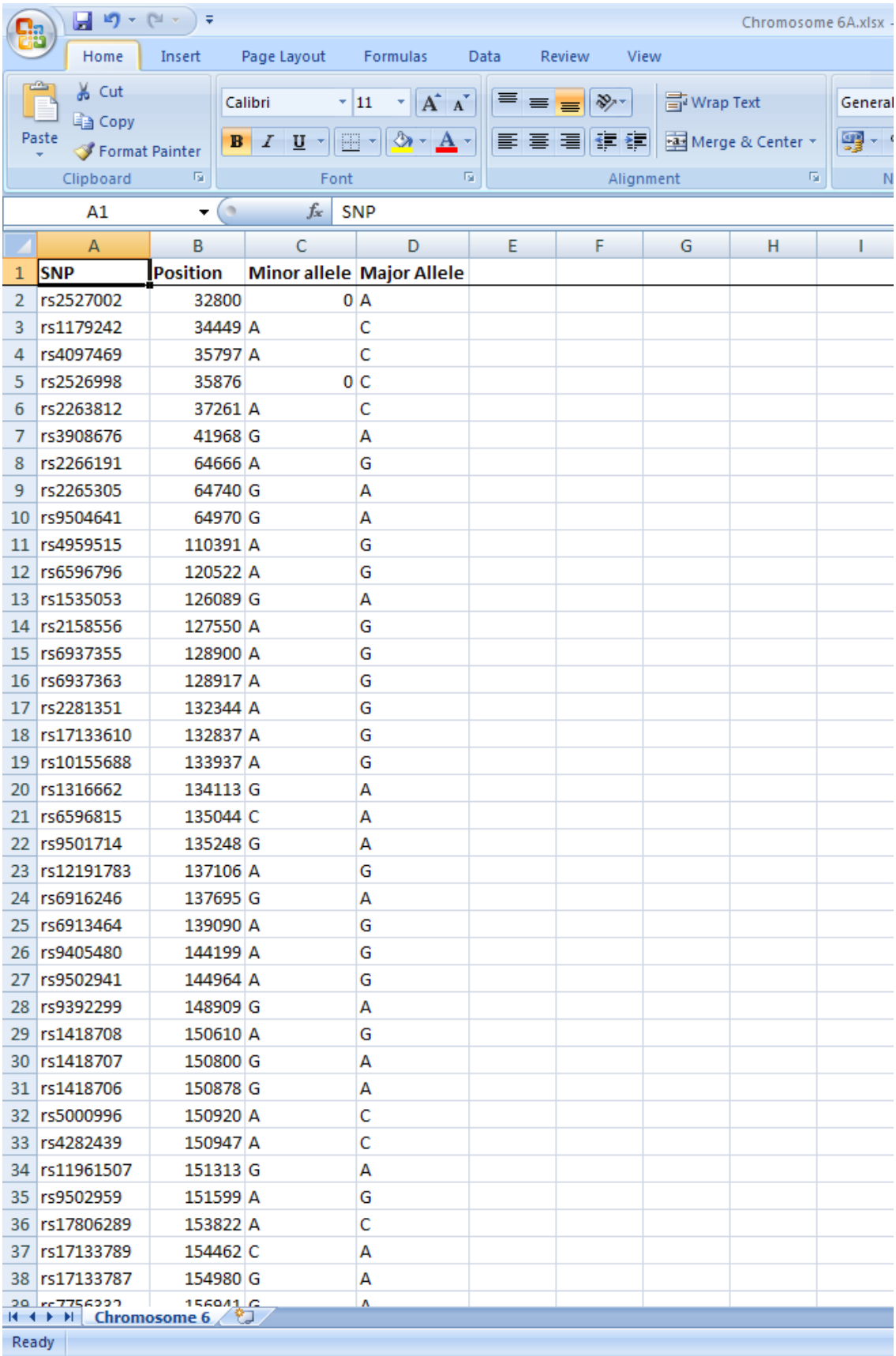
A1 | SNP

	A	B	C	D	E	F	G	H	I
1	SNP	Position	Minor Allele	Major Allele					
2	rs13423995	23012	G	A					
3	rs13386087	24503	A	C					
4	rs11900053	26787	A	G					
5	rs11542478	28938	C	A					
6	rs11897072	37808	A	G					
7	rs6713375	75793	G	A					
8	rs300789	79910	G	A					
9	rs7588043	94979	A	G					
10	rs6755432	96692	G	A					
11	rs300780	100819	A	G					
12	rs300773	105035	G	A					
13	rs10519439	108913	A	G					
14	rs435733	111222	G	A					
15	rs300711	129346	G	A					
16	rs437409	132137	G	A					
17	rs409954	139546	A	G					
18	rs408209	143279	C	A					
19	rs409572	143302	A	G					
20	rs10495480	160616	A	C					
21	rs300692	171415	G	A					
22	rs13033859	174390	G	A					
23	rs6744829	178223	G	A					
24	rs440617	179818	G	A					
25	rs415660	179972	G	A					
26	rs384526	183882	A	G					
27	rs385909	189678	A	G					
28	rs434818	190477	G	A					
29	rs385272	196704	G	A					
30	rs2952790	202106	A	G					
31	rs428697	203302	G	A					
32	rs451144	204609	0	A					
33	rs2290911	214919	G	A					
34	rs3791221	216933	G	A					
35	rs6742278	216999	G	A					
36	rs10205435	220442	A	G					
37	rs1461028	228975	G	A					
38	rs10181051	232800	G	A					

Chromosome 2

Ready

Screen-shot of GWAS database detailing the SNPs genotyped on Chromosome 6, position of each SNP, and the minor and major allele possibilities at each locus.



Chromosome 6A.xlsx

Home | Insert | Page Layout | Formulas | Data | Review | View

Cut | Copy | Format Painter | Paste | Clipboard | Calibri | 11 | Font | Wrap Text | Merge & Center | Alignment | General

A1 | SNP

	A	B	C	D	E	F	G	H	I
1	SNP	Position	Minor allele	Major Allele					
2	rs2527002	32800	0	A					
3	rs1179242	34449	A	C					
4	rs4097469	35797	A	C					
5	rs2526998	35876	0	C					
6	rs2263812	37261	A	C					
7	rs3908676	41968	G	A					
8	rs2266191	64666	A	G					
9	rs2265305	64740	G	A					
10	rs9504641	64970	G	A					
11	rs4959515	110391	A	G					
12	rs6596796	120522	A	G					
13	rs1535053	126089	G	A					
14	rs2158556	127550	A	G					
15	rs6937355	128900	A	G					
16	rs6937363	128917	A	G					
17	rs2281351	132344	A	G					
18	rs17133610	132837	A	G					
19	rs10155688	133937	A	G					
20	rs1316662	134113	G	A					
21	rs6596815	135044	C	A					
22	rs9501714	135248	G	A					
23	rs12191783	137106	A	G					
24	rs6916246	137695	G	A					
25	rs6913464	139090	A	G					
26	rs9405480	144199	A	G					
27	rs9502941	144964	A	G					
28	rs9392299	148909	G	A					
29	rs1418708	150610	A	G					
30	rs1418707	150800	G	A					
31	rs1418706	150878	G	A					
32	rs5000996	150920	A	C					
33	rs4282439	150947	A	C					
34	rs11961507	151313	G	A					
35	rs9502959	151599	A	G					
36	rs17806289	153822	A	C					
37	rs17133789	154462	C	A					
38	rs17133787	154980	G	A					

Chromosome 6

Ready

Appendix V

Plate Design

	1	2	3	4	5	6	7	8	9	10	11	12
A	EPI5821	SIEPI 045 AJC	SIEPI 060 AJC	SIEPI 020 AJC	SIEPI 023 AJC	EPI9497	BLANK	EPI30169	BLANK	EPI34419	EPI5584	EPI5107
B	EPI9500	EPI4618	EPI5290	EPI5113	EPI5659	EPI5731	BLANK	EPI5911	EPI10106	EPI34446	EPI32099	EPI30454
C	BLANK	EPI4747	EPI10070	EPI30169	EPI5911	EPI9701	SIEPI 045 AJC	EPI9701	EPI17578	BLANK	EPI34572	BLANK
D	EPI10172	EPI5566	EPI4777	EPI5938	EPI10196	EPI9410	BLANK	BLANK	BLANK	EPI34530	EPI34212	EPI32015
E	EPI16975	EPI4753	EPI9662	EPI9710	EPI10106	EPI17578	SIEPI 020 AJC	EPI10172	EPI31132	EPI45499	EPI5362	EPI32507
F	EPI31117	EPI31132	BLANK	EPI5344	EPI30457	EPI10127	BLANK	EPI5566	EPI5344	EPI5653	EPI44863	EPI101368
G	EPI34464	EPI45574	EPI30955	EPI34560	EPI34434	EPI34425	SIEPI 023 AJC	EPI4777	EPI30457	EPI101395	EPI101122	EPI101167
H	EPI34419	EPI34446	EPI34530	EPI45499	EPI5653	EPI101395	EPI9497	EPI5938	BLANK	EPI121932	EPI4984	EPI100942
I	EPI121932	EPI121719	EPI31973	EPI32297	EPI32522	EPI34239	EPI4618	EPI10196	EPI10127	EPI121719	BLANK	EPI101368
J	EPI45466	EPI16972	EPI5584	EPI32099	EPI34572	EPI34212	BLANK	BLANK	EPI34464	EPI31973	EPI201462	EPI201462
K	EPI5362	EPI44863	EPI101122	EPI4984	EPI201462	EPG183957	EPI5113	BLANK	EPI45574	BLANK	EPG183957	EPI32060
L	EPI9479	EPI32060	EPI5800	EPI10058	EPI30499	EPI9791	BLANK	EPI16975	EPI30955	EPI32297	EPI9479	EPI34572
M	EPI5107	EPI30454	BLANK	EPI32015	EPI32507	EPI101368	BLANK	EPI4753	BLANK	EPI32522	EPI32060	BLANK
N	EPI101167	EPI100942	EPI5821	SIEPI 060 AJC	EPI5659	EPI10070	EPI5731	EPI9662	EPI34560	EPI34239	EPI5800	EPI34434
O	EPI9701	EPI9410	EPI31117	EPI10127	EPI45574	BLANK	EPI4747	EPI9710	EPI34434	EPI16972	EPI10058	EPI45499
P	EPI32507	EPI34446	EPI101395	EPI34239	EPI9479	EPI100942	BLANK	BLANK	EPI34425	BLANK	EPI9791	EPI34464

Plex 2 shown in grey

Appendix VI

SLC22A1:

Appendix Table 8: SLC22A1 one-way ANOVAs

SNP (Minor allele)	F-Statistic	Degrees of Freedom		p-value	p-value rank	FDR-corrected p-value
		Between Groups	Within Groups			
rs3798173 (A)	5.102	2	92	0.008	1	1.424
rs628031 (A)	4.567	2	93	0.013	4	0.579
rs456598 (A)	4.264	2	93	0.017	5	0.605
rs806383 (A)	3.393	2	93	0.038	12	0.564
rs2197296 (A)	2.646	2	93	0.076	21	0.644
rs1564348 (G)	2.306	2	93	0.105	31	0.603
rs3798167 (A)	1.679	2	93	0.192	50	0.684
rs622342 (C)	1.503	2	93	0.228	52	0.780
rs654993 (A)	1.494	2	93	0.230	54	0.758
rs3818678 (C)	0.940	2	93	0.394	80	0.877
rs683369 (C)	0.801	2	93	0.452	88	0.914
rs2297374 (A)	0.741	2	93	0.480	90	0.949
rs3101826 (A)	0.719	2	93	0.490	92	0.948
rs1443844 (G)	0.698	2	93	0.500	96	0.927
rs10455868 (T)	0.447	2	67	0.641	116	0.984
rs651164 (A)	0.308	2	93	0.735	137	0.955
rs3798164 (A)	0.231	2	91	0.794	147	0.961
rs10455864 (T)	0.168	2	67	0.846	157	0.959
rs6455682 (A)	0.165	2	92	0.848	158	0.955
rs4646275 (A)	0.155	2	93	0.856	160	0.952
rs619598 (T)	0.155	2	68	0.857	161	0.947
rs3798174 (A)	0.127	2	93	0.881	163	0.962
rs1871389 (A)	0.112	2	93	0.894	167	0.953
rs2083867 (G)	0.112	2	93	0.894	167	0.953
rs4646283 (G)	0.085	2	92	0.919	174	0.940
rs622591 (A)	0.062	2	93	0.940	176	0.951
rs6937722 (A)	0.039	2	93	0.962	177	0.967

Appendix Table 9: SLC22A1 Student's t-tests

Variable	t-statistic	Degrees of freedom	p-value	Mean difference	95% Confidence interval for mean difference		p-value rank	FDR-corrected p-value
					Lower	Upper		
rs461473 (A)	-2.575	67	0.012	-0.36837	-0.65394	-0.08280	3	0.712
rs34059508 (A)	-2.151	69	0.035	-0.55731	-1.07413	-0.04049	11	0.566
rs3777392 (A)	1.689	94	0.095	0.18918	-0.03325	0.41161	28	0.604
rs4646272 (C)	-0.799	94	0.427	-0.10934	-0.38120	0.16251	84	0.905
rs9457839 (A)	-0.680	94	0.498	-0.15965	-0.62559	0.30629	95	0.933
rs9456505 (A)	0.601	69	0.549	0.08618	-0.19965	0.37200	100	0.977
rs4646273 (A)	-0.470	94	0.640	-0.07564	-0.39548	0.24419	114	0.999
rs3798168 (A)	-0.470	94	0.640	-0.07564	-0.39548	0.24419	114	0.999
rs3798169 (C)	-0.427	93	0.671	-0.06845	-0.38707	0.25018	124	0.963
rs9457843 (A)	-0.334	94	0.739	-0.03737	-0.25930	0.18457	138	0.953
rs35956182 (A)	0.118	67	0.906	0.02519	-0.40073	0.45112	173	0.932

UGT1A4:

Appendix Table 10: UGT1A4 one-way ANOVAs

SNP (Minor allele)	F-Statistic	Degrees of Freedom		p-value	p-value rank	FDR-corrected p-value
		Between Groups	Within Groups			
rs10209214 (G)	2.880	2	93	0.061	17	0.639
rs6717546 (A)	2.642	2	93	0.077	22	0.623
rs6431631 (C)	2.385	2	93	0.098	29	0.602
rs6719561 (A)	2.323	2	93	0.104	30	0.617
rs8330 (C)	2.140	2	92	0.124	39	0.566
rs4148329 (G)	1.418	2	91	0.248	56	0.788
rs4148325 (A)	0.543	2	93	0.583	101	1.027
rs6742078 (A)	0.543	2	93	0.583	101	1.027
rs887829 (A)	0.543	2	93	0.583	101	1.027
rs3771341 (A)	0.463	2	93	0.631	109	1.030
rs4148324 (C)	0.455	2	92	0.636	110	1.029
rs2013018 (G)	0.438	2	93	0.647	117	0.984
rs3755321 (G)	0.438	2	93	0.647	117	0.984
rs4663945 (A)	0.438	2	93	0.647	117	0.984
rs2221198 (A)	0.427	2	93	0.654	120	0.970
rs4124874 (C)	0.417	2	92	0.660	122	0.963
rs2008595 (A)	0.388	2	93	0.679	125	0.967
rs3755319 (C)	0.388	2	93	0.679	125	0.967
rs4148326 (G)	0.388	2	93	0.679	125	0.967
rs4663327 (A)	0.387	2	92	0.680	129	0.938
rs2018985 (G)	0.382	2	93	0.684	130	0.937
rs4294999 (G)	0.294	2	92	0.746	139	0.955
rs4663963 (C)	0.294	2	92	0.746	139	0.955
rs6744284 (A)	0.274	2	93	0.761	141	0.961
rs4148328 (A)	0.225	2	93	0.799	148	0.961
rs3806591 (G)	0.174	2	68	0.840	156	0.958
rs4663335 (A)	0.137	2	93	0.872	162	0.958
rs1018124 (G)	0.111	2	93	0.895	169	0.943
rs4663968 (G)	0.111	2	93	0.895	169	0.943
rs11563251 (A)	0.102	2	93	0.903	171	0.940

Appendix Table 11: UGT1A4 Student's t-tests

Variable	t-statistic	Degrees of freedom	p-value	Mean difference	95% Confidence interval for mean difference		p-value rank	FDR-corrected p-value
					Lower	Upper		
rs17862884 (A)	-1.136	94	0.259	-0.26543	-0.72934	0.19849	60	0.768
rs6431633 (A)	0.627	67	0.533	0.07426	-0.16198	0.31049	98	0.968
rs2302538 (G)	-0.293	92	0.770	-0.03246	-0.25268	0.18775	144	0.952
rs11568318 (A)	-0.215	92	0.830	-0.02985	-0.30536	0.24566	154	0.959

SCN1A:

Appendix Table 12: SCN1A one-way ANOVAs

SNP (Minor allele)	F-Statistic	Degrees of Freedom		p-value	p-value rank	FDR-corrected p-value
		Between Groups	Within Groups			
rs922224 (G)	2.889	2	93	0.061	17	0.639
rs478389 (C)	1.952	2	67	0.150	42	0.636
rs10176603 (A)	1.498	2	93	0.229	53	0.769
rs484926 (G)	1.380	2	93	0.257	58	0.789
rs6731591 (G)	1.380	2	93	0.257	58	0.789
rs13397210 (A)	1.324	2	93	0.271	62	0.778
rs1841548 (G)	1.324	2	93	0.271	62	0.778
rs10183551 (G)	1.253	2	93	0.290	64	0.807
rs7607455 (G)	1.143	2	93	0.323	68	0.846
rs16851382 (A)	1.134	2	93	0.326	71	0.817
rs11692675 (G)	0.995	2	93	0.374	74	0.900
rs10497275 (G)	0.815	2	93	0.446	86	0.923
rs10497276 (A)	0.815	2	93	0.446	86	0.923
rs12998913 (G)	0.709	2	93	0.495	93	0.947
rs536744 (A)	0.539	2	92	0.585	104	1.001
rs6722462 (G)	0.498	2	93	0.609	106	1.023
rs577306 (A)	0.485	2	93	0.617	107	1.026
rs1020853 (A)	0.476	2	93	0.623	108	1.027
rs13421166 (C)	0.455	2	93	0.636	110	1.029
rs1834840 (G)	0.455	2	93	0.636	110	1.029
rs4667867 (A)	0.455	2	93	0.636	110	1.029
rs1381109 (A)	0.388	2	93	0.679	125	0.967
rs1824549 (C)	0.359	2	93	0.700	131	0.951
rs545331 (A)	0.268	2	93	0.766	142	0.960

Appendix Table 13: SCN1A Student's t-tests

Variable	t-statistic	Degrees of freedom	p-value	Mean difference	95% Confidence interval for mean difference		p-value rank	FDR-corrected p-value
					Lower	Upper		
rs17791817 (T)	1.568	68	0.122	0.21626	-0.05897	0.49149	38	0.571
rs12151636 (C)	1.125	69	0.264	0.23459	-0.18137	0.65054	61	0.770
rs7606193 (G)	0.236	94	0.814	0.07769	-0.57561	0.73099	151	0.960
rs9678982 (G)	0.236	94	0.814	0.07769	-0.57561	0.73099	151	0.960
rs6735544 (G)	-0.148	94	0.883	-0.06848	-0.98767	0.85071	164	0.958

SCN2A:

Appendix Table 14: SCN2A one-way ANOVAs

SNP (Minor allele)	F-Statistic	Degrees of Freedom		p-value	p-value rank	FDR-corrected p-value
		Between Groups	Within Groups			
rs12993173 (C)	4.256	2	92	0.017	5	0.605
rs1439805 (C)	4.184	2	93	0.018	7	0.458
rs9287856 (A)	3.724	2	93	0.028	8	0.623
rs9287857 (G)	3.724	2	93	0.028	8	0.623
rs1866603 (G)	3.629	2	93	0.030	10	0.534
rs17184707 (A)	3.384	2	93	0.039	13	0.534
rs10174400 (A)	3.108	2	91	0.049	14	0.623
rs10182570 (C)	3.040	2	93	0.053	15	0.629
rs10184275 (G)	2.894	2	93	0.060	16	0.668
rs7607897 (A)	2.861	2	92	0.062	19	0.581
rs10207911 (G)	2.760	2	93	0.068	20	0.605
rs353119 (A)	2.622	2	93	0.078	23	0.604
rs353121 (G)	2.622	2	93	0.078	23	0.604
rs1439804 (A)	2.453	2	93	0.092	27	0.607
rs353111 (C)	2.290	2	93	0.107	33	0.577
rs2119067 (G)	2.208	2	93	0.116	37	0.558
rs6741147 (T)	2.003	2	68	0.143	40	0.636
rs353123 (A)	1.912	2	93	0.154	43	0.637
rs353139 (G)	1.734	2	93	0.182	49	0.661
rs13025009 (G)	1.676	2	93	0.193	51	0.674
rs7600082 (T)	1.429	2	68	0.247	55	0.799
rs3769941 (G)	1.411	2	93	0.249	57	0.778
rs2304010 (A)	1.142	2	93	0.324	69	0.836
rs2892961 (A)	1.142	2	93	0.324	69	0.836
rs1965757 (G)	1.089	2	93	0.341	73	0.831
rs10497259 (A)	0.963	2	93	0.385	75	0.914
rs1838847 (A)	0.963	2	93	0.385	75	0.914
rs3769931 (G)	0.966	2	93	0.385	75	0.914
rs7573433 (G)	0.909	2	68	0.408	81	0.897
rs4667481 (G)	0.858	2	93	0.427	84	0.905
rs13387970 (A)	0.793	2	93	0.456	89	0.912
rs12469667 (A)	0.689	2	92	0.505	97	0.927
rs6718960 (G)	0.619	2	92	0.541	99	0.973
rs10497258 (G)	0.533	2	93	0.589	105	0.998
rs1368234 (C)	0.418	2	68	0.660	122	0.963
rs764660 (C)	0.333	2	93	0.718	133	0.961
rs10203729 (C)	0.319	2	93	0.728	134	0.967
rs10204322 (A)	0.319	2	93	0.728	134	0.967
rs2116658 (A)	0.319	2	93	0.728	134	0.967
rs2116659 (A)	0.268	2	93	0.766	142	0.960
rs10930162 (G)	0.261	2	93	0.771	145	0.946
rs4480996 (G)	0.261	2	93	0.771	145	0.946
rs17183814 (A)	0.115	2	68	0.892	166	0.956
rs6740895 (G)	0.100	2	93	0.905	172	0.937
rs2060199 (A)	0.016	2	68	0.984	178	0.984

Appendix Table 15: SCN2A Student's t-tests

Variable	t-statistic	Degrees of freedom	p-value	Mean difference	95% Confidence interval for mean difference		p-value rank	FDR-corrected p-value
					Lower	Upper		
rs2304016 (G)	1.400	93	0.165	0.63869	-0.26702	1.54339	44	0.668
rs13424762 (G)	1.379	94	0.171	0.63196	-0.27818	1.54211	46	0.662
rs6714217 (G)	1.379	94	0.171	0.63196	-0.27818	1.54211	46	0.662
rs10803795 (G)	0.242	94	0.810	0.04366	-0.31522	0.40253	150	0.961
rs17242693 (A)	-0.189	94	0.850	-0.02913	-0.33461	0.27636	159	0.952
rs4303727 (G)	-0.146	94	0.884	-0.01456	-0.21235	0.18341	165	0.954

SCN3A:

Appendix Table 16: SCN3A one-way ANOVAs

SNP (Minor allele)	F-Statistic	Degrees of Freedom		p-value	p-value rank	FDR-corrected p-value
		Between Groups	Within Groups			
rs11903851 (A)	5.018	2	93	0.009	2	0.801
rs1899013 (G)	2.555	2	93	0.083	25	0.591
rs1982208 (C)	2.555	2	93	0.083	25	0.591
rs1946892 (A)	2.283	2	68	0.110	34	0.576
rs11686777 (A)	2.218	2	93	0.115	35	0.585
rs1347992 (G)	2.218	2	93	0.115	35	0.585
rs17829596 (G)	1.989	2	93	0.143	40	0.636
rs1439806 (A)	1.836	2	93	0.165	44	0.668
rs11677254 (G)	1.814	2	68	0.171	46	0.662
rs13011371 (G)	1.247	2	91	0.292	65	0.800
rs7574918 (A)	1.204	2	93	0.304	66	0.820
rs4667792 (G)	1.151	2	92	0.321	67	0.853
rs13015737 (A)	0.945	2	93	0.392	78	0.895
rs4667786 (G)	0.943	2	93	0.393	79	0.885
rs10497257 (G)	0.867	2	93	0.423	82	0.918
rs4145346 (G)	0.867	2	93	0.423	82	0.918
rs6727857 (G)	0.706	2	93	0.496	94	0.939
rs6755352 (G)	0.426	2	93	0.654	120	0.970
rs2165208 (A)	0.219	2	93	0.803	149	0.959
rs2043932 (A)	0.182	2	93	0.834	155	0.958
rs2390165 (A)	0.075	2	93	0.928	175	0.944

Appendix VII

SLC22A1

rs806383

ANOVA identified a significant association between maintenance dose of LTG and genotype at this SNP locus (p=0.038, uncorrected). This intronic SNP is found in *SLC22A1* and did not tag any additional SNPs. The homozygous-minor allele group (AA) appeared to require a higher maintenance dose of LTG than the homozygous-major allele group (GG) (*Appendix Table 17* and *Appendix Table 18*). Comparison with the heterozygous group (GA) failed to reach statistical significance but there was modest evidence of an allele-specific effect on dose, with carriers of the A-allele requiring higher maintenance doses of LTG (*Figure 10*).

Appendix Table 17: Back transformed dose data for rs806383

Genotype	Number of Participants	Mean Dose (mg/day)	Minimum Dose (mg/day)	Maximum Dose (mg/day)
GG	46	176	50	500
GA	39	183	100	500
AA	11	259	150	675
Total	96	187	50	675

Appendix Table 18: Back-transformed dose data for rs806383 - between groups comparison

Compared Genotypes		Ratio of Geometric Means	95% Confidence Interval for Ratio of Geometric Means		Significance
			Lower Boundary	Upper Boundary	
GG	GA	0.959	0.757	1.214	0.962
	AA	0.679*	0.471	0.978	0.034
GA	GG	1.043	0.823	1.321	0.962
	AA	0.708	0.489	1.026	0.076
AA	GG	1.473*	1.023	2.121	0.034
	GA	1.412	0.975	2.046	0.076

* The mean difference is significant at the 0.05 level.
Number of participants: GG = 46, GA = 39, AA = 11, Total = 96

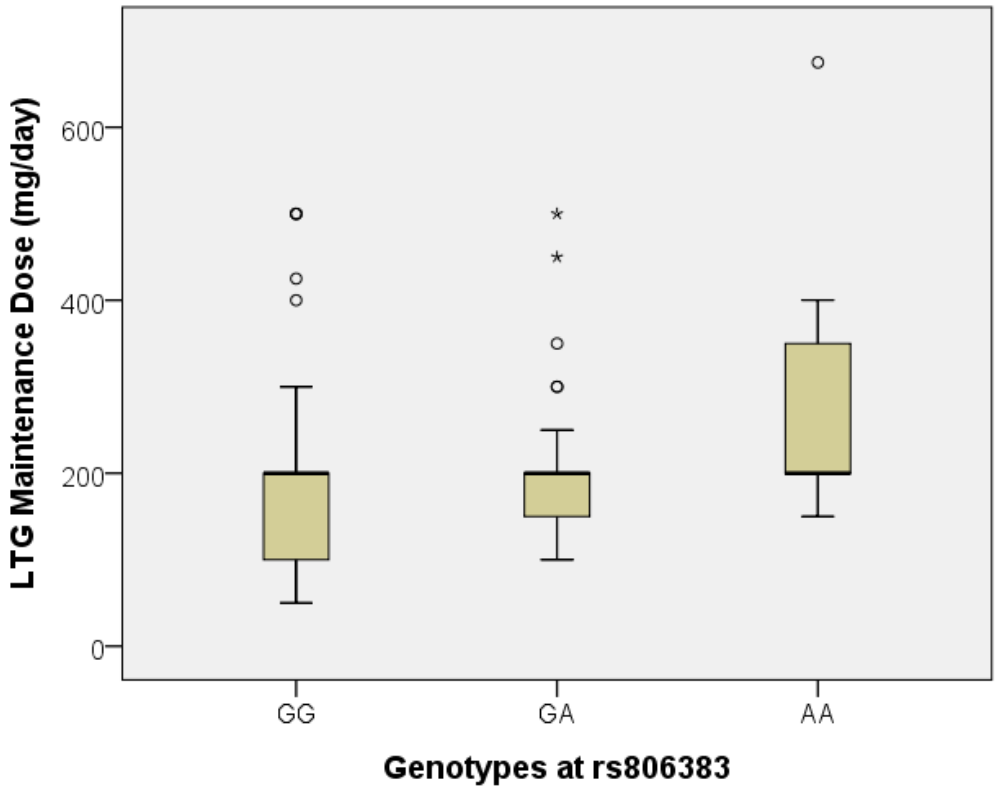


Figure 10: Dose distribution between rs806383 genotypes
Top box boundary = third quartile, lower boundary = first quartile, and bold line = median. Whiskers = values within 1.5 x inter-quartile range (IQR) outside the box. Circles = values 1.5 to 3 x IQR outside the box. Stars = values more than 3 x IQR outside the box.
Number of participants: GG = 46, GA = 39, AA = 11, Total = 96

rs34059508

Student's t-test identified a significant association between maintenance dose and genotype at this SNP locus (p=0.035, uncorrected). This is a coding non-synonymous SNP in *SLC22A1* that results in a G465R amino acid substitution in the OCT1 protein. It did not tag any additional SNPs. The heterozygous group (GA, geometric mean = 311mg/day, n=3) appeared to require a higher maintenance dose of LTG than the homozygous-major allele group (GG, geometric mean = 178mg/day, n=68) (*Appendix Table 19*). Again, with only two genotypes at this locus, it was not possible to infer allele-specific effects.

Appendix Table 19: Back-transformed dose data for rs34059508 – between groups comparison

Variable	t-statistic	Degrees of freedom	Ratio of Geometric means	95% CI of Ratio of Geometric Means		p-value	p-value rank	FDR-corrected p-value
				Lower	Upper			
rs34059508 (A)	-2.151	69	0.573	0.342	0.960	0.035	3	0.712

Number of participants: GG = 68, GA = 3, Total = 71

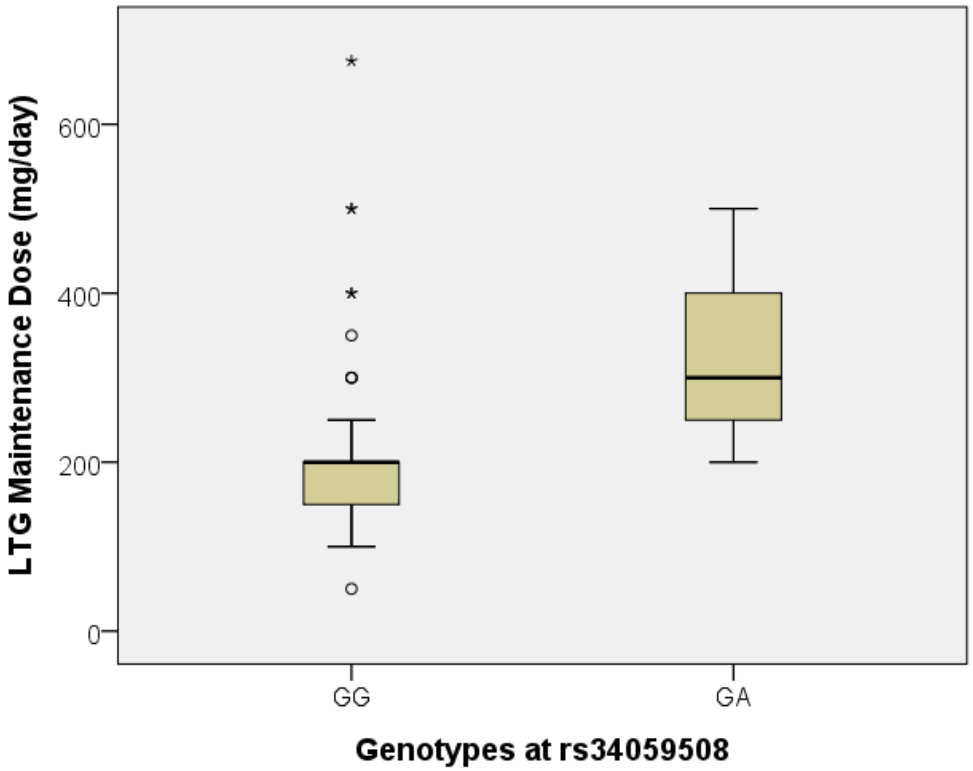


Figure 11: Dose distribution between rs34059508 genotypes
Top box boundary = third quartile, lower boundary = first quartile, and bold line = median. Whiskers = values within 1.5 x inter-quartile range (IQR) outside the box. Circles = values 1.5 to 3 x IQR outside the box. Stars = values more than 3 x IQR outside the box.
Number of participants: GG = 68, GA = 3, Total = 71

SCN2A

rs9287856

ANOVA identified a significant association between maintenance dose of LTG and genotype at this SNP locus (p=0.028, uncorrected). This intronic SNP is found in *SCN2A* and tagged an additional four SNPs (*Appendix Table 6*). A statistically significant difference between the individual genotype groups could not be identified in post-hoc tests (*Appendix Table 20* and *Appendix Table 21*). There was some evidence of an allele-specific effect on dose, with carriers of the A-allele requiring higher maintenance doses of LTG (*Figure 12*).

Appendix Table 20: Back-transformed dose data for rs9287856

Genotype	Number of Participants	Mean Dose (mg/day)	Minimum Dose (mg/day)	Maximum Dose (mg/day)
GG	66	174	50	500
GA	26	207	100	675
AA	4	300	200	450
Total	96	187	50	675

Appendix Table 21: Back-transformed dose data for rs9287856 – between groups comparison

Compared Genotypes		Ratio of Geometric Means	95% Confidence Interval for Ratio of Geometric Means		Significance
			Lower Boundary	Upper Boundary	
GG	GA	0.843	0.644	1.104	0.284
	AA	0.582	0.304	1.112	0.083
GA	GG	1.186	0.906	1.554	0.284
	AA	0.690	0.375	1.270	0.218
AA	GG	1.720	0.899	3.289	0.083
	GA	1.450	0.788	2.668	0.218

Number of participants: GG = 66, GA = 26, AA = 4, Total = 96

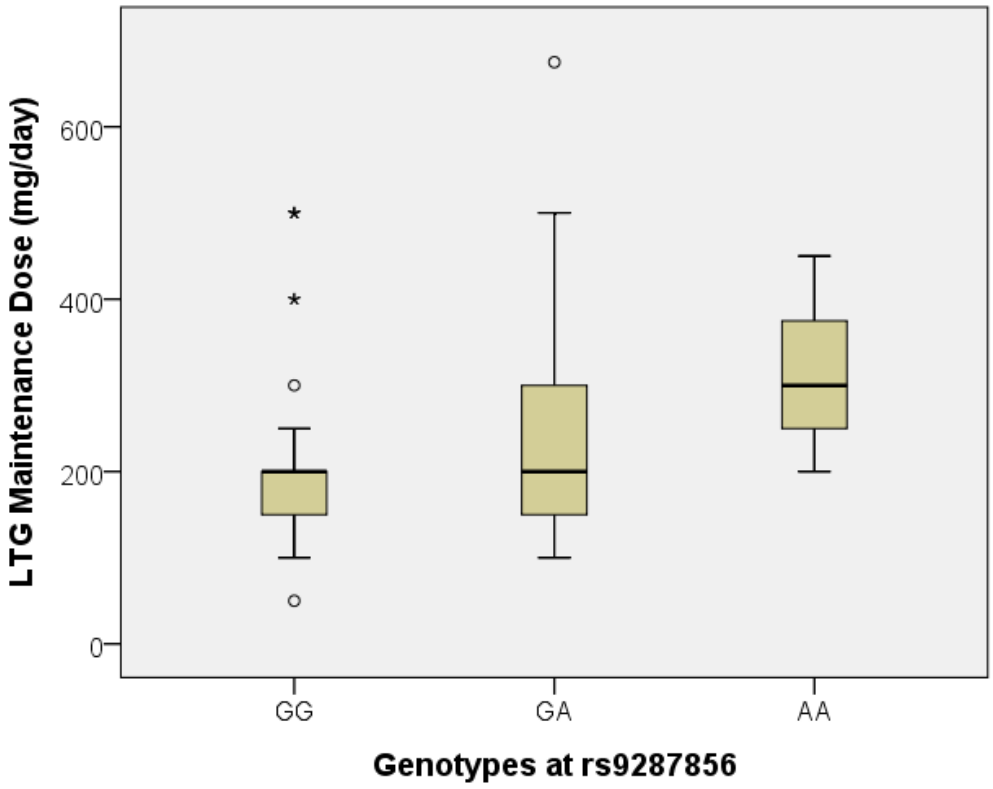


Figure 12: Dose distribution between rs9287856 genotypes
Top box boundary = third quartile, lower boundary = first quartile, and bold line = median. Whiskers = values within 1.5 x inter-quartile range (IQR) outside the box. Circles = values 1.5 to 3 x IQR outside the box. Stars = values more than 3 x IQR outside the box.
Number of participants: GG = 66, GA = 26, AA = 4, Total = 96

rs9287857

ANOVA identified a significant association between maintenance dose of LTG and genotype at this SNP locus (p=0.028, uncorrected). This intronic SNP is found in *SCN2A* (*Appendix Table 6*) and tagged an additional four SNPs. A statistically significant difference between the individual genotype groups could not be identified in post-hoc tests (*Appendix Table 22* and *Appendix Table 23*). There was some evidence of an allele-specific effect on dose, with carriers of the A-allele requiring higher maintenance doses of LTG (*Figure 13*).

Appendix Table 22: Back-transformed dose data for rs9287857

Genotype	Number of Participants	Mean Dose (mg/day)	Minimum Dose (mg/day)	Maximum Dose (mg/day)
AA	66	174	50	500
AG	26	207	100	675
GG	4	300	200	450
Total	96	187	50	675

Appendix Table 23: Back-transformed dose data for rs9287857 – between groups comparison

Compared Genotypes		Ratio of Geometric Means	95% Confidence Interval for Ratio of Geometric Means		Significance
			Lower Boundary	Upper Boundary	
AA	AG	0.843	0.656	1.083	0.272
	GG	0.582	0.333	1.016	0.059
AG	AA	1.186	0.923	1.524	0.272
	GG	0.690	0.386	1.234	0.326
GG	AA	1.720	0.985	3.003	0.059
	AG	1.450	0.810	2.593	0.326

Number of participants: AA = 66, AG = 26, GG = 4, Total = 96

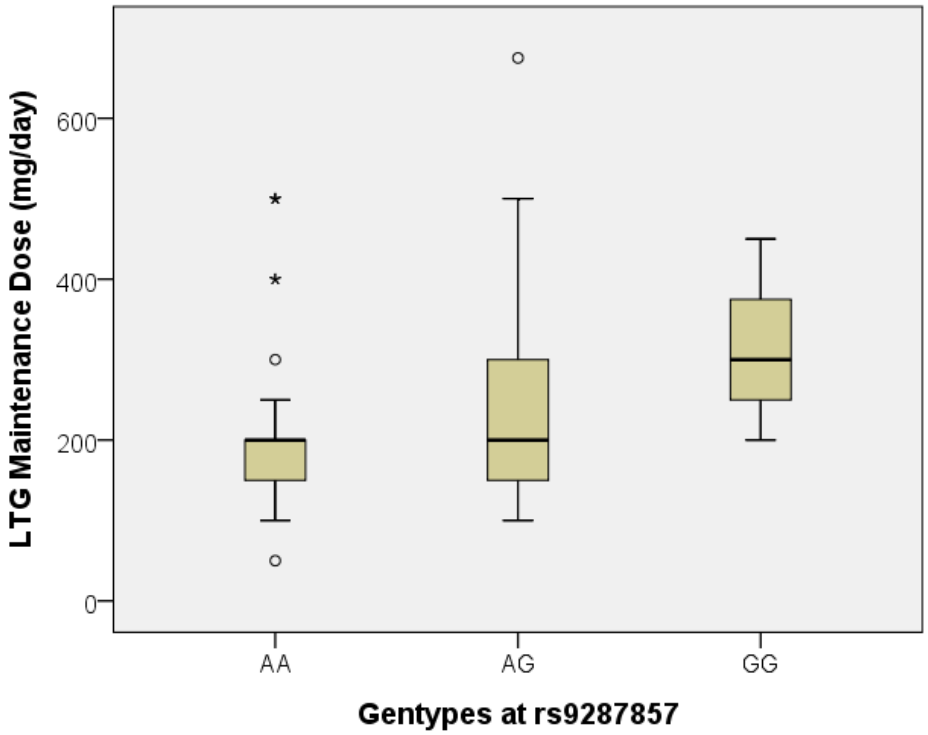


Figure 13: Dose distribution between rs9287857 genotypes
Top box boundary = third quartile, lower boundary = first quartile, and bold line = median. Whiskers = values within 1.5 x inter-quartile range (IQR) outside the box. Circles = values 1.5 to 3 x IQR outside the box. Stars = values more than 3 x IQR outside the box.
Number of participants: AA = 66, AG = 26, GG = 4, Total = 96

rs1866603

ANOVA identified a significant association between maintenance dose of LTG and genotype at this SNP locus (p=0.030, uncorrected). This intronic SNP is found in *SCN2A* and tagged an additional four SNPs (*Appendix Table 6*). The homozygous-minor allele group (GG) appeared to require a higher maintenance dose of LTG than the heterozygous group (AG) (*Appendix Table 24* and *Appendix Table 25*). Comparison with the homozygous-major allele group (AA) failed to reach statistical significance and there no was no evidence of an allele-specific effect on dose (*Figure 14*).

Appendix Table 24: Back-transformed dose data for rs1866603

Genotype	Number of Participants	Mean Dose (mg/day)	Minimum Dose (mg/day)	Maximum Dose (mg/day)
AA	22	187	100	500
AG	55	172	50	675
GG	19	237	150	500
Total	96	187	50	675

Appendix Table 25: Back-transformed dose data for rs1866603 - between groups comparison

Compared Genotypes		Ratio of Geometric Means	95% Confidence Interval for Ratio of Geometric Means		Significance
			Lower Boundary	Upper Boundary	
AA	AG	1.087	0.827	1.429	0.839
	GG	0.790	0.563	1.109	0.256
AG	AA	0.920	0.700	1.209	0.839
	GG	0.726*	0.544	0.969	0.025
GG	AA	1.266	0.902	1.778	0.256
	AG	1.377*	1.032	1.837	0.025

*The mean difference is significant at the 0.05 level.
Number of participants: AA = 22, AG = 55, GG = 19, Total = 96

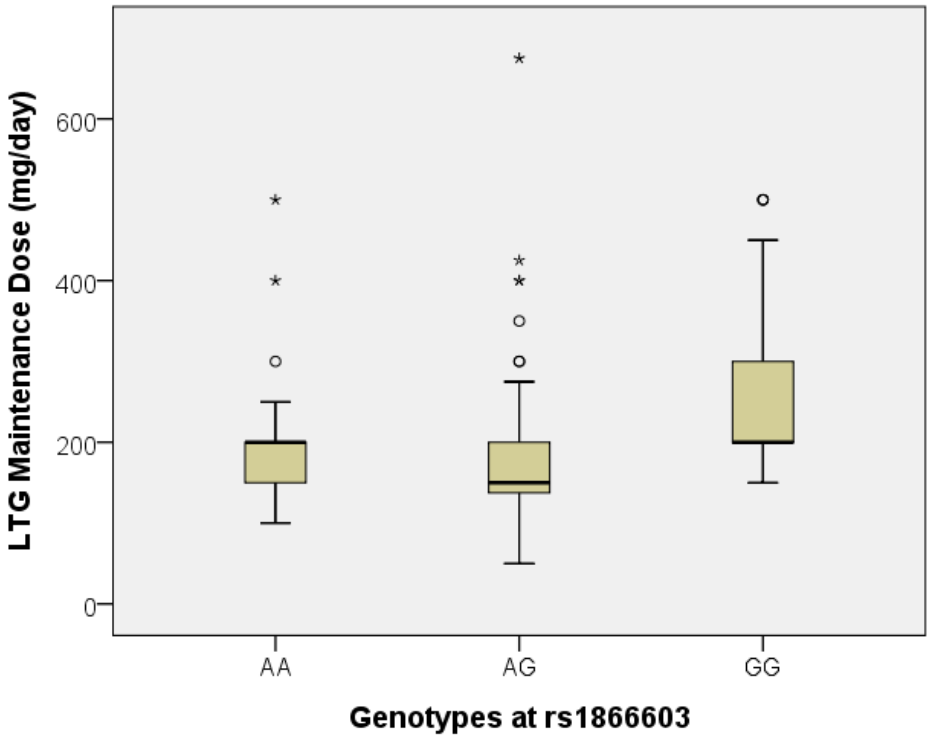


Figure 14: Dose distribution between rs1866603 genotypes
Top box boundary = third quartile, lower boundary = first quartile, and bold line = median. Whiskers = values within 1.5 x inter-quartile range (IQR) outside the box. Circles = values 1.5 to 3 x IQR outside the box. Stars = values more than 3 x IQR outside the box.
Number of participants: AA = 22, AG = 55, GG = 19, Total = 96

rs17184707

ANOVA identified a significant association between maintenance dose of LTG and genotype at this SNP locus (p=0.039, uncorrected). This intronic SNP is found in *SCN2A* and did not tag any additional SNPs. A statistically significant difference between the individual genotype groups could not be identified in post-hoc tests (*Appendix Table 26* and *Appendix Table 27*). There was some evidence of an allele-specific effect on dose, with carriers of the A-allele requiring higher maintenance doses of LTG (*Figure 15*).

Appendix Table 26: Back transformed dose data for rs17184707

Genotype	Number of Participants	Mean Dose (mg/day)	Minimum Dose (mg/day)	Maximum Dose (mg/day)
GG	53	173	50	500
GA	39	198	100	500
AA	4	300	150	675
Total	96	187	50	675

Appendix Table 27: Back-transformed dose data for rs17184707 - between groups comparison

Compared Genotypes		Ratio of Geometric Means	95% Confidence Interval for Ratio of Geometric Means		Significance
			Lower Boundary	Upper Boundary	
GG	GA	0.874	0.695	1.100	0.401
	AA	0.577	0.328	1.013	0.057
GA	GG	1.144	0.909	1.438	0.401
	AA	0.659	0.373	1.166	0.218
AA	GG	1.734	0.987	3.047	0.057
	GA	1.517	0.857	2.683	0.218

Number of participants: GG = 53, GA = 39, AA = 4, Total = 96

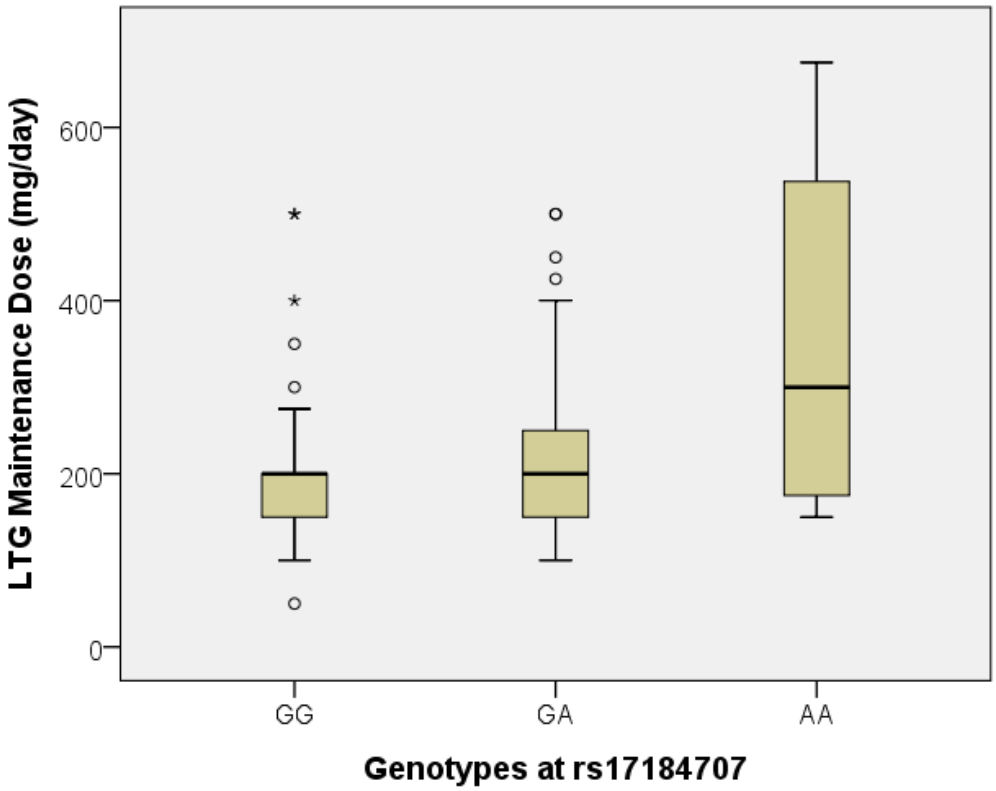


Figure 15: Dose distribution between rs17184707 genotypes
Top box boundary = third quartile, lower boundary = first quartile, and bold line = median. Whiskers = values within 1.5 x inter-quartile range (IQR) outside the box. Circles = values 1.5 to 3 x IQR outside the box. Stars = values more than 3 x IQR outside the box.
Number of participants: GG = 53, GA = 39, AA = 4, Total = 96

rs10174400

ANOVA identified a significant association between maintenance dose of LTG and genotype at this SNP locus (p=0.049, uncorrected). This intronic SNP is found in *SCN2A* and tagged an additional ten SNPs (*Appendix Table 6*). The homozygous-minor allele group (AA) appeared to require a higher maintenance dose of LTG than the homozygous-major allele group (GG) (*Appendix Table 28* and *Appendix Table 29*). Comparison with the heterozygous group (GA) failed to reach statistical significance but there was modest evidence of an allele-specific effect on dose, with carriers of the G-allele requiring higher maintenance doses of LTG (*Figure 16*).

Appendix Table 28: Back transformed dose data for rs10174400

Genotypes	Number of Participants	Mean Dose (mg/day)	Minimum Dose (mg/day)	Maximum Dose (mg/day)
GG	43	170	50	500
GA	44	189	100	675
AA	7	265	150	500
Total	94	185	50	675

Appendix Table 29: Back-transformed dose data for rs10174400 - between groups comparison

Compared Genotypes		Ratio of Geometric Means	95% Confidence Interval for Ratio of Geometric Means		Significance
			Lower Boundary	Upper Boundary	
GG	GA	0.901	0.716	1.134	0.616
	AA	0.643*	0.415	0.997	0.048
GA	GG	1.110	0.882	1.397	0.616
	AA	0.714	0.461	1.105	0.179
AA	GG	1.554*	1.003	2.407	0.048
	GA	1.401	0.905	2.168	0.179

*. The mean difference is significant at the 0.05 level.
Number of participants: GG = 43, GA = 44, AA = 7, Total = 94

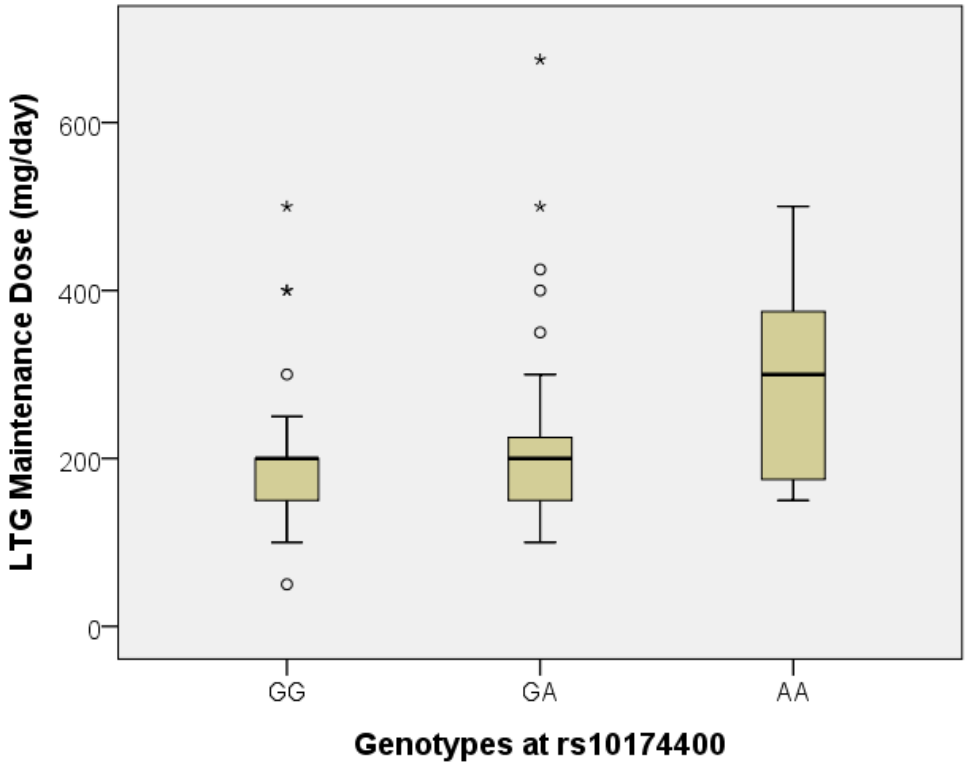


Figure 16: Dose distribution between rs10174400 genotypes

Top box boundary = third quartile, lower boundary = first quartile, and bold line = median. Whiskers = values within 1.5 x inter-quartile range (IQR) outside the box. Circles = values 1.5 to 3 x IQR outside the box. Stars = values more than 3 x IQR outside the box.
Number of participants: GG = 43, GA = 44, AA = 7, Total = 94