Health Economics of Pharmacogenetic Testing

Dong Di
BSc. (Hons), National University of Singapore

Dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy
Duke-NUS Graduate Medical School
National University of Singapore
August 2015

ABSTRACT

Some commonly used medications can cause a life-threatening adverse drug reaction named Stevens-Johnson syndrome (SJS). Though the genetic risk factors have been well established, adoption of pharmacogenetic testing to reduce SJS risk in clinical care has been slow. I conduct health economic evaluations to inform clinical and regulatory decision making on whether pharmacogenetic testing should be done to reduce the risk of SJS in Singapore.

I first assessed the cost-effectiveness of two pharmacogenetic tests from a health system perspective, and found that genotyping epilepsy patients for HLA-B*1502 prior to carbamazepine treatment was highly cost-effective, whereas genotyping chronic gout patients for HLA-B*5801 prior to allopurinol treatment was not cost-effective in Singapore. I then measured patients' preferences for pharmacogenetic testing in chronic gout treatment. Results suggested that risk of SJS, cost of genotyping, cost of gout treatment, doctor's recommendation and choice of peer patients all influenced patients' decisions. I predicted the test uptake rate to be around 65% if the test was to be offered. These results would contribute to the change in regulatory recommendation on genetic testing and the clinical practice in Singapore.

Health Economics of Pharmacogenetic Testing

by

DONG Di

PhD Program in Integrated Biology and Medicine
Duke-NUS Graduate Medical School

Date: Oct. 07, 2014

Approved:

Eric Finkelstein, Thesis Advisor

Angelique Chan

Cynthia Sung

David Matchar

Marcel Bilger

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy
Duke-NUS Graduate Medical School
National University Singapore
October 2015

Health Economics of Pharmacogenetic Testing

by

DONG Di

PhD Program in Integrated Biology and Medicine
Duke-NUS Graduate Medical School

Date: Oct. 07, 2015

Approved:

Eric Finkelstein, Thesis Advisor

Angelique Chan

Cynthia Sung

David Matchar

Marcel Bilger

Dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy
Duke-NUS Graduate Medical School
National University of Singapore
October 2015

Abstract

Burgeoning genetic research is enabling the personalization of medical treatments based on patients' individual genetic profile. One potential application that is likely to make significant impact in transforming patient care in the near future is pharmacogenetics, where patients' genetic traits can predict their responses to drugs, and appropriate treatment that maximizes effectiveness and minimizes side effects can be selected based on genetic testing results. The pharmacogenetics of several life-threatening adverse drug reactions have been well established, however, the adoption of pharmacogenetic testing in clinical care has been slow. The main reason being that the clinical utility, adverse consequences and economic value of genetic testing are unclear. In many cases, the decision of technology adoption often involves tradeoffs between the above factors, which is difficult without a systematic evaluation of various factors all together and a commonly accepted standard.

The objective of the thesis is to conduct health economic evaluations to generate evidence to inform clinical and regulatory decision making on whether pharmacogenetic testing should be routinely conducted in order to reduce the risk of a life-threatening adverse drug reaction named Stevens-Johnson syndrome (SJS) and Toxic Epidermal Necrolysis (TEN) in the context of Singapore. To assess the value of pharmacogenetic testing in Singapore, two health economics evaluation methods are employed: cost-effectiveness analysis (CEA) and discrete choice experiment (DCE).

Cost-effectiveness analysis adopts a health system perspective to estimate the long-term cost and effectiveness related to pharmacogenetic testing in the population, with consideration of test accuracy, predictive power of test results, population risk allele prevalence, efficacy of various drugs, side effects of various drugs and their sequelae, patients' quality of life, survival, and treatment costs. Cost-effectiveness evaluates the incremental effectiveness and incremental costs

associated with genetic testing relative to the status quo treatment strategy to reveal the incremental value of genetic testing. A threshold of cost-effectiveness that reflects the societal willingness-to-pay can then be applied to judge whether genetic testing is cost-effective in the health system. Cost-effectiveness analysis favors technologies that achieve high effectiveness at low costs at the health system level, and is useful for policy makers to make resource allocation between various healthcare needs, and make efficient use of public healthcare resources. However, cost-effectiveness does not speak to what individuals would do or should do. To understand individual level decision making, discrete choice experiment can be used to elicit patients' preferences and willingness-to-pay for genetic test to reduce risk of adverse drug reactions.

The thesis consists of six chapters. Chapter 1 lays the general background of the thesis. It outlines the advancement and challenges in the adoption of pharmacogenetic testing in clinical practice, and describes how health economics evaluations can inform the genetic testing decision making. Chapter 2 assesses the cost-effectiveness of HLA-B*1502 testing prior to carbamazepine treatment for epilepsy patients from a health system perspective to inform clinical and regulatory policy making in Singapore. Results suggest that compared with the status quo strategy of providing carbamazepine to all patients without genotyping, genotyping and targeted treatment is highly cost-effective for Chinese and Malays, but not Indians in Singapore. The study, together with other related studies, has led to a regulatory recommendation of HLA-B*1502 testing prior to carbamazepine initiation among epilepsy patients, and subsequent adoption in clinical practice. These changes as well as intended outcomes and unintended consequences are briefly reviewed at the end of Chapter 2. Chapter 3 assesses the cost-effectiveness of HLA-B*5801 testing, and other risk-mitigation strategies for allopurinol among chronic gout patients. Results suggest that HLA-B*5801 testing-guided treatment selection, in which allopurinol is avoided in test positive patients, is not cost-effective at the

population level, as the limited choice of alternative drugs to allopurinol will result in poorer serum urate control and worse gout treatment outcomes in some patients. On the other hand, a combination of genetic testing and a safety monitoring program is favored from the cost-effectiveness perspective under certain circumstances. Chapter 4 adopts a different perspective to review the literature on patients' preferences for pharmacogenetic testing, and motivate the study in Chapter 5. Chapter 5 describes a discrete choice experiment to quantitatively measure patients' preferences for genetic testing prior to initiating allopurinol in chronic gout treatment. Empirical data suggest that there is significant heterogeneity in patients' preferences. A group of patients are risk averse, and have high willingness-to-pay for genetic testing even though the test is not perfectly predictive and treatment costs are significant higher. On the contrary, other patients are cost conscious, and consider cost containment to be more important than risk reduction. The preferences of both groups of patients are quantified in Chapter 5. In addition, this study also revealed the strong impact of doctor's recommendation and herd effect on patients' decision making. The thesis is concluded in Chapter 6, with recommendations for future research.

Dedication

Dedicated to my parents Dong Jian and Wang Yanping, my husband Wang Kewei (Andrew), my parents-in-law Wang Jiaguo and Li Xiang, and my cat Nemo.

Acknowledgement

I would like to express my deepest gratitude to my mentor Professor Eric Finkelstein. Eric accepted me as his student when I had little background in economics, and planned my curriculum to equip me with skills to conduct health services research. Eric trained me to think as an economist. His emphasis on critical thinking has motivated me to reframe my specific research questions and hypotheses, and develop more convincing arguments. I would also like to thank him for always being available and willing to provide insightful suggestions. Eric also introduced me to new opportunities and broadened my horizon. I would like to thank A/Professor Cynthia Sung for introducing me to the field of pharmacogenetics and initiated the excellent collaboration between my team at Duke-NUS and the Health Sciences Authority. Cynthia's knowledge and passion in pharmacogenetics and pharmacovigilance has led us to develop research studies with significant policy impact in Singapore. I am grateful to Dr Marcel Bilger as he is always very patient and welcoming when I approach him with econometrics questions. He is always capable of recommending systematic and creative solutions. I appreciate his help and inspiration. I would like to thank Professor David Matchar, and A/Professor Angelique Chan for giving valuable suggestions to my research from different perspectives, and providing constant encouragement along the way. I want to thank Dr Semra Ozdemir for her help in my DCE project and offering recommendations on experimental design, survey interview and data analyses. I would like to thank Dr Bee Yong Mong and Dr Sue-Ann Toh for their support in my data collection, without which the DCE survey could not be completed. I would like to thank Dr Teng Gim

Gee and Ms Tan-Koi Wei Chuen, who I enjoyed working with on the cost-effectiveness analysis. Discussion with them has inspired me to come up with new ideas to improve the study. I would like to thank Mr Chay Junxing for his encouragement and help with my DCE project and data analysis in general. I would also like to thank all my colleagues at Duke-NUS Health Services and Systems Research and Lien Centre for Palliative Care for supporting me through my PhD journey.

I would like to confer my special thanks to A/Prof Silke Vogel, A/Prof David Silver, Ms Megahwani Catherine and Ms Rosilah Binte Mohd Mustapha, who have taken good care of me and the PhD cohort. I would like to thank all my classmates in the IBM 2014 cohort for their company and inspiration. Last but not least, my gratitude goes to Duke-NUS for offering scholarship to support my study and living, without which I would not be able to pursuit my PhD.

Table of Contents

List of Tables

List of Figures

Chapter 1 . Challenges in adoption of Pharmacogenetic testing and the role of health economics analyses

1.1 Chapter summary

This chapter is an introduction to the general background of the thesis. It outlines the advancement of pharmacogenetics, particularly the application of pharmacogenetics in preventing severe adverse drug reactions. Then the current status of adoption and challenges are discussed, followed by an introduction of two health economics evaluation methods that can facilitate the decision of genetic testing adoption.

1.2 Trend in personalized medicine and pharmacogenetics

Burgeoning genetic research has started to transform medicine and enable the personalization of medical treatment based on patients' individual genetic traits. Genetic testing is the process of identifying individual genetic variability, for a broad spectrum of medical applications using various testing methods.[1] Diagnostic genetic testing can be used to confirm suspected diagnosis. Predictive genetic testing can be used to screen for genetic markers to predict susceptibility to a future disease.

Pharmacogenetic testing, which can be used as companion diagnostic, predicts patient responses to a particular treatment. Among various applications, pharmacogenetic testing has direct and clear guidance on prescribing behavior, and is likely to have a more immediate impact in transforming clinical practice.[2]

Pharmacogenetics study how genetic differences influence the variability in patients' responses to drugs, including individual variability in drug dose requirement, efficacy and risk of adverse reactions.[3] Pharmacogenetic information may help to identify the patients who are most likely to respond to a certain drug, and/or to have adverse reactions, and therefore facilitate drug selection and optimize drug dosing to achieve better efficacy and lower risk of side effects.[4]

Large amount of pharmacogenomic information is available. Of 1200 drugs reviewed by the United States Food and Drug Administration (US FDA) between 1945 and 2005, 10% have pharmacogenomic information in their drug label.[4,5] A drug utilization review based on the prescription claims database of a large pharmacy benefits manager in the US showed that a quarter of all outpatients received at least one drug with pharmacogenomic information on the label.[5] Applying pharmacogenetic information is therefore promising to have significant impact on medication usage and treatment outcomes.

One success of pharmacogenetic testing in influencing clinical practice is the targeted treatment of cancers.[6-9] Breast tumors that overexpress human epidermal growth factor receptor type 2 (HER2) have better response to the drug trastuzumab. HER2 gene-amplification test and HER2 protein immunochemistry tests are now used to identify patients whose tumor cells overexpress HER2 and are therefore more likely to benefit from trasuzumab treatment.[6,8] The American Society of Clinical Oncology recommends KRAS mutation testing for all patients with metastic colorectal carcinoma before anti-EGFR antibody therapy, and states that those with mutations in codon 12 or 13 should not receive anti-EGFR antibody therapy.[9] Immunochemistry tests for two other proteins: EGFR and c-kit are also approved as “companion diagnostics” for the colorectal cancer drug Erbitux and the gastrointestinal stromal tumor drug Gleevec, respectively.[6]

1.3 ADRs and pharmacogenetics

Another area of pharmacogenetics with potential is to reduce the risk of adverse drug reactions (ADRs) and improve drug safety. ADRs incur significant health care burden and cost to the health system. It was estimated that in the US in 1994, overall 2216000 hospitalized patients had serious ADRs, among which 106,000 had fatal ADRs.[10] 5%-7% of hospital admissions in US and Europe are due to ADRs each year, which ranks among the top six causes of inpatient death.[10,11] Other than the

threats to the quality of care and medical cost, ADRs have also resulted in the withdrawals of many effective drugs. Between 1999 and 2012, 43 drugs were withdrawn from the market due to ADRs,[12] with an even larger number of drugs experiencing decreased usage after severe ADRs were reported. ADRs therefore have significant adverse impact on availability of drugs and the appropriate use of effective drugs at the health system level, in addition to the direct medical consequences.

In Singapore, 3155 cases of ADRs are voluntarily reported each year to the Health Sciences Authority (HSA), with 49% being classified as severe ADRs, and 22.4% being skin-related disorders.[13] A review of admission causes in a general hospital in Singapore revealed that 0.42% of inpatients had drug allergy, with cutaneous eruptions being the most clinical presentation (95.7%).[14] Serious cutaneous ADRs occurred in 5.2% of patients who had drug allergy.[14] Cutaneous ADRs are therefore among the most concerned ADRs in Singapore.

The most severe forms of cutaneous ADRs are Stevens-Johnson syndrome (SJS), and Toxic Epidermal Necrolysis (TEN). SJS and TEN are life-threatening hypersensitivity reactions, characterized by erosions of the mucous membranes, and extensive detachment of the epidermis.[15,16] SJS is the milder form, where less than 10% of body surface area has skin detachment, with an average mortality of 5%. TEN is the severe form, with skin detachment in more than 30% of body surface area, and a mortality of up to 40%.[15,17] SJS-TEN overlap is a transition between SJS and TEN, with average mortality of 15%. Even though the incidence is low, SJS/TEN is a significant public health concern due to the high mortality, expensive hospitalization and treatment, as well as the fear and reluctance to treatment.[18]

Medications are major causes of SJS/TEN.[17] Notably, many SJS/TEN-causing drugs are commonly used, such as carbamazepine, allopurinol, phenytoin, phenobarbital, amoxicillin, coamoxiclav, cotrimoxazole, and non-steroidal anti-inflammatory drugs (NSAIDs) of the oxicam type.[13,18] Data from the multinational

EuroSCAR study revealed that allopurinol, the first line urate-lowering therapy for chronic gout management, is the most common cause of SJS in Europe and Israel.[19] In Singapore, among all voluntarily reported SJS cases between 2003 and 2008, carbamazepine, a drug commonly indicated for epilepsy, neuropathic pain and bipolar disorders, was the leading drug cause of SJS, followed by phenytoin, cortrimoxazole, and allopurinol.[13] A retrospective study of case records of SJS patients in India revealed that carbamazepine was the most common cause of SJS.[20] The incidence of SJS varies across populations. For instance, the incidence is significantly higher in Han Chinese than Caucasians (8 per million person-years vs 1-6 per million person-years),[17,21] presenting more challenges to drug safety in Asian countries.

The associations between genetic factors and SJS induced by some commonly used drugs have been discovered in the past decade. Chung et al first discovered the strong association between carbamazepine-induced Stevens–Johnson syndrome and the human leukocyte antigen HLA–B*1502 allele among Han Chinese in Taiwan, with the odd ratio being 2,504.[21] The strong association was subsequently confirmed in various other Asian populations, including Han Chinese in Taiwan,[22] Hong Kong,[23] southern China,[24] central China,[25] northern China,[26] Thai,[27,28] Malaysian,[29,30] Singaporean,[31,32] Korean,[33] and Indian[34] populations. On the other hand, the association was not found among Japanese[35] or Caucasian[36,37] populations. The association between allopurinol-induced SJS and the HLA-B*5801 allele was also identified in Han Chinese in Taiwan,[38] Hong Kong,[39] and mainland China,[40] Korean,[41,42] and Thai,[43] where a moderate percentage of population are carriers of the HLA-B*5801 allele (8-20%), as well as Japanese population[35,44] even though the carrier frequency is low.[45] The strong genetic association particularly in Asian populations presents promising opportunities to use genetic testing to guide drug selection to reduce risk of SJS/TEN in various Asian countries with higher incidence of SJS/TEN.

1.4 Adoption of Pharmacogenetic Testing and Barriers

Despite rapid advancement in genetic research, the adoption of pharmacogenetic testing in routine clinical practice is still in its early stage. The number of genetic tests that are commonly used for routine clinical care is relatively small. Of all drug labels with pharmacogenomic information in US and Europe that were licensed between 1998 and 2012, only 14 labels direct clinicians to test prior to prescribing.[46] Often times, the information doesn't lead to specific actions or changes in clinical practice.

There were several barriers in the adoption of pharmacogenetic testing. Four aspects were commonly emphasized when evaluating whether a genetic test should be used in clinical decision making: analytical validity, clinical validity, clinical utility and ethical, legal and social implications.[47,48] A*nalytical validity* and *clinical validity* requires that a test can accurately and reliably detect the genotype, and the association between genotype and clinical manifestation is statistically significant. *Clinical utility* is the extent to which the test can improve treatment outcome or reduce ADR risks for specific patients. Genetic tests that make non-actionable predictions will have limited clinical utility. Testing HLA-B*1502 and HLA-B*5801 allele for drug-induced SJS have relatively well-established analytical validity, clinical validity and clinical utility through case-control genetic association studies and randomized control trials in general patient population. However, three issues remain unclear. Firstly, the current clinical studies focus on the immediate outcome of SJS, but often ignore the long-term medical consequences. Genetic testing may influence the choice of medications, which also influence the long-term treatment efficacy and patients outcomes. Weighing different domains of clinical outcomes can be challenging. In an era of rising health care cost, the cost impact of genetic testing is important. Genetic test may results in higher or lower medical costs. When extra costs are incurred, the value of the service, or whether the benefit justify the cost, becomes an important issues. With limited

healthcare budget and resources, spending on the high value or cost-effective services will achieve the biggest outcome improvement. It is therefore important to quantify the benefit and cost of testing at a societal level to better inform the economic value of testing.

The above barriers to the adoption of HLA-B*1502 and HLA-B*1502 pharmacogenetic testing for SJS are not pure medical decisions, and involve the judgement and tradeoffs between various domains of clinical outcomes, long and short-term outcomes, as well as cost consequences, both at population level and individual patient level.

In this dissertation, I employed a series of economic analyses and economic criteria to evaluate the value of pharmacogenetic testing for the two leading drug causes of SJS/TEN (carbamazepine and allopurinol) in Singapore from a health system perspective and an individual patient perspective. First, cost-effectiveness analyses were conducted to evaluate the long-term cost and benefit of pharmacogenetic testing for these two drugs at the population level. Subsequently, patients preferences for allopurinol pharmacogenetic testing, and tradeoffs made between various factors are quantified using a discrete choice experiment.

1.5 Health economic evaluations to inform decision making

1.5.1 Cost-effectiveness analysis

Cost-effectiveness analysis is a commonly used decision tool in health economics to evaluate new technologies and programs. It systematically compares the costs and effectiveness associated with each available alternative strategy to manage the same condition. Effectiveness is measured in Quality-adjusted life years (QALYs), which unifies various dimensions of clinical outcomes (such as treatment efficacy, side effects, mortality, disability, quality of life, disease duration) into one measure. Cost-effectiveness of a new technology or program is usually calculated in incremental terms relative to a status quo strategy, to reveal the incremental value added by the

new treatment. A threshold which reflects the societal willingness-to-pay for one quality-adjusted life year is then applied to judge the incremental cost-effectiveness of the new technology or program. New technologies or programs that significantly improves effectiveness at low costs are considered to have high value. Cost-effectiveness offers a criteria to allocate scare resources based on efficiency. Spending on high value treatment and services will lead to efficient use of healthcare resources.

Cost-effectiveness analysis is increasingly used to facilitate decision making at various levels. Clinicians can evaluate the long-term cost and effectiveness of various treatment alternatives and choose the most cost-effective treatment. Regulatory agencies such the National Institute for Health and Clinical Excellence (NICE) in the UK uses a cost-effectiveness threshold in its assessment and guidance.[49] Public and private payers can rely on cost-effectiveness criteria to determine whether a new technology or service will be reimbursed, based on the ground of value and efficiency.

1.5.2 Discrete choice experiment

Discrete choice experiment is a stated-preference method to quantify individual preferences using a series of choice questions.[50-52] When revealed preferences or actual market behaviors are not observable, such as when a market does not exist, or when a product is not yet available, stated preference method can provide useful insights on preferences by offering hypothetical choice sets. Discrete choice experiment is also referred to as choice-based conjoint analysis. The name “Conjoint analysis“ arose from the key characteristics of this type of study that different features of products or services are “CONsidered JOINTly”.[53] Each feature is referred to as an attribute. And each choice alternative is composed of combinations of levels or each attribute. Compared to other stated-preference methods, such as contingency valuation, DCE is advantageous in measuring preferences for each attribute level (the marginal value), the relative important of various attribute, and the tradeoffs between different

attributes.[54] DCE elicit preferences using choice questions, which is a more intuitive and realistic way of everyday decision making, compared to other methods such as rating, or ranking.[53]

First developed in marketing, later adopted by public and environmental economists, conjoint analysis and DCE have been increasingly used in health care in the recent decade. The preference of patients and other stakeholders regarding medical treatments, screening and preventive services, health service delivery, have been used to inform clinical practice and priority setting.[53,55-59] Recently, DCE has gained popularity in informing regulatory decisions. US FDA has published a draft guidance on the use of patients' preference information in 2015.[60] The DCE methodology and the applications in weighing benefit and risk of drugs and devices were reviewed.

Common attributes included in DCEs are health care outcome-related attributes (such as treatment efficacy, side effects, survival etc), health care process-related attributes (such as waiting time, quality of care, mode of service, and type of health care professional), cost attributes, and others. DCE allows the explicit quantification of tradeoffs individual make between different attributes. The tradeoff between an attribute and the cost attribute provides estimates on the monetary value of the attribute level, or the willingness-to-pay (WTP). The DCE results have also been used to predict the choice probability or the uptake rate of a certain product or service.

1.6 Objective and structure of this thesis

The objective of the thesis is to conduct health economic evaluations to generate evidence to inform clinical and regulatory decision making on whether pharmacogenetic testing should be routinely done in order to reduce the risk of drug-induced SJS/TEN in the context of Singapore. The two leading causative agents, carbamazepine and allopurinol, were the focus of my studies.

The thesis consists of four research chapters. Chapter 2 assesses the cost-effectiveness of HLA-B*1502 testing prior to carbamazepine treatment for epilepsy

patients from a health system perspective to inform clinical and regulatory policy making in Singapore. The study, together with other related studies, has led to a change in regulatory recommendation, and subsequent changes in clinical practice. These changes as well as intended outcomes and unintended consequences were briefly reviewed at the end of Chapter 2. Chapter 3 assesses the cost-effectiveness of HLA-B*5801 testing for allopurinol among chronic gout patients. Nevertheless, cost-effectiveness analysis does not speak to what individual patients should do or will do. Chapter 4 adopts a different perspective to review the literature on patients' preferences for pharmacogenetic testing, and motivate the study in Chapter 5. Chapter 5 describes a discrete choice experiment to quantitatively measure patients' preferences for genetic testing prior to initiating allopurinol in chronic gout treatment. The thesis is concluded in Chapter 6, with recommendations for future research.

Chapter 2 . Cost-effectiveness of HLA-B*1502 Genotyping Newly Diagnosed Adult Epilepsy Patients in Singapore

2.1 Abstract

Objective

Asians who carry the HLA-B*1502 allele have an elevated risk of developing Stevens-Johnson syndrome (SJS) and toxic epidermal necrolysis (TEN) when treated with the antiepileptic drugs (AEDs) carbamazepine (CBZ) and phenytoin (PHT). Using data in Singapore, this study evaluates the cost-effectiveness of HLA-B*1502 genotyping, and identifies circumstances in which genotyping and targeted treatment with alternative antiepileptic drugs that do not induce SJS/TEN is likely to be more cost-effective.

Methods

A decision tree model was developed in TreeAge. The model takes into account costs of epilepsy treatments and genotyping, reductions in quality of life (QoL) and increased costs resulting from SJS/TEN complications, the prevalence of the risk allele, the positive predictive value (PPV) of genotyping, life expectancy and other factors.

Results

Compared with the status quo strategy of providing CBZ to all patients without genotyping, genotyping and targeted treatment results in an incremental cost-effectiveness ratio of $37,030/QALY for Chinese patients, $7,930/QALY for Malays and $136,630/QALY for Indians in Singapore.

Conclusions

Due to the different population allele frequencies of HLA-B*1502, genotyping for HLA-B*1502 and targeted epilepsy treatment is cost-effective for

Singaporean Chinese and Malays, but not for Singaporean Indians. Based on the study results, policies and clinical practices have been changed.

2.2 Introduction

The World Health Organization (WHO) estimates the population prevalence of active epilepsy to be 4 to 10 per 1,000 worldwide, but higher in developing countries.[61] Roughly 50 million people worldwide suffer from epilepsy, with more than half living in Asia.[61,62] The first line treatment for epilepsy consists of first generation antiepileptic drugs (AEDs) aimed at reducing the frequency of seizures. Due to their effectiveness and low cost, the most frequently prescribed drugs are carbamazepine (CBZ) and phenytoin (PHT).[62,63] However, they are not without side effects, including cutaneous hypersensitivity reactions, ranging from mild rash to rare but potentially fatal Steven-Johnson syndrome (SJS), toxic epidermal necrolysis (TEN) and SJS-TEN overlap. These conditions are characterized by blistering exanthema of purpuric macules, mucosal involvement and skin detachment.[16,64] The fatality rate is reported to be roughly 5% for SJS, 30% for TEN,[17,64] and somewhere in between for SJS-TEN overlap.[65]

Epidemiologic data reveal that epilepsy patients from certain Asian populations have a higher risk of developing SJS or TEN following CBZ treatment compared with Caucasians.[28,66] In 2004, a strong association between the HLA-B*1502 allele and risk for CBZ-induced SJS and TEN was discovered among Han Chinese in Taiwan (odds ratio of 2,504, positive predicted value of 5.6%, and negative predicted value of 99.9%)[66,67]. This association was later confirmed in various other Asian populations including Han Chinese in Hong Kong,[23] southern China,[24] central China,[25] northern China,[26] Thai,[27,28] Malaysian,[29,30] Singaporean,[31,32] Korean,[33] and Indian[34] populations. On the other hand, the association was not found among

Japanese[35] or Caucasian[36,37] populations. For these groups with established genetic associations, the prevalence of the allele ranges between 5.7 % and 27.5%,[68]whereas it is virtually absent in Caucasians and Japanese. These differences, along with differences in CBZ prescribing patterns, largely explain the differences in CBZ induced SJS/TEN across countries.

Based on the evidence on genetic associations, in 2007, the US Food and Drug Administration (FDA) amended the prescribing information for CBZ, recommending (but not requiring) genotyping in populations in which HLA-B*1502 is present before prescribing CBZ. [69] Though it is not yet accepted practice in Asia, given the above findings, it might seem appropriate to genotype for the HLA-B*1502 risk allele and provide an alternative to CBZ to those who are HLA-B*1502 positive. Phenytoin, another anti-epileptic drug, can also induce SJS/TEN and has also been associated with the HLA-B*1502 allele.(Hung 2010) In Singapore, based on a registry of adverse drug reactions maintained by the Health Sciences Authority, between 2003 and 2009, 262 reports of SJS, 35 reports of SJS-TEN overlap and 74 cases of TEN were received. CBZ was the leading suspected causative agent in 18% of the reports, whereas phenytoin (PHT) was suspected in 9.6% of cases.[70] Therefore HLA-B*1502 genotyping may be considered prior to CBZ or PHT treatment for epilepsy patients in Singapore to reduce risk of SJS/TEN.

Despite the risk reduction, there are several concerns related to the adoption of genotyping, particularly on the higher costs, and the predictive power of the genetic test. There are alternatives drugs to CBZ and PHT for those suffering from epileptic seizures, including sodium valproate (SVP), lamotrigine (LTG), topiramate (TPM), levetiracetam (LEV) and gabapentin

(GPT). These drugs have comparable efficacy to CBZ but lower or no risk of SJS/TEN. However, they are substantially more expensive. It is also not clear whether genotyping and using these alternative medications for those who test positive for the HLA-B*1502 risk allele is cost-effective. The other issue concerns the predictive power of the test. The risk allele is present in 5.2% of Singaporean Chinese (Singapore Immunology Network) and 15.7% of Malays,[71] In contrast, the incidence of SJS/TEN is around 0.2% among Han Chinese, implying that even among risk allele-carriers, more than 90% will not develop SJS/TEN.

The goal of this analysis is to present a cost-effectiveness model to allow for identifying those circumstances in which genetic testing and targeted treatment with an alternative medication for those who test positive is likely to be more cost-effective that: 1) treatment with CBZ or PHT without genotyping and 2) providing alternative drugs with no SJS/TEN risk without genotyping. Although the model is populated using cost and SJS or TEN data from Singapore, through sensitivity analyses it identifies the threshold conditions in which genotyping and targeted therapy would be cost-effective in other settings. The model and results will be useful for all countries and health plans considering the decision of whether or not to genotype for the risk allele.

2.3 Methods

A decision tree model was developed in TreeAge Pro 2011 (TreeAge Software, Williamstown, MA) to estimate the cost-effectiveness of HLA-B*1502 genotyping for newly diagnosed adult epilepsy patients in Singapore for whom CBZ or PHT is considered suitable as first-line monotherapy (Figure 1). CBZ and PHT are assumed to be perfect substitutes in the model and denoted as

CBZ/PHT, for their similar cost, efficacy and safety profiles.[72] A local anti-epileptic drug usage study revealed that CBZ and PHT was used as first line monotherapy in 74% of adult patients with newly diagnosed epilepsy, most of whom had partial seizures. [73] The model also assumes VPA to have comparable efficacy and safety profile with CBZ/PHT but without SJS/TEN risk.[74,75] Though the evidence on relative efficacy of various drugs remains inconclusive, this assumption is supported by clinical trials and meta-analyses on most seizure types.[76] [72,76,77]

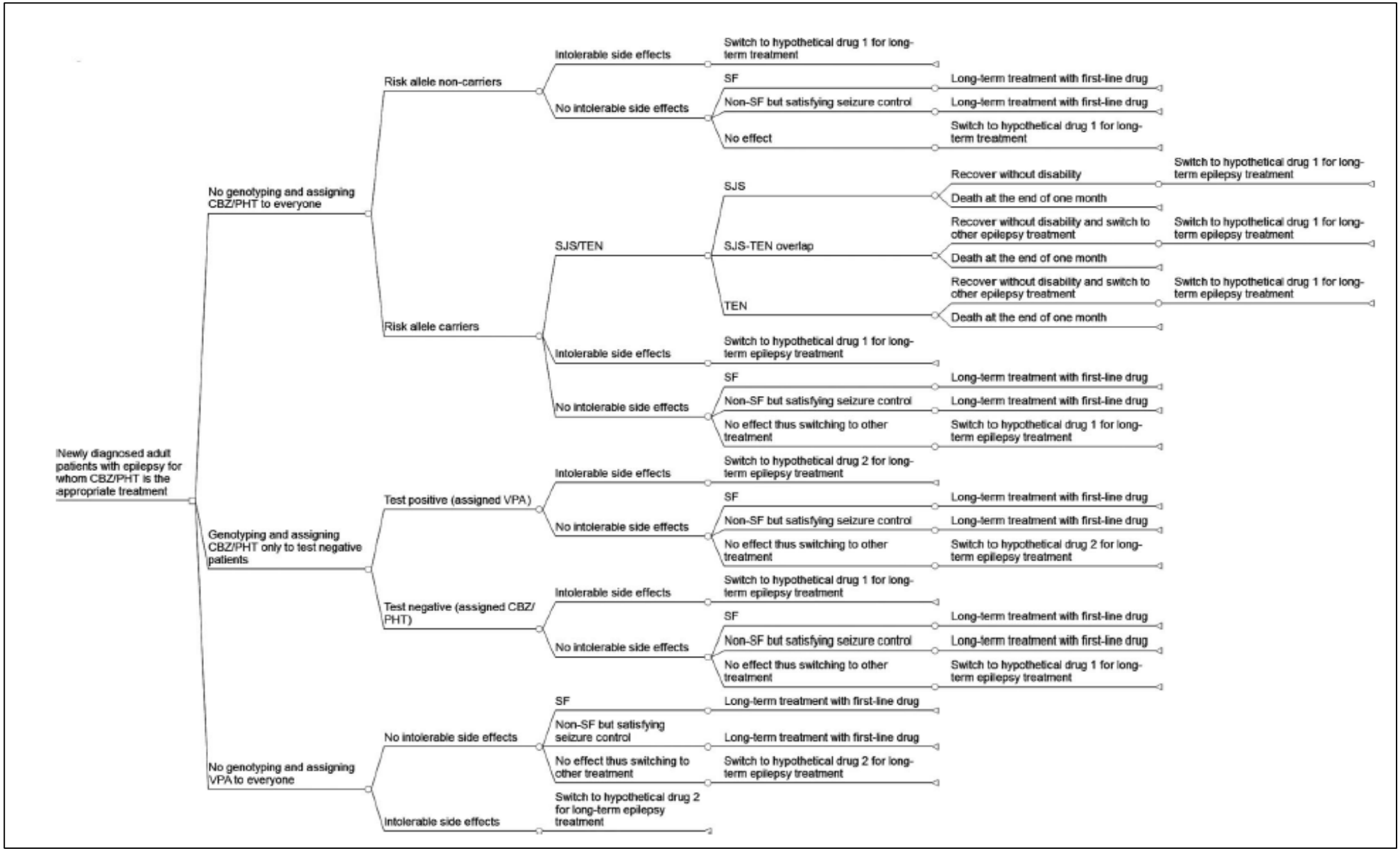


Figure 1. Decision tree model of three treatment strategies for newly diagnosed adult epilepsy patients in Singapore for whom CBZ/PHT is considered appropriate treatment. CBZ: carbamazepine; PHT: phenytoin; VPA: valproate acid; SJS: Stevens-Johnson syndrome; TEN: toxic epidermal necrolysis; SF: seizure free.

Model structure

Figure 1 shows the different strategies modeled, and the treatment pathway and patients outcomes for each treatment strategy. The upper branch of the decision tree represents the status quo practice, which is using CBZ/PHT as first line treatment without genotyping. The middle branch considers the genetic testing strategy, where all newly diagnosed adult epilepsy patients are genotyped for HLA-B*1502 allele before treatment initiation. Test positive patients will receive VPA as first line drug, and test negative patients receives CBZ/PHT due to minimal risk of SJS/TEN. The lower branch examines an alternative risk-mitigation strategy that is likely to occur in real clinic settings, where CBZ/PHT is avoided and all patients receive VPA as first line therapy without genotyping.

Based on the clinical literature, five treatment outcomes are modeled post initial treatment (1) being seizure-free (SF) after treatment and continuously taking the same drug for the long term; (2) being non-SF but achieving satisfactory seizure control (defined as achieving greater than 50% reduction in seizure frequency), and taking the same drug for the long term; (3) showing no satisfactory response (defined as < 50% reduction in seizure frequency) to the drug and switching to an alternative drug; (4) having intolerable side effects (such as rash, fever, fatigue, dizziness, alopecia as documented in clinical trials), and switching to an alternative drug; (5) development of SJS /TEN and complete recovery, followed by alternative epilepsy treatment; (6) death due to SJS/TEN. The first four outcomes are common to all three strategies and the last 2 only possible in the status quo branch.

When SJS/TEN occurs, extensive and expensive medical care is required, but usually last for only a few days or weeks.[65] Patients are assumed to either die within 1 month after CBZ/PHT initiation or to fully recover by the end of 1 month. For patients requiring second line epilepsy treatments (with outcome 3, 4 and 5), we modeled a hypothetical second drug whose cost is a weighted average of the commonly used anti-epileptics, and producing an efficacy reflecting the average efficacy of different drugs. For patients who fail CBZ/PHT treatment, the alternatives include VPA, Lamotrigine (LTG), Levetiracetam (LEV) and Topiramate (TPM), whereas for patients who fail with VPA treatment and intend to avoid all SJS/TEN-inducing drugs, LEV and TPM are the assumed alternatives. The choice of second line drugs and treatment pathway in each scenario, drug dosage and usage patterns were advised by physicians.

To mirror clinical practice, we explicitly modeled three distinct treatment periods. The first period spans the first month after treatment initiation, after which clinicians evaluate the risk of intolerable side effects and life-threatening SJS/TEN. One month is chosen as the literature shows most SJS/TEN cases develop within 3 weeks.[64] The second period encompasses months 2 through 6, which allows physicians to evaluate short-term drug efficacy. Short-term efficacy data are from a clinical trial conducted in the UK.[78] Based on treatment efficacy in the second period, treatment is adjusted in the third period and continues for an additional seven years, which is roughly the median cumulative treatment duration. Even though some guidelines and physicians support life-time treatment even for those who remain seizure-free, it is not common in practice. Beyond the treatment period we assume that treatment is

discontinued and health related quality of life is restored to perfect health and lasts for another 30 years. The time horizon of 30 years is chosen as the average onset age for adult epilepsy is around 40 and average life expectancy for epilepsy patients is 70 (ten years shorter than that of the general population).[79] The impact of these assumptions on results was evaluated in sensitivity analyses.

Model Inputs

Table 1 lists all input variables and sensitivity ranges. Several key variables are described below.

Table 1. Model Inputs.

Variable Name	Base-case value	Range for sensitivity analysis	Source
Cost (in 2010 US dollars)			
Average annual cost of CBZ/PHT (Daily median dosage = 420mg /300mg)	170	85-340	Selling prices were from IMS HEALTH and median daily dosage prescribed by local clinicians
Average annual cost of VPA (Daily median dosage=1050 mg)	470	235-940	
Average annual cost of hypothetical therapy for patients who fail CBZ treatment	1,100	550-2,200	
Average annual cost of hypothetical therapy for patients who fail VPA treatment	1,860	930-3,720	
Cost of HLA-B*1502 genotyping	270	80-380	
Cost of per case SJS treatment	3,480	1,740-5,220	Singapore public hospital discharge data
Cost of per case SJS-TEN overlap treatment	10,250	5,125-15,380	
Cost of per case TEN treatment	17,030	8,510-25,540	
Cost of therapeutic drug monitoring test	15	8-23	Public hospitals in Singapore
cost of neurologist consultation (per visit)	80	38-115	
QoL			
SF with tolerable side effects	0.9418	0.8836-1	[80,81]
Non-SF but show >50% reduction in seizure frequency	0.907	0.814-1	
No effect	0.8288	0.7576-0.9	
On hypothetical treatment	0.909	0.868-0.95	
Intolerable side effects	0.8	0.7-0.9	
SJS (duration=8.9 days)	0.35	0.175-0.525	Estimated with reference to QoL of burn patients[82]
SJS-TEN overlap (duration=9.2days)	0.3	0.15-0.45	
TEN (duration=12.4 days)	0.25	0.125-0.375	
SJS/TEN fatality and incidence			
Fatality for SJS	5%	2.5-7.5%	[64]
Fatality for SJS-TEN overlap	15%	7.5-22.5%	
Fatality for TEN	30%	15-45%	
Percentage of SJS-TEN overlap among SJS/TEN overlap and TEN	10%	5-15%	Singapore Health Sciences Authority (2003-2009 data)[70]
Percentage of TEN among SJS/TEN	20%	10-30%	
HLA-B*1502 genotyping			
Population frequency of HLA-B*1502[a]	14.87%	11-18.74%	[83] and unpublished data from Singapore Genome Variation Project and Singapore Immunology Network
Positive predictive value of positive genotyping results in CBZ/PHT users[a]	5. 96%	4-7.92%	[84]
Efficacy and safety of CBZ (clinical response at 6 months post treatment initiation among patients with partial seizures)			
Non-seizure-free but show >50% reduction in seizure frequency and stay on treatment	48%	38-58%	[78]
No effect	8%	2-14%	
Intolerable side effects	25%	5-45%	
Other inputs			
Duration modeled (in years)	30	20-40	[79]
Percentage of duration on epilepsy treatment	23.3%	25%-41%	Clinician's recommendation
Annual discount rate	3%	0-5%	

All monetary amounts are presented in US dollars. Data in Singapore dollars were converted to US dollars using exchange rate $1.3 Singapore dollars =$1 US dollar.

Incidence of SJS/TEN

Calculating the incidence of CBZ/PHT-induced SJS/TEN in Singapore is challenging, as the exact number of new CBZ/PHT users are not measured, and cases are reported on voluntary basis. A study in Taiwan used the national insurance claims database and estimated the incidence of CBZ-induced SJS to be 0.23%. Taiwan Chinese and Singaporean Chinese have similar origin, and genetic profiles. Therefore, the incidence among Singapore Chinese is assumed to be the same as that in Taiwan (0.23%).[84] To estimate the incidence in Singaporean Malays and Indians, we used data from a voluntary adverse drug reaction registry maintained by the Singapore Health Sciences Authority.[70] There may be under-reporting in voluntary registries, we therefore assumed the incidence to be the same among Singaporean Chinese and Taiwan Chinese, and scaled the estimated incidence for Singaporean Malays and Indians assuming equal degree of under-reporting for different ethnic groups. The adjusted incidence of CBZ/PHT-induced SJS/TEN among Singapore Malays and Indians patients initiating CBZ/PHT are 0.61% and 0.14%. More than 95% of Singapore resident population are Chinese, Malays or Indians.[85] Among ethnicity-weighted Singapore CBZ/PHT users, the incidence is 0.27%.

*Positive predictive value (PPV) of HLA-B*1502 genotyping*

PPV is defined as the probability of actually developing the condition when the test predicts the condition. Based on the sensitivity and specificity established in Taiwan Chinese (98.3% and 95.8% respectively),[84] PPVs of HLA-B*1502 genotyping were estimated as 5.96% for the entire population, 5.1% for Singapore Chinese, 12.5% for Singapore Malays and 3.2% for

Singapore Indians. In the base case analysis, we considered the ethnicity-weighted Singapore population. The ethnicity-weighted Singapore population was considered for main analysis. The negative predictive value (NPV) is close to 100%.

Costs and Utilization

Wholesale prices of available anti-epileptic drugs in Singapore in 2010 were obtained from IMS HEALTH. To approximate the retail prices, the obtained wholesale prices were multiplied by 1.2 to account for the markup. Average daily costs for each drug was calculated by multiplying the unit price by the median dosage for each drug, as commonly prescribed by local clinicians (Supplementary Table e1). The costs of hypothetical drugs were calculated as a weighted (by utilization) average of the several commonly used alternative drugs. SJS and TEN treatment costs were estimated based on National University Hospital discharge data for 20 cases. None of these cases were fatal, and we assumed the costs for cases that ended in a fatality to be double of the base case value due to additional resources required at the end of life. We made the assumptions that each patient required one therapeutic drug monitoring test immediately after treatment initiation, four specialists visits in the first year of treatment and 2 visits per year thereafter during treatment period. All costs were converted to US dollars at the exchange rate of $1.3 Singapore dollars to $1 US dollar as of October 2010.

Sensitivity Analyses

The robustness of the cost-effectiveness results and the impact of specific parameters were tested through one-way sensitivity analyses and a probabilistic sensitivity analysis (PSA). In one way sensitivity analyses,

variables were varied one at a time, within reasonable sensitivity ranges, and the incremental cost-effectiveness ratios were calculated. Several key variables of interests were further analyzed using threshold analysis to identify the threshold at which the cost-effectiveness results will be altered. Scenario analysis and two-way sensitivity analyses were also conducted.

The probabilistic sensitivity analyses allowed all variables to vary simultaneously based on 10,000 repeated draws from assigned distributions. All variables except percentage of remaining life expectancy on treatment were assumed to follow triangular distributions. The base-case value was used as the likeliest value in the triangular distribution, and the lower and upper bounds of the sensitivity ranges were used as the min and max (Table 1). The percentage of remaining life time a patient is on epilepsy treatment is approximated using a bimodal distribution, which was constructed as a combination of two triangular distributions to account for patient heterogeneity in drug responses and epilepsy recurrence. The first triangular distribution (min=2 yrs; mode=3.5 yrs; max=5 yrs) represents patients with good responses to drugs and no recurrence, whereas the second triangular distribution (min=10 yrs; mode=15 yrs; max=20 yrs) corresponds to patients who require longer term treatment. We assumed that 60% of patients fall in the first distribution and 40% in the second, based on expert opinions. We conducted sensitivity analyses to explore the impact of this assumption on the cost-effectiveness results (Supplementary Table e2). This study was reviewed and granted exemption by the National University of Singapore Institutional Review Board (NUS IRB).

2.4 Results

Base case cost-effectiveness results are shown in Table 2. Effectiveness is measured in quality-adjusted life years (QALYs), which is, the remaining life years after adjusting for quality of life (Qol) within that time period. Qol is a quality weight between 0 and 1, with 0 indicates death and 1 represents perfect health). Our results show that genotyping and prescribing VPA for those who test positive generates a modest improvement in QALYs (0.019 QALYs) at a $570 marginal increase in cost relative to the status quo practice, resulting in an incremental cost-effectiveness ratio (ICER) of $29,750/QALY. The strategy of providing VPA to all patients without genotyping is not favorable as it gives the same QALYs as the genotyping strategy but at a higher cost. This is referred to as a dominated strategy. If the annual cost of VPA drops to within $37 of the cost of CBZ, this strategy would become cost-effective.

Table 2. Cost-effectiveness of 3 strategies for newly diagnosed epilepsy patients in Singapore for whom CBZ/PHT is considered appropriate treatment

Strategy	Cost (US dollars)	Incremental Cost (US dollars)	QALYs	Incremental QALYs	ICER (US dollars/QALY)	Dominance
No genotyping and CBZ/PHT for all patients	4,110	-	18.846	-	-	Not Dominated
Genotyping and VPA for test positive patients and CBZ/PHT to test negative patients	4,680	570	18.865	0.019	29,750	Not Dominated
No genotyping and VPA for all patients	6,780	2,100	18.865	0	0	Dominated

ICER: incremental cost-effectiveness ratio, QALY: quality-adjusted life-years

The cost-effectiveness results for Singapore Chinese, Malays and Indians are shown separately in Table 3. Relative to the status quo strategy,

genotyping is cost-effective for Chinese and Malays and but not for Indians in Singapore.

Table 3. Incremental cost-effectiveness of genotyping versus no genotyping strategy for 3 major ethnical populations in Singapore

Ethnicity	Cost (US dollars)	Incremental Cost (US dollars)	QALYs	Incremental QALYs	ICER (US dollars/QALY)
Singapore Chinese	4,650	560	18.865	0.015	37,030
Singapore Malays	- 5,050	610	18.865	0.077	7,930
Singapore Indians	4,370	360	18.865	0.00263	136,630

ICER: incremental cost-effectiveness ratio, QALY: quality-adjusted life-years

Sensitivity analyses

A commonly used cost-effective threshold is $50,000/QALY[86]. Using this threshold to define what signifies cost-effectiveness, the one-way sensitivity analyses (Figure 2 Panel A) show that any single variable when varied within the assigned sensitivity does not increase the ICER beyond the cost-effectiveness threshold.

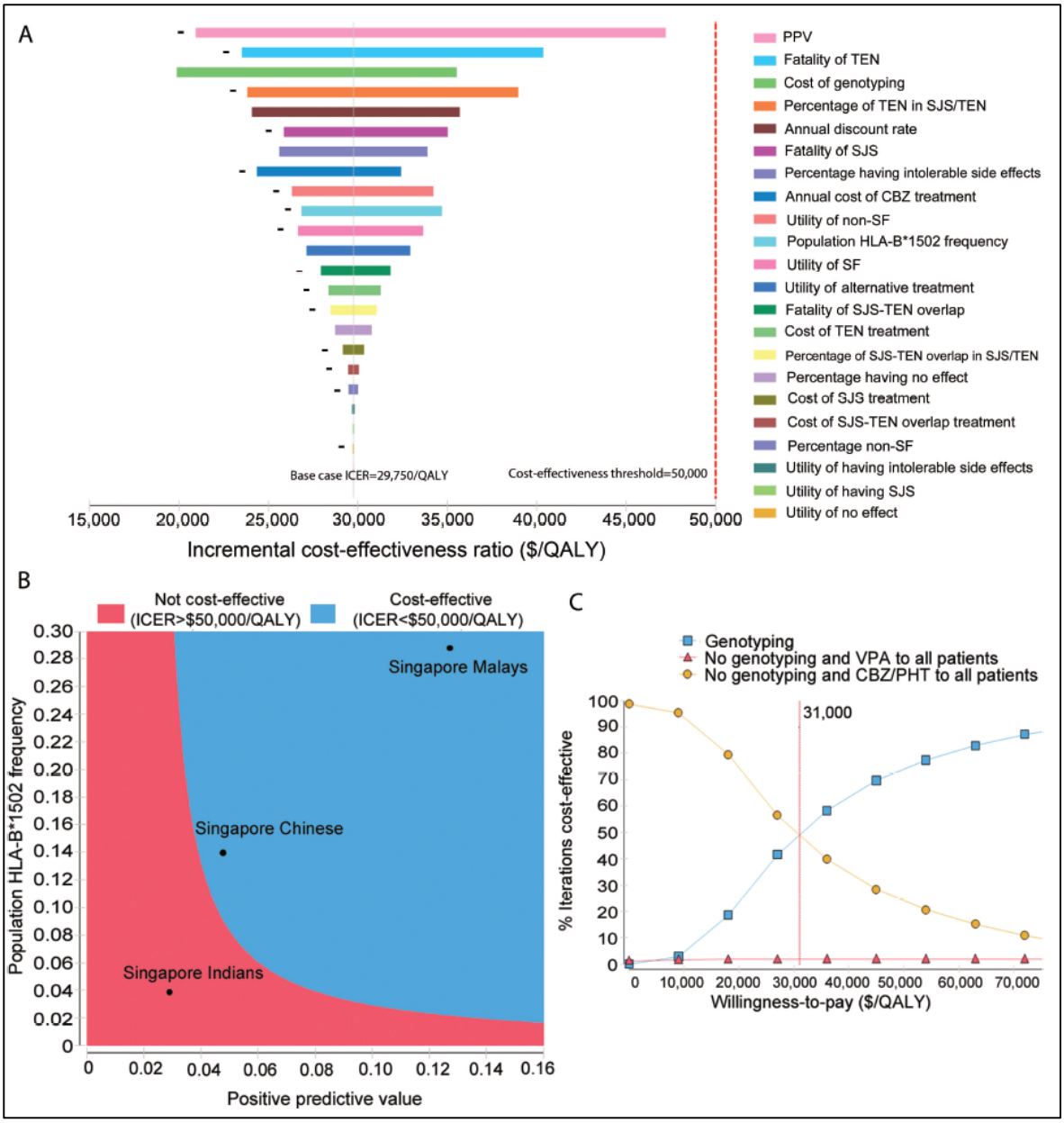


Figure 2. Sensitivity analyses. (A) One-way sensitivity analysis of all uncorrelated variables on Incremental cost-effectiveness ratio (ICER) of the genotyping strategy. Variables that are correlated with other variable(s) were not shown. The minus sign at the left side of the bar indicates ICER decreases when the variable increases. SF: seizure free. (B) Two-way sensitivity analysis of the effects of positive predictive value (PPV) and population HLA-B*1502 frequency on ICER. (C) Cost-effectiveness acceptability curves of 3 treatment strategies from probabilistic sensitivity analysis. Red vertical line: the willingness-to-pay at which genotyping is cost-effective for 50% of the iterations.

In efforts to generalize the model beyond Singapore, various sensitivity analyses were conducted. Two population-specific variables, PPV and frequency of HLA-B*1502 allele in the population, were varied within wider ranges to identify the threshold values at which the cost-effectiveness results would alter. Holding other variables constant at the base case values, a PPV below 3.8% will increase the ICER to above $50,000/QALY, as will an HLA-B*1502 population frequency lower than 6.1%. To show the impact of combinations of PPV and allele frequency on the ICER, two-way sensitivity analyses are shown in Figure 2B. Genotyping is cost-effective in populations with higher test PPV and higher HLA-B*1502 frequency. A higher allele frequency could compensate for a lower PPV to make genotyping cost-effective. However, if the PPV is below 3%, genotyping is likely to be not cost-effective regardless of allele frequency. This result holds under current genotyping costs, however, a lower genotyping cost could compensate for a lower PPV.

Using the base case values of input variables, the probabilistic sensitivity analysis (PSA) reveals that, assuming a willingness-to-pay of $50,000/QALY, genotyping is cost-effective in 75% of iterations. As long as the societal willingness-to-pay is higher than $31,000/QALY, genotyping would be preferred in more than 50% of iterations among 10,000 draws in the simulation (Figure 2C). Additionally, the genotyping cost of $270 currently represents roughly 6% of the expected epilepsy treatment cost. This percentage is likely to decrease in the future due to technological advancements and increased availability of genotyping services, which will make genotyping even more cost-effective.

2.5 Discussion

The study estimated the incremental cost-effectiveness of genotyping for each ethnic group in Singapore, and revealed the differences in cost-effectiveness for each ethnic groups. This is due to the differences in population characteristics HLA-B*1502 frequency and PPV of genotyping. The product of the two determines the likelihood of an average CBZ/PHT user developing SJS/TEN. In the base case, it is estimated that the proportion of new patients who would develop SJS/TEN after initiating CBZ/PHT is 0.70% among Chinese, 3.55% among Malays and 0.12% among Indians. Among those who develop SJS/TEN, fatality is expected to be 9.5%. To prevent one case of SJS/TEN, 142 Chinese patients, 28 Malay patients, or 833 Indian patients would need to be genotyped on average. To avoid one death due to CBZ/PHT-induced SJS/TEN, 1,500 Chinese patients, 297 Malays patients and 8,770 Indians patients would need to be genotyped prior to initiating CBZ/PHT.

Besides population HLA-B*1502 allele frequency and PPV of HLA-B*1502 genotyping, two additional factors influential on results are the treatment duration and remaining life expectancy (Table e2). Longer treatment duration increases the long-term costs of providing expensive alternative medications. However, among those expected to live a long life, such as young people, death due to SJS/TEN generates a large loss of QALYs. The cost-effectiveness results from the tradeoff of these two factors. Nevertheless, our model shows that if treatment is life-long, then genotyping is not cost-effective regardless of remaining life expectancy. This is because when the life expectancy is short, preventing an SJS/TEN-induced death

results in few QALYs saved; whereas for long life expectancies, the increased cost of lifetime alternative treatments, drives the ICER beyond the acceptable threshold. As a general rule of thumb, as the percentage of remaining life on treatment increases, genotyping becomes less cost-effective.

The above analysis assumes drug prices to be at the base case values. However, there is substantial variation in medical practices and drug prices across countries and across health plans within countries. Higher drug prices (due to either higher dosage prescribed or higher prices per unit) would reduce the cost-effectiveness of HLA-B*1502 genotyping, all else equal. For example, if all drugs cost 5 times that of the base case values, the ICER of genotyping would increase to $107,520/QALY, which is no longer cost-effective. However, locations with high anti-epileptic drug costs are likely to also have higher costs of SJS/TEN treatments, which may drive the cost-effectiveness ratio back to acceptable levels depending on the magnitude of the increase. If the costs of a particular drug change differentially from the other drugs, such as when patent expires or when the demand changes, ICERs will change accordingly. For example, if VPA and CBZ have the same price, then providing VPA to all patients as first line therapy would be the preferred strategy as it avoids the risk of SJS/TEN without the need to genotype. On the other hand, if the prices of all alternative drugs increase beyond $1,420 per year while the price of CBZ remains unchanged, genotyping be not cost-effective, and using CBZ/PHT without genotyping would become the optimal strategy, as the higher long-term costs of epilepsy treatment would outweigh the benefit of genotyping.

In the study, it is assumed that health related qualify of life is restored to perfect health after anti-epileptic treatment. Actual health related qualify of life may be lower due to imperfect responses to drugs, recurrence of epilepsy, or other health problems. If true, the QALY gains due to prevention of SJS/TEN and related death might be over-estimated. However, scenario analysis shows that even assuming that the low QoL during treatment is sustained until death, the ICER for genotyping ($27,980) is still below the cost-effectiveness threshold. Additional sensitivity analyses reveal that, based on our assumptions, as long as the QoL for successful treatment is greater than 0.83, genotyping is cost-effective. We also made assumptions on the clinical treatment pathways, following clinical guidelines and experts opinions. However, in reality, treatment decisions depends on many factors and may substantially deviate from our base case assumptions. For instance, when selecting anti-epileptics, patients who does not tolerate a single seizure may request to switch to the (more expensive) alternative drugs even when they have a substantial reduction in seizure frequency. Among those patients, genotyping would not be cost-effective as they are more likely to switch to the more expensive drugs irrespective of the genotyping results. In addition to the above, the model includes several additional assumptions and simplifications. The model simplifies the treatment rules for who receives which drugs. In real clinic settings, many factors, including seizure type may influence the treatment regimens. Besides, treatment options are sometimes more complicated than what's captured in the model, such as when more than two lines of treatments and combination therapies using multiple drugs are involved. In addition, based on available literature, this study assumes VPA,

CBZ, and PHT to have similar efficacy for the epileptic conditions concerned in this study. This assumption is supported by clinical trials and meta-analyses for generalized seizures and secondary generalized tonic-clonic seizures[76,77]. For partial seizures, some evidence suggests CBZ is superior to VPA in the short term for complex partial seizures[75-77]. In cases where CBZ is superior to VPA, genotyping will be less cost-effective. Moreover, the effectiveness data is from clinical trials in Caucasian populations. Though no evidence suggests differences on drug response and QoL perception among epilepsy patients across different ethnicities, we cannot rule out the possibilities of population variations in drug response, cultural differences on QoL values, or different clinical practices.

A final limitation is that genotyping results are assumed to be immediately accessible at the time of diagnosis, which may be challenging in clinical practice. While waiting for testing results (several days), it may be appropriate to provide an alternative treatment to CBZ/PHT until when the genotyping results can be obtained, and then, switch to CBZ/PHT for those who test negative.

Barring the above limitations, this model provides a template to assess the cost-effectiveness of HLA-B*1502 genotyping in other Asian countries, though local clinical practice and medical costs should be considered. In general, in countries with high HLA-B*1502 frequency and high incidence of CBZ/PHT-induced SJS/TEN, genotyping is more likely to be cost-effective. This includes many Southeast Asian countries (such as Singapore, Malaysia, Vietnam, Thailand, Indonesia, and the Philippines) and southern eastern regions of Asia (such as Hong Kong, Taiwan and certain southern

provinces of china). The frequency of HLA-B*1502 in these populations is generally higher than 5% and even above 20% in some ethnic groups[287]. Contrarily, prevalence is below 2.5% in India (except certain ethnic groups) and northern Asian countries including Japan, South Korea, and northern regions of China, suggesting that genotyping is unlikely to be cost-effective in these regions.

*2.6 Changes in HLA-B*1502 genotyping policies and practices in Singapore*

Various regulatory actions have been undertaken after the completion of this study, and clinical practice has changed as a result of this study and other related studies. Through the collaborative effort of multiple sectors to implement HLA-B*1502 testing in clinical practice, valuable lessons have been learned.

Before this study

In March 2009, the Health Sciences Authority (HSA) published a Product Safety Alert on serious adverse skin reactions associated with carbamazepine based on international studies on the genetic association, local ADR reports and US FDA recommendations.[88] The package insert of Tegretol® (carbamazepine) was updated in Singapore by the manufacturer to reflect the association observed between HLA-B*1502 allele and CBZ-induced SJS, the prevalence of this allele in various Asian population, and a recommendation to consider testing for the presence of HLA-B*1502 allele in patients with Asian ancestry prior to prescribing Tegretol®. In the package insert, it was also stated that the use of carbamazepine should be avoided in

tested patients who are found to be positive for HLA-B*1502 unless the benefits clearly outweigh the risks.

However, as the test was new to Singapore, the service was not readily available in hospitals or clinics in Singapore where carbamazepine was prescribed. The only accredited lab in Singapore that offers the HLA-B*1502 test was the HSA's Tissue Typing Laboratory, which conducted comprehensive HLA typing mostly for patients prior to organ transplantation and bone marrow transplantation.

There were several perceived barriers to the uptake of the test. Firstly, the test could only be done outside practitioners' institutions, which adds additional administrative workload for physicians and hospital staff to order the test, transport samples, and receive hard copy test results. The independency of the IT systems between different institutions created difficulties to the delivery of test results, the incorporation of results into electronic medical record, and the sharing of test results between different providers such as the tertiary hospitals and primary care clinics. Secondly, the available test service was not tailored for carbamazepine testing. The tissue-typing based procedure has high accuracy, but high cost (S$350) and long turnover time (3-7 working days). A cost of S$350 was considered high relative to the cost of carbamazepine. In addition, for epilepsy patients who require immediate relief, a turnover time of 3-7 days may cause delay in their critical treatment. Indeed, a low take-up rate was observed. On the other hand, an unpublished analysis of the anti-epileptics sales data from IMS HEALTH database and communications with neurologist both suggested a drop in the use of carbamazepine. Similar trend was also observed in Hong

Kong.[89] With the risk information provided, and barriers to testing, an easy alternative solution was to switch away from carbamazepine to alternative medicines. Carbamazepine is an old generic drug with long proven clinical efficacy and low cost. Switching from carbamazepine to alternative drugs that are often branded and more expensive will elevate medical costs. This is considered an unintended consequence of policy. More efforts were needed to promote the appropriate use of the risk information.

After the study

With stronger evidence on genetic association, clinical utility, and cost-effectiveness (thanks to our study), in April 2013, Singapore Ministry of Health (MOH) made an announcement that HLA-B*1502 genotyping prior to the initiation of carbamazepine therapy in new patients of for Asian ancestry was the new standard of care. HSA, together with MOH, issued a Dear Healthcare Professional Letter to communicate the new recommendations for HLA-B*1502 genotyping the use of test results.[90] Meanwhile the National University Hospital (NUH) Molecular Diagnosis Centre (MDC) started to offer the test at a cost of S$187 (excluding GST) with a turnover time of 2-4 working days. The decrease in price was due to economies of scale and improvements in testing methods. To ensure the access to test service by low-income patients, 75% of the test cost was subsidized for patients from the MOH-funded restructured hospitals and institutions. The test later became available in several other hospitals, with improvement in IT system and results delivery.

6 months after the new recommendation was announced, a preliminary evaluation was published by HSA in November 2013.[91] A total of 307 tests were performed, with 9.8% of samples tested positive for HLA-B*1502.

Contrary to a historical average of 15 CBZ-SJS/TEN reports to HSA per year, no SJS/TEN report related to carbamazepine was received by the time of publication.

Chapter 3 . HLA-B*5801 genetic testing and safety program when initiating allopurinol therapy for chronic gout management: a cost-effectiveness analysis

3.1 Abstract

Objective

Allopurinol is an efficacious urate-lowering therapy (ULT), but on rare occasions, patients develop potentially fatal adverse reactions. The risk of reactions such as Stevens-Johnson syndrome (SJS) is significantly higher among HLA-B*5801 carriers. We assessed the cost-effectiveness of risk-mitigation strategies that use HLA-B*5801 genetic testing, an enhanced safety program or a combination of both.

Methods

The analysis adopted a health systems perspective and considered Singaporean patients with chronic gout, over a lifetime horizon, where allopurinol and probenecid are appropriate medications. The model incorporated SJS outcomes, long-term gout treatment outcomes, HLA-B*5801 allele frequencies, drug prices, and other medical costs.

Results and Conclusions

Based on a cost-effectiveness threshold of US$50,000/QALY, HLA-B*5801 guided ULT selection or enhanced safety program were not cost-effective in the base case analysis. Avoidance of ULTs was the least preferred strategy as uncontrolled gout leads to lower QALYs and higher costs. Conditions under which genotyping or enhanced safety program would become cost-effective were identified.

3.2 Introduction

Gout is a common rheumatic disease with increasing prevalence worldwide due to increased longevity, dietary changes, and greater use of medications with urate

retentive effects such as diuretics and low-dose aspirin.[92-94] Gout increases medical costs, reduces patients' quality of life (QoL), [95-97] and is an independent risk factor for all-cause and cardiovascular mortality.[98]

Pharmacologic management of chronic gout aims to reduce serum uric acid (SUA) levels to prevent formation and promote crystal dissolution.[97] Allopurinol is generally well-tolerated and the most commonly used urate-lowering therapy (ULT).[99] However, Allopurinol was one of the drugs most commonly associated with Stevens-Johnson Syndrome (SJS) and Toxic Epidermal Necrolysis (TEN),[17,18,100-103] which are rare but serious cutaneous reactions with average fatality at 30% for TEN.[64,104]

Strong genetic association between HLA-B*5801 and allopurinol-induced SJS/TEN was confirmed in various populations,[42,43,102,105-107] suggesting that genotyping may mitigate risks of allopurinol-induced SJS/TEN. The test has a negative predictive value (NPV) of close to 100% but a positive predictive value (PPV) of only 1.52% for SJS/TEN among Han Chinese in Taiwan.[43,106] The American College of Rheumatology recommends HLA-B*5801 genotyping as a risk management measure for at-risk populations;[94] however the European Medicines Agency (EMA) Pharmacovigilance Working Party cautions against routine HLA-B*5801 genotyping given the lack of suitable alternative therapies to allopurinol and the lack of evidence of clinical utility.[108] The Taiwan Food and Drug Administration has issued a notice that HLA-B*5801 should be considered prior to allopurinol treatment, but testing is not mandatory.[109] Given the seriousness of SJS/TEN, but low PPV of the HLA-B*5801 genetic test, it is still unclear what role HLA-B*5801 genetic testing should play in clinical practice, especially in Asian populations with high prevalence of HLA-B*5801 allele, such as the Han Chinese, Southeast Asian, and Korean.[110] When available, HLA-B*5801 genetic testing results will influence physicians' choice of ULTs in gout management, which has impact not only on rates of adverse drug reactions (ADRs), but also the long-term clinical outcomes and treatment costs of chronic gout. These long-term implications of HLA-B*5801 testing are not well envisaged, and are often neglected in evaluations of genetic testing.

An alternative risk mitigation strategy is enhanced safety monitoring of SJS/TEN symptoms for early drug withdrawal and SJS/TEN management. It has been shown that early withdrawal of causative drugs among SJS/TEN patients is associated with lower risk of dying.[111]

This study examined the incremental cost-effectiveness of six strategies, including those involving genetic testing and safety monitoring program, to mitigate the risk of allopurinol-induced SJS/TEN and to identify the conditions in which each strategy is incrementally cost-effective over a life time horizon, from the Singapore health system perspective.

3.3 Methods

A decision tree model was developed for a hypothetical cohort of gout patients who were eligible for allopurinol and probenecid, using TreeAge Pro 2013 (TreeAge Software, Williamstown) to evaluate incremental cost-effectiveness of five strategies over a 30-year time horizon. 30 years roughly represents the remaining life expectancy of gout patients, given an average onset age of 50,[112] and life expectancy of 80.[113]

Treatment strategies

The strategies modeled were: (a) Standard ULT with allopurinol as first-line drug (Standard ULT); (b) Standard ULT with allopurinol as first-line drug coupled to a safety program (ULT+SP). The hypothetical 3-month safety program (SP) comprised of one nurse-led patient education session on SJS/TEN, 6 fortnightly phone calls to check for early signs of SJS/TEN and a hotline for adverse reaction reporting and triaging for medical attention when needed; (c) HLA-B*5801 genetic testing-guided ULT treatment (G→ULT) in which patients received different first-line ULT based on test results (probenecid for test positive, allopurinol for test negative); (d) HLA-B*5801 genetic testing to enroll test positive patients in SP when initiating allopurinol (G→SP); test negative patients would receive allopurinol without SP; (e) HLA-B*5801 genetic test-guided ULT with the enhanced safety program (G→ULT→SP), in which test positive patients are initially given probenecid as in the G→ULT strategy, but non-responders

are subsequently switched to allopurinol and monitored via the enhanced safety program; (f) No ULT and treatment of acute flares only (no ULT) (Table 4).

Table 4. Components of six strategies

Strategy	1st line therapy	Genetic testing	Safety Program
ULT	Allopurinol	No	No
ULT + SP	Allopurinol	No	Yes
G→ULT	Allopurinol (for HLA-B*5801 negative patients) Probenecid (for HLA-B*5801 positive patients)	Yes	No
G→SP	Allopurinol	Yes	Yes (for HLA-B*5801 positive patients only)
G→ULT→SP	Allopurinol (for HLA-B*5801 negative patients) Probenecid (for HLA-B*5801 positive patients)	Yes	Yes (for HLA-B*5801 positive patients who do not respond to probenecid only)
No ULT	Treatment of Acute flares only	No	No

ULT, urate-lowering therapy; SP, safety program; G, HLA-B*5801 genetic testing.

Treatment sequence

Treatment sequence was based on international gout management guidelines and local clinical practices.[97,99,114] Response to ULT treatment was defined as achieving target SUA ≤6 mg/dl (360 μmol/l) and non-response referred to SUA > 6 mg/dl.[97] First-line ULT was assumed to be allopurinol at 300 mg/day.[115] As higher doses may be necessary to reach SUA target for some patients,[116,117] allopurinol up to 600 mg/day was modeled as next treatment step for non-responders. Probenecid (up to 2g/day) was modeled as the second-line treatment.[118]

Model structure

The decision tree in Figure 3 describes the treatment pathways. To mirror clinical treatment pathway, a titration period and a maintenance period were modeled. In titration period, patients on genetic-testing guided ULT strategy (G→ULT) received allopurinol if test negative or probenecid if test positive. Patients on other strategies except no ULT, received allopurinol. After 3 months, patients' response was evaluated and next step in the treatment sequence was initiated for non-responders, and those with side effects. In maintenance period, appropriate ULT identified in titration period was maintained over 20 years. When no appropriate ULT was identified, no ULT was given in maintenance period and only acute flares were treated.

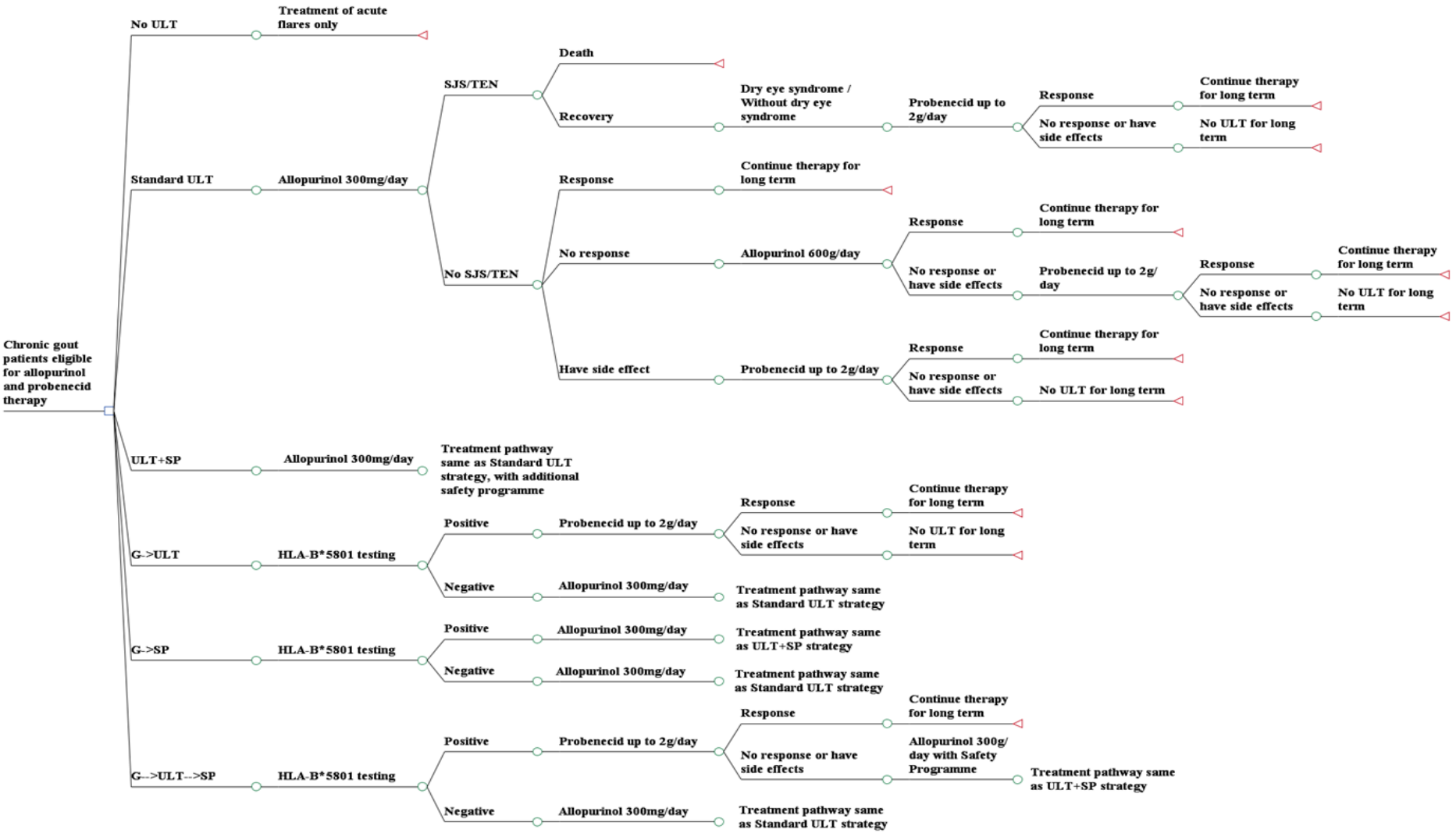


Figure 3. Decision tree model. ULT, urate-lowering therapy; SP, safety program; G, HLA-B*5801 genetic testing; SJS, Stevens-Johnson syndrome; TEN, Toxic Epidermal Necrolysis.

SJS/TEN may occur shortly after allopurinol exposure.[119,120] Most allopurinol-induced SJS/TEN cases occur within 60 days of allopurinol exposure,[119] and the average duration for SJS/TEN treatment is within 2 weeks.[65] Patients who develop SJS/TEN are therefore assumed to succumb within 3 months after allopurinol initiation or to recover after treatment. Various complications such as ocular complications may occur among patients recovered from SJS/TEN, and have long-term implications on quality of life and medical costs.[121,122] The common and lasting condition, dry eye syndrome, was modeled to account for the impact of SJS complications.

Model inputs

Model input variables and the sensitivity ranges are listed in Table 5. Quantification of key variables is described below.

Table 5. Input variables and sensitivity ranges

Variable name	Base case value	Minimum	Maximum	References
Cost, 2012 US$				
Cost of per case SJS treatment	3,477	1,738	6,962	[123]
Cost of per case SJS-TEN overlap treatment	10,254	5,123	20,500	
Cost of per case TEN treatment	17,031	8,515	34,062	
Cost of HLA-B*5801 genetic testing	270	135	404	Singapore Health Sciences Authority tissue typing lab
Average annual cost of allopurinol (daily dosage=up to 300mg)	33	16	66	Median selling price in public healthcare institutions in Singapore
Average annual cost of allopurinol (daily dosage=up to 600mg)	66	33	132	
Average annual cost of probenecid (daily dosage=up to 2g)	132	66	265	
Average drug cost of acute gout flare treatment (7 days)	22	11	92	
Cost of doctor consultation (per visit)	46	23	123	Public healthcare institutions in Singapore
Cost of safety program (per 3 month)	62	31	123	Cost estimate based on similar programs in public healthcare institutions in Singapore
Average annual cost to manage dry eye syndrome	200	100	800	[124]
QoL/Utility				
SJS (duration=8.9 days)	0.35	0.25	0.45	[123]
SJS-TEN overlap (duration=9.2 days)	0.3	0.2	0.4	
TEN (duration=12.4 days)	0.25	0.15	0.35	
Achieving SUA target *	0.7463	0.6463	0.8463	[125]
Not achieving SUA target *	0.7	0.6	0.8	
Utility discounting factor for dry eye syndrome	0.8	0.7	0.9	[126]
Treatment outcomes of ULTs (clinical response at 3 months post treatment initiation among patients with gout)				
Proportion of patients achieving SUA target with allopurinol daily dose up to 300mg/day*	0.38	0.2	0.5	[127,128]
Proportion of patients who achieve SUA target with allopurinol daily dose up to 600mg/day*	0.76	0.4	0.85	[129]
Proportion of patients who achieve SUA target with probenecid daily dose up to 2g/day*	0.68	0.4	0.85	[130]
Proportion of patients having side effects (excluding SJS/TEN) upon taking allopurinol	0.05	0.025	0.1	[130]
Proportion of patients having side effects upon taking probenecid	0.12	0.035	0.14	[131,132]
Annual number of flares experienced by chronic gout patients with uncontrolled SUA	4	2	10	Assumption
SJS/TEN fatality and incidence				
Incidence of allopurinol-induced SJS/TEN among patients who initiate allopurinol	0.002	0.001	0.004	[43,106]
Fatality of SJS	0.05	0.025	0.1	[123]
Fatality of SJS-TEN overlap	0.15	0.075	0.3	
Fatality of TEN	0.3	0.15	0.6	
Proportion of SJS among SJS/TEN	0.7	0.65	0.75	
Proportion of TEN among SJS/TEN	0.2	0.15	0.25	
Percentage of SJS/TEN patients developing dry eye syndrome	0.59	0.3	0.8	[121]
HLA-B*5801 genotyping				
Proportion of HLA-B*5801 carriers in the Singapore population (ethnicity-weighted)	0.185	0.1	0.3	[133]
Incidence of allopurinol-induced SJS/TEN among patients who initiate allopurinol for the first time	0.002	0.001	0.004	[106,134]
Effectiveness of safety program				
Percentage reduction in SJS/TEN mortality	0.3	0.1	0.8	[111]
Other inputs				
Duration modeled , years	30	10	40	Assumption
Annual discount rate	0.03	0.01	0.05	Assumption

*Predictive value of HLA-B*5801 genetic testing*

Prevalence of HLA-B*5801 carriers is 22.3%, 7.3% and 3.5% among Singaporean Chinese, Malays and Indians respectively, based on published allele frequencies and Hardy-Weinberg Equilibrium,[133] resulting in an ethnicity-weighted prevalence of 18.5%. Among Asian populations with SJS/TEN incidence data, Taiwan has the closest ethnic makeup to Singapore. The incidence of allopurinol-induced SJS/TEN in Singapore was assumed to be the same as Taiwan, or 0.2%.[134] Sensitivity and specificity of HLA-B*5801 test were assumed to be 100% and 85% respectively with resulting PPV of 1.52% and NPV of 100%.[134]

Safety program

Early withdrawal of causative drugs among SJS/TEN patients is associated with lower risk of dying (odds ratio 0.69 per day),[111] though early withdrawal may not stop disease progression.[135] We therefore assumed that the hypothetical safety program did not reduce the incidence of SJS/TEN but reduced SJS/TEN mortality by 30%.

Costs and utilization

In first year of treatment, patients are assumed to require four doctor consultations for ULT initiation and dose titration. Patients achieving satisfactory response with ULTs were assumed to continue life-time ULT treatment with the same ULT, and maintained satisfactory SUA levels. As hyperuricemia is a major risk factor for flares, patients meeting SUA target were assumed to have no flares in maintenance period and require two routine doctor visits annually. Patients who failed ULTs or had side effects were assumed to receive no ULT in the long-term, and have four flares on average, which were treated using colchicine, non-steroidal anti-inflammatory drugs (cyclooxygenase-1 and 2 inhibitors),or glucocorticoids. In addition to four doctor consultations for acute gout treatment, these patients were assumed to have 3 hospital admissions every 10 years. This estimate was based on a study which reported average number of hospital admission for gout or gout-related complications to be 1.5 over 10 year.[136] We doubled this number to reflect the higher accessibility of hospital care in Singapore. Costs of doctor consultations, ULTs, medications

for acute flares management, and gout-related admissions were obtained from public healthcare institutions in Singapore. Based on the bills of 11 gout-related admissions between 2012 and 2013, average cost per admission was around US$1,484 and the average length of stay being 3.36. All costs were displayed in US dollars with 1US$ equivalent to 1.27 Singapore dollars as of 2 October, 2014.[137]

Cost-effectiveness analysis

The total costs and QALYs associated with each treatment strategy were calculated over 30-year time horizon. QALYs is define as life years adjusted for QoL, and was calculated as

$$Total\ QALYs = \sum_{t=1}^{30} \frac{QoL_t}{(1+d)^{(t-1)}}$$

where QoL or utility score ranges between 0 and 1, with 1 indicating perfect health and 0 indicating death; d, annual discount rate is 3%; and t indicates years since treatment initiation. Incremental Cost-effectiveness Ratios (ICERs) was calculated as incremental cost over incremental QALYs.

Sensitivity analysis

To examine the robustness of results over various assumptions, one-way sensitivity and probabilistic sensitivity analyses were conducted. In the one-way sensitivity analyses, variable were varied within the sensitivity ranges, one at a time, and ICERs were generated (Figure 4). In the probabilistic sensitivity analysis, all variables were varied simultaneously, and the distribution of ICERs based on 10,000 repeated draws from assigned distributions were obtained. All variables were assumed to follow triangular distributions, with most likely values being base case; minimum and maximum being lower and upper bounds of sensitivity ranges.

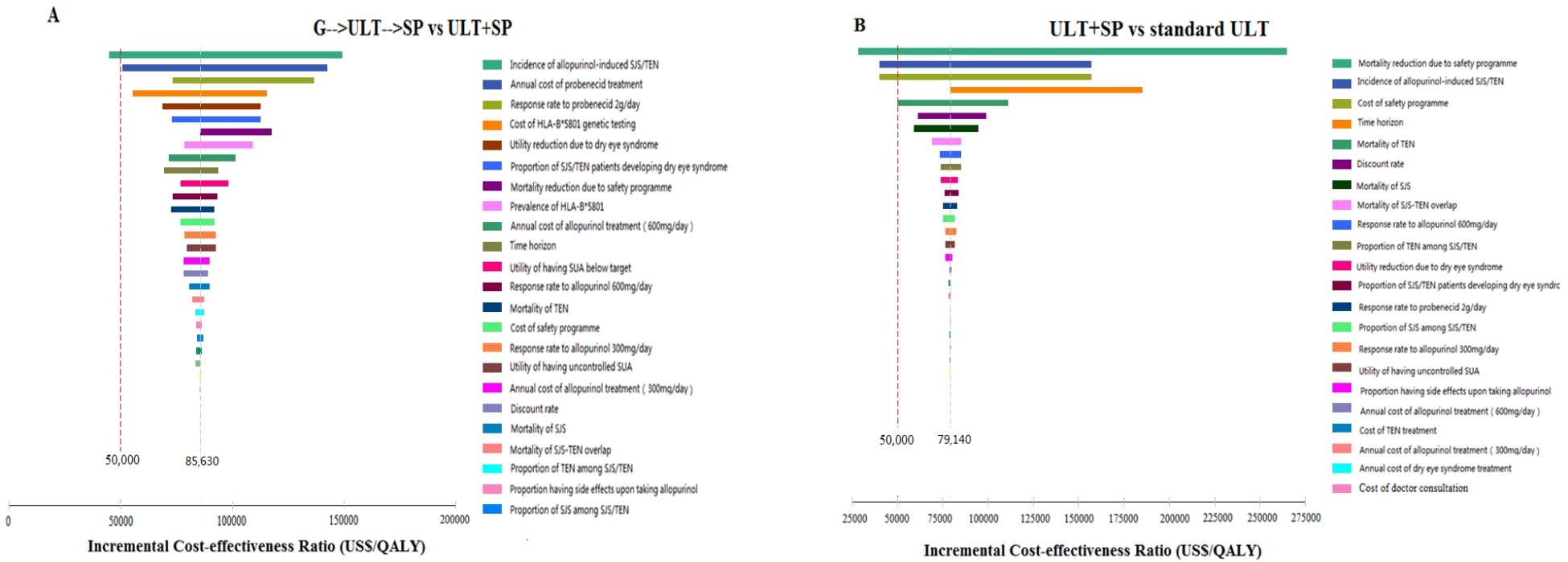


Figure 4. One-way sensitivity analysis of assumptions on the incremental cost-effectiveness ratio (ICER). (A) G->ULT->SP compared to standard ULT; (B) ULT+SP compared to Standard ULT. ULT, urate-lowering therapy; SP, safety program, SJS, Stevens-Johnson syndrome.

3.4 Results

Cost-effectiveness

Consistent with recommendations,[138] strategies were listed according to increasing order of costs, and ICERs were calculated regards to the next most costly strategy (Table 6). Standard ULT coupled to a safety program (ULT+SP) compared to standard ULT alone yields an ICER of US$79,140/QALY, relative to standard ULT. G→ULT→SP had an ICER of US$85,630/QALY compared to ULT+SP. Three strategies were dominated (more expensive and less QALYs than another strategy): genetic testing to enroll test positive allopurinol patients in SP (G→SP); genetic testing-guided ULT treatment (G→ULT); and no ULT. US$50,000 is a commonly used ICER threshold to identify cost-effective interventions.[86] This is very similar to the National Institute of Health and Clinical Excellence ICER threshold of £20,000-30,000,[49] which is approximately US$48,000 at the currency exchange rate on 2 October 2014.[137] In the base case, genotyping and safety program are both not cost-effective, by any of the commonly used cost-effectiveness thresholds.

Table 6. Cost-effectiveness of six strategies for ethnicity-weighted Singaporean patients requiring ULTs

Strategy	Cost ($)	Incremental cost ($)	QALYs	Incremental QALYs	ICER ($/QALY)	Dominance
Standard ULT	4,130	-	14.9966	-	-	Undominated
ULT+SP	4,200	60	14.9974	0.0008	79,140	Undominated
G→SP	4,420	220	14.9974	0	-	Dominated
G→ULT→SP	4,590	390	15.0020	0.0046	85,630	Undominated
G→ ULT	5,160	570	14.9597	-0.0423	-13,510	Dominated
No ULT	15,310	10,720	14.1319	-0.8701	-12,320	Dominated

ULT, urate-lowering therapy; SP, safety program; G, HLA-B*5801 genetic testing; QALY, quality-adjusted life years; ICER, incremental cost-effectiveness ratio.

Sensitivity analyses

One-way sensitivity analysis shows that cost of safety program, mortality reduction due to safety program, and the incidence of SJS/TEN were the most influential factors on

ICERs, and variations in all others inputs within the defined ranges did not alter the cost-effectiveness results (Figure 4). Probabilistic sensitivity analysis showed based on willingness-to-pay of US$50,000/QALY, allopurinol without genetic testing (standard ULT) is the preferred strategy in 40.7% of iterations, compared to 38.5% of iterations and 20.8% of iterations for ULT+SP and G→ULT→SP, respectively. (Figure 5)

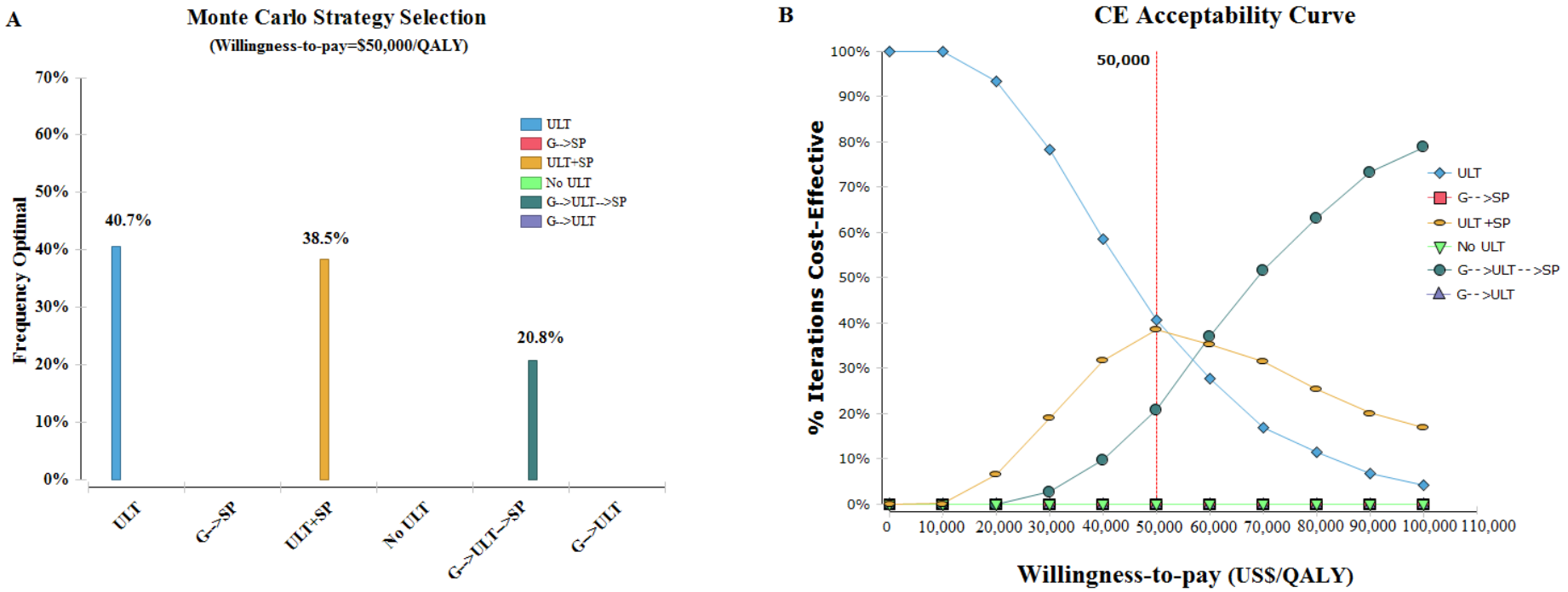


Figure 5. Probabilistic Sensitivity Analysis. (A) Strategy Selection at Willingness-to-Pay of $50000/QALY. (B) Cost-effectiveness acceptability curves of six strategies. ULT, urate-lowering therapy; SP, safety program; G, HLA-B*5801 genetic testing; QALY, quality-adjusted life years.

3.5 Discussion

This study compares six potential risk mitigation strategies in chronic gout management. Among all strategies, no ULT resulted in the lowest QALYs, and surprisingly the highest long-term cost. This is because high SUA levels result in more frequent flares, and higher costs due to flare treatments and hospitalizations. This suggests that forgoing ULT treatment because of the fear of SJS/TEN risk would results in worse outcome and higher life-time gout treatment costs.

Genetic testing strategies

The three strategies involving genetic testing were either dominated or not cost-effective under the base case scenarios. Genetic testing-guided ULT treatment (G→ULT) incurs additional testing costs and higher drug costs as the alternative drugs for HLA-B*5801 positive patients are more expensive than allopurinol. Paradoxically, if allopurinol is completely avoided among test positive patients, these patients have lower QALYs as they are restricted to fewer alternative ULT options, and consequently, will have poorer SUA management outcomes. Patients who test positive of HLA-B*5801 and fail to respond to probenecid would receive no ULT in the long term and have more frequent flares when they might have benefitted from allopurinol. Given the HLA-B*5801 prevalence in the Singapore population (18.5%) and the low PPV of the test (1.52%), a G→ULT strategy, in which genetic test results dictate the selection of the initial ULT, will switch 18.5% of patients away from allopurinol when only 1.52% of them would be expected to develop SJS/TEN.

We also considered whether genetic testing might be a useful tool for prioritizing high-risk patients for an enhanced safety program upon allopurinol initiation (G→SP) when it is operationally challenging to enroll all gout patients in safety program in busy clinic settings. We found that G→SP is more expensive, as the current cost of genetic testing (US$270) is relatively high compared to that of an enhanced safety program (US$63). If the cost of genetic testing drops to US$23, G→SP achieves the

same cost as the safety program for all patients. At this cost, genotyping may be a useful strategy in busy clinics to screen patients and prioritize monitoring resources to the most at-risk patients.

As completely avoiding allopurinol in test positive patients is not favorable from cost-effectiveness perspective, the alternative strategy G→ULT→SP provides an option for clinicians who would like to achieve good SUA control with a lowered risk of SJS/TEN. This involves using probenecid first in test-positive patients, before embarking on allopurinol therapy with an enhanced safety program for those who do not respond to probenecid. This strategy has an ICER of US$85,630/QALY and is not cost-effective at an ICER threshold of US$50,000. However, it would become cost-effective if the cost of the genetic test drops below US$90, which is possible.

The main reason why genetic guided ULT selection reduced QALYs is the limited alternative options. Febuxostat, widely used in Europe and USA as a alternative drug, is not readily available in Singapore, and the current cost is 40 times higher than allopurinol. However, when febuxostat was modeled as third-line ULT, genetic testing-guided ULT still yields fewer QALYs than standard ULT and at higher cost, which implies that using allopurinol, probenecid and febuxostat to optimize treatment for patients achieves higher overall response rate than using probenecid and febuxostat only. Moreover, hypersensitivity reactions associated with febuxostat, including Stevens-Johnson syndrome, have also been reported.[139]

In contrast to our results, Saokaew [140] et al. concluded genetic testing is very cost-effective (ICER=US$5,062/QALY) in preventing allopurinol-induced SJS/TEN in Thai population. [140]The divergent findings result from differences in 1) treatment costs, 2) incidence of allopurinol-induced SJS/TEN in the respective populations, 3) population frequency of HLA-B*5801, and 4) assumptions on gout treatment outcomes. To the latter point, whereas Saokaew et al. didn't distinguish responders and non-responders to ULTs, we assigned different QoL and treatment costs to the two groups,

which we believe is more realistic. This is a key difference that generated the divergent results.

In addition to SJS/TEN, HLA-B*5801 is also associated with other adverse skin reactions such as Drug Rash with Eosinophilia and Systemic Symptoms (DRESS).[106] Data on incidence, costs of treatment, and long-term complications of DRESS are scarce. As an approximation of incidence, we examined the number of DRESS cases in the national Singapore voluntary adverse drug reaction database. Between 1993 and 2014, DRESS constituted 30% of all serious cutaneous adverse reactions (SCAR) associated with allopurinol. In the base case model, we assumed that the incidence of SJS/TEN alone is 0.2%; an estimate of the combined incidence of DRESS and SJS/TEN therefore is 0.28%. As noted above, the ICER of the G→ULT→SP strategy only drops below the cost-effectiveness threshold when the incidence is higher than 0.35%.

Safety program

Based on the commonly cited cost-effectiveness threshold of US$50,000/QALY, enrolment of all gout patients into safety program when initiating allopurinol is not cost-effective compared with ULT alone, under base case assumptions.[86] However it would become cost-effective compared with standard ULT, if cost of safety program were reduced to below US$39 per patient, or if safety program resulted in over 47% reduction in SJS/TEN mortality, or 24% reduction in SJS/TEN incidence. In fact a 30% reduction in mortality assumed under base case may not fully capture the benefits of safety program, which may reduce seriousness and costs of treating other adverse reactions reported within the wide clinical spectrum of AHS.[141,142]

Study limitations

This study has several limitations. First, the study presumes the efficacy of ULTs in the short term continues for the long-term when metabolic changes, comorbidities and other medical therapies may ensue with aging. Second, long-term

complications of SJS/TEN and other side effects of gout management may be underestimated due to lack of long-term data. However, sensitivity analysis on treatment costs showed the results are robust even when the costs were doubled. Third, the treatment outcomes modelled were based on published clinical studies. In practice, effectiveness of ULTs may be lower than controlled circumstances; as monitoring of SUA and up-titration of drug dosage may not be universally performed, and patients' non-adherence to ULT is an issue.[143-145] Nonetheless, as these factors pertain to ULT treatment in general, conclusions on genetic testing and safety program are not likely to be influenced. Finally, this study is likely to underestimate the benefits of HLA-B*5801 genetic testing in reducing the mortality and morbidities attributable to other serious cutaneous adverse reactions.[141]

3.6 Conclusion

Complete avoidance of ULT due to the fear of SJS/TEN in chronic gout management results in the worst outcome and highest long-term costs. An enhanced safety program for all patients initiating ULT may become cost-effective if program costs are low or if significant mortality reduction can be achieved. HLA-B*5801 genetic testing for all gout patients commencing ULT, if used to avoid allopurinol in all test-positive patients, reduces the overall QALYs at a population level. Test positive patients (18.5%) would have fewer alternative treatment options, and thus worse gout outcomes, while SJS/TEN would be avoided in 1.5% of patients. HLA-B*5801 genetic testing and prescribing probenecid in test-positive patients initially, but switching non-responders to allopurinol coupled with an enhanced safety program, although not cost effective currently, would become cost-effective if testing costs drop substantially. Our results do not preclude individuals from seeking genetic test should they choose to do so nor implementation of safety program for extra clinical vigilance, only that the use of public resources is not justifiable based on existing data and thresholds for cost-effectiveness.

Chapter 4 . Introduction to patients' preference for using pharmacogenetic testing to reduce severe adverse drug reactions

4.1 Introduction

Severe adverse drug reactions (ADRs) have long been a medical and public health concern. With the advancement of genetic research, genetic testing has been shown promising to select drugs for safer gout treatment.[106] Cost-effectiveness analysis (CEA) described in Chapter 3 provides information on the value of HLA-B*5801 testing from the health system. The negative cost-effectiveness results suggests that implementing HLA-B*5801 testing at the system level will not bring high value from the public resource allocation perspective. However, this does not necessarily imply that individual patients and doctors should not test. In fact, cost-effectiveness analyses are not aimed to answer the question whether patients should or would use the test. The adoption of genetic test is a complex issue, concerning the interplay between various stakeholders (patients, physicians, providers, payers, regulators). Patients are the consumers, and often times the payers too. Patients' preferences are therefore crucial to determine the uptake of HLA-B*5801 genetic test and inform testing policies.

This chapter is an introduction to patients' preferences for using pharmacogenetic testing to reduce risk of severe ADRs. It motivates the empirical study in Chapter 5, and facilitates the formulation of research questions and hypotheses. This chapter starts by outlining the importance of understanding patients' preferences, and then reviews the theoretical framework to analyze patients' preferences, followed by the literature on patients' attitudes towards genetic testing. I then review the evidence on the determinants of patients' preferences for genetic testing, with a focus on the methods using which these determinants were studied. At

the end of this chapter, I summarized the research question and hypotheses, which are studied in the next chapter.

4.2 Why is patients' preference important?

4.2.1 Why may individual preferences for HLA-B*5801 testing differ from assessment at the health system level?

Cost-effectiveness analysis assesses the benefit and cost of genetic testing to the health system, and applies a societal willingness-to-pay to determine whether a service is of high value from a public resource allocation perspective. However, individual patients may not go through the same process in their decision making, and significant heterogeneity can be expected. Patients' decisions may be different from system level cost-effectiveness analysis for several reasons.

First, cost-effectiveness applies a threshold (such as $50,000/QALY to define cost-effectiveness),[86] which is meant to represent the societal willingness-to-pay for one quality-adjusted life year (QALY). However, there may be individual variations in the perceived value and benefit of pharmacogenetic test, and therefore the worthiness of testing. The willingness-to-pay may also correlate with individual's ability to pay and other socio-demographic characteristics.

Second, when uncertainty is involved, the decision making usually deviates from expected value calculation, as used in CEAs. In CEAs, expected reduction in utility due to SJS is calculated as the chance of SJS multiplying by the utility reduction associated with SJS, which is consistent with the expected utility theory.[146] As the chance of SJS is only 0.2%, the adverse negative impact of SJS at the population level is small. Despite the low incidence, life-threatening adverse drug reactions, is a big safety concern among some patients and physicians. In prospect theory, the probabilities of outcomes happening are transformed into decision weights, which can be thought as the decision maker's perception about the probabilities. [147,148] Therefore the utility can be written as

$$U = \sum_{i=1}^{n} w(p_i)v(x_i)$$

Where $w(p_i)$ is the decision weight of probability p_i, and $v(x_i)$ is the valuation of outcome x_i. For small probability events, such a life-threatening adverse drug reaction, individuals tend to overvalue the small probability (ie. $w(p_i) > p_i$) Therefore, individuals may have a high willingness-to-pay to avoid the small chance of developing SJS.

Third, the judgement of cost-effectiveness is relative to a comparator which is often the current practice. In the CEA described in the previous chapter, it is assumed the status quo to be allopurinol treatment for all eligible chronic gout patients, based on clinical guidelines. In reality, some doctors and patients are not comfortable with prescribing or taking allopurinol knowing the risk of SJS. The fear may results in lack of ULT treatment, which the cost-effectiveness analysis showed to be the most costly and least effective strategy. Genetic testing, in addition to reduce risk of SJS, may also improve gout control, due to the more confident use of allopurinol. Therefore, the actual benefit of genetic testing may be higher than modelled in CEA. However without data, these cannot be precisely quantified.

Fourth, cost-effectiveness evaluates the benefit of testing in terms of the clinical utility, which is the potential of the test results to improve treatment outcome.[149] Consequently, those who have negative test results, which will not alter their treatment, receive no health benefit from testing. However, from patients' perspective, there may be a "value of knowing",[150] which is, those who test negative derive utility from the assurance that they are not at risk.

Cost-effectiveness analysis offers a convenient and standardized tool for policy makers to efficiently allocate scarce public resources among competing needs to achieve the most value or best health outcomes within budget constraints. The fact that HLA-B*5801 is not cost-effective suggests public resources spent on reimbursing the test would not achieve high value for the health system compared to a cost-effective intervention. However question remains whether services should be made available for

voluntary test, and whether clinical practice guidelines should encourage routine genetic testing.

4.2.2 Why is understanding patients' preferences important for medical practices and policy-making?

Knowing the availability of the genetic test, there is an urgent demand for information and guidance on the use of genetic test and genetic-test guided treatment regimen from the physician community. However, existing evidence does not provide sufficient indications on the appropriate use of genetic test. To inform policy making, we sought to understand patients' preferences. Patients, as the consumer of medical services, derive utility from the services, even though often times the treatment is not directly chosen by patients. Treatment effectiveness, side effects, financial burden, and care experiences all influence patient's utility. Minimizing risk of severe side effect does not necessarily maximize patient's utility if the treatment effectiveness is compromised, or if significant financial burden is incurred. Patient's preference information are useful to physicians for several reasons. Firstly, knowing the tradeoffs can enable physicians to communicate risk and mitigation strategies more effectively, and choose the most suitable treatment based on each patient's medical profile and preference profile. Secondly, inaccuracy (such as false positive, false negative results) of test can lead to difficult medical decisions, as false results may lead to suboptimal treatment. Directly eliciting patients' preferences on accuracy parameters allow the identification of the maximum acceptable risk, which can facilitate clinical decision making.

Patients' preferences are also useful for other stakeholders. From a service provider perspective, a forecast of uptake rate is desirable to facilitate operation planning, and price setting. From the regulator perspective, understanding patients' preferences can inform the formulation of risk communication letters to health professionals, and revise drug package insert to incorporate genetic information and usage advice. Knowledge on patient's preference could also inform the design of

effective vigilance and risk minimization programs. In many cases, the uptake rates for public health-promoting programs are below target. For instance, colorectal cancer screening is promoted, and reimbursed in the United States, yet around 50% of individuals older than 50 years have never been screened.[151] This implies some preference or motivation factors are underlying the screening decision. Without understanding the decision making process, information provision and financial incentive may not be successful in meeting the intended uptake target. Patients' preferences can also inform the research and development of genetic tests. Identifying what test features patients value most, and understanding patients' willingness-to-pay for test features can facilitate the development of more useful tests.[150] In a broader context, understanding patients' preferences may help to set agenda and prioritize pharmacogenetics research.

Recently, patients' preferences have garnered more attention, and have been increasingly considered by the medical community and regulators. In clinical guidelines, patients' preferences are often mentioned, especially in situations when tradeoffs between risk and benefit are involved.[152] One area of application by regulators is to weigh the benefit and risk for new drugs and medical devices.[153,154] For instance, some effective treatments may be associated with risk of life threatening side effects. The regulatory and clinical perspective is usually to minimize risk or weigh the benefit and risk, which often runs into difficulty, as it is unclear how therapeutic benefits and risk of side effects should be traded off. Patients may be willing to accept higher risk of severe side effects in exchange for better treatment outcome, especially for conditions with limited alternative therapeutic options. Measuring patients' preferences is one potential solution to quantify the tolerable risk in exchange for better disease management. US FDA has published guidance on the use of patients' reported outcome in regulatory decisions in 2010, and released a draft guidance on patients' preference information in 2015.[60,154] The guidance reviews the methods to measure patient's preferences, and the use in premarket approval applications (PMA), Humanitarian Device Exemption (HDE), and *de novo* review processes.[154]

4.3 Patients' general attitudes and preferences for the use of pharmacogenetic testing to reduce severe ADRs

Literature on the preference for using pharmacogenetic testing to reduce severe ADRs is relatively new and limited, mainly because the basic science and clinical evidence were only developed in the last decade, and not yet widely applied clinically. Nevertheless, qualitative and quantitative studies on patients' attitudes towards pharmacogenetic testing revealed wide public interest.

Qualitative studies on patients' perceptions about pharmacogenetic testing identified the lack of prior knowledge on pharmacogenetic testing.[155] However, when educated about the definition and applications of pharmacogenetic tests, the public were generally enthusiastic towards pharmacogenetic testing.[155-157] In a phone interview with 328 German patients with asthma or chronic obstructive pulmonary disease, 96% of patients appreciated the availability of pharmacogenetic tests, and claimed the willingness to take a test prior to receiving asthma medication.[155] In this group, the ability of the test to avoid side effects is an important consideration, and majority of patients were worried about the possibility that the test could not find the suitable drug with best therapeutic outcome and lowest risk of side effects. Similarly, a random-digit-dial telephone survey of 1,139 US adults showed that 85% of respondents were willing to take a pharmacogenetic test to predict serious side effects.[157]

A few studies have examined patients' preferences for pharmacogenetic tests in specific clinical scenarios quantitatively, and also confirmed patients' preferences for taking genetic test to reduce risk of adverse drug reactions. Payne et al studied patients' preferences for using pharmacogenetic test to identify the side effect neutropenia associated with the immunosuppressant azathioprine.[158] In the study, various dimensions of test were listed, and patients were found to pay significant attention to the predictive accuracy of the test (ie. the ability of the test to predict risk of side effect). Herbild *et al.* measured Danish populations' preference for pharmacogenetic testing prior to depression treatment, and found that patients were willing to pay a significant amount

of money to avoid change of medication due to lack of effectiveness or unacceptable side effect.[159]

4.4 Determinants of preferences for genetic testing to reduce risk of severe ADR

4.4.1 Approach

To identify the determinants of patients' pharmacogenetic testing decisions, three synergistic approaches were adopted. Firstly, I reviewed the conceptual models of health behaviors, the determining factors outlined in the model, and operationalized these factors in the context of allopurinol pharmacogenetic testing in Singapore. Secondly, the empirical literature on determinants of patients' attitudes and preferences for pharmacogenetic testing and other screening services are reviewed. Lastly, the identified factors were verified via in-depth interview with diabetes patients.

4.4.2 The health belief model

Various models have been proposed to explain health behaviors, such as the acceptance of screening or preventive services. Some commonly used models are the Health belief model (HBM)[160,161], Anderson's health behavior model[162,163], and the theory of planned behavior[164]. The Health Belief Model was used here to conceptualize individual patients' genetic testing decisions.

Health Belief Model is one of the most commonly used models to explain and predict individuals' health behaviors. It was first developed in the 1950's and 1960's by Rosenstock *et al.* at the United States Public Health Service to explain the series of failures of programs to promote disease preventives or screening tests for tuberculosis (TB), cervical cancer, dental disease, rheumatic fever, polio and influenza, even though these services were provided free of charge or at very low cost for demonstration.[160,161] HBM focuses on individual-level belief and decision making, and assumes the decision makers to be rational.[165] HBM has outlined six key variables that will determine whether an individual will take preventive actions.[161] Four variables concerning individual's

perceptions are: 1) perceived susceptibility to disease, 2) perceived seriousness of disease, 3) perceived benefits of taking actions, and 4) perceived barriers to taking actions. Two variables to trigger the actions are 5) cues to action, and 6) self efficacy. Besides the six key considerations, there are also modifiable factors, which can modify the perceived threat, and benefit of taken the action, and subsequently influence the likelihood of health behaviours. Using the HBM framework, I identified the factors that may determine whether or not an individual will take a genetic test before initiating allopurinol treatment to avoid potential life-threatening adverse reaction SJS that can be induced by allopurinol (Figure 6). Empirical evidence on each factor was also briefly reviewed.

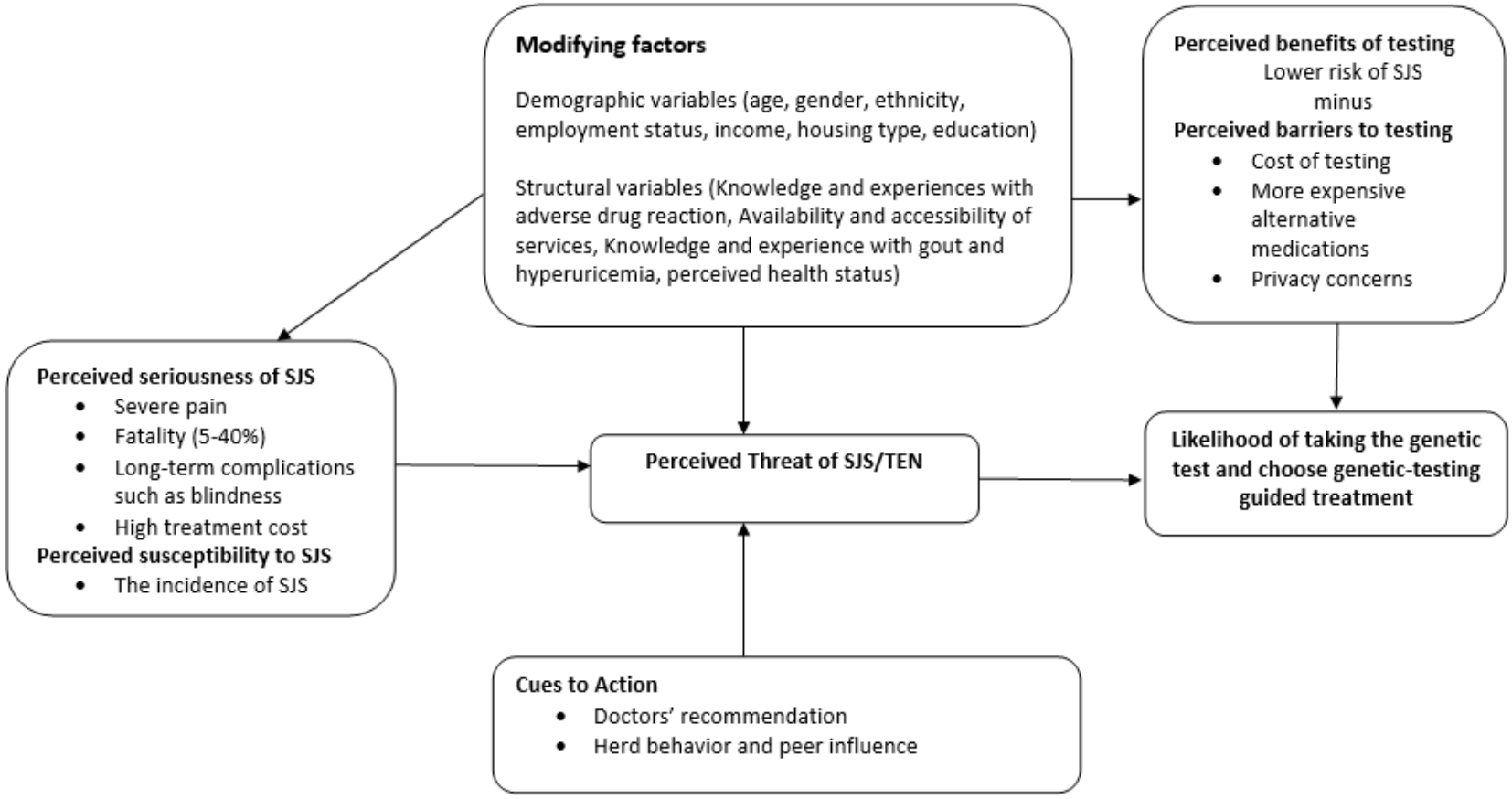


Figure 6. Determinants of genetic testing decisions identified using the Health Belief Model framework.

4.4.3 Literature review and conceptualization of determinants of pharmacogenetic testing decisions in the health belief model (HBM) framework

The perceived susceptibility of the perceived seriousness

Perceived threat of disease is the main motivation to take a screening test. The level of threat depends on the seriousness of disease and perceived individual susceptibility. Haga *et al*. found in a phone survey that more US individuals were more interested in using pharmacogenetic testing to predict serious side effect than mild side effects (85% vs 73%).[157] Hall *et al*. compared the preferences of the general public and a high risk population (the Jewish population) to test for Tay Sachs disease, and discovered that Jewish respondents were more likely to be tested.[166]

SJS/TEN are serious conditions that have an average mortality of 10% (5%-40%), cause severe pain during onset, and may have long-term sequelae such as dry eye syndrome and blindness.[16,17,122] Moreover, SJS/TEN treatment is costly, mainly due to hospitalization and the use of antibiotics. The seriousness of the condition is the primary motivation for taking the genetic test to predict the risk of SJS/TEN, and select appropriate drug to minimize SJS/TEN.

The susceptibility can be best quantified by the likelihood of developing SJS/TEN upon initiating allopurinol. Among Taiwan Han Chinese patients receiving allopurinol treatment, around 0.2% would develop SJS/TEN.[106,167] 0.2% is a small probability that individuals do not commonly encounter in everyday life. It's unclear how individuals interpret their susceptibility. Psychology and behavioural economics evidence suggests that when very small risk is involved, individual may exaggerate the probability or neglect the probability in their decision making.[147,168-170] It is therefore not clear whether or not individuals are concerned about this level of risk, and motivated to take preventive actions.

Perceived benefit of taking actions

The major benefit of testing prior to initiate allopurinol is to reduce the risk of allopurinol-induced SJS/TEN. The benefit of testing can be measured by the reduction in the risk of SJS/TEN. The extent of risk reduction depends on the accuracy and predictive power of the test. Indeed, many studies on patients' preferences on diagnostic tests have highlighted the importance of test accuracy in influencing patients' testing behaviours using various accuracy indicators. Hall *et al.* studied the false negative rate of genetic test, defined as the chance that someone carries the risk gene when the test is negative, and found that higher false negative rate significantly discouraged testing.[166] Knight *et al*. also used the false negative rate as an indicator for accuracy for a colorectal screening test.[171] Payne *et al.* varied the predictive accuracy (defined as the ability of the test to predict the risk of the side effect) in their study, and discovered that patients were willing to compromise test experiences (eg. waiting longer) for a small improvement in predictive accuracy of test.[158] Marshall *et al.* examined patients' decisions on colorectal cancer screening, and found the sensitivity and specificity of screening test to be crucial information.[172]

Despite patients' strong preferences for more information on test being provided,[158] it has long been recognized that the framing of risk and accuracy information can influence patients' perceptions and decisions.[173] For instance, there are several ways to describe the accuracy of the HLA-B*5801 test for allopurinol-induced SJS/TEN: (a) the test-guided treatment can reduce the risk of SJS/TEN from 0.2% to almost 0; (b) the test-guided treatment can reduce the risk of SJS/TEN by more than 99%; (c) the test have a sensitivity of 98%, and a specificity of 95.8%; and (d) the test has a 0% false negative rate, but a 98% false positive rate. All statements are true, yet people may react differently.[169] It is therefore worth considering the appropriate form of accuracy information communication, and the impact of framing on the responses elicited. Though technical terms such as sensitivity, specificity, false positive/negative rate, and positive/negative predictive value are often used to describe the accuracy of test,

laypeople and even health care professionals may not understand the precise meaning of these terms. In the case of HLA-B*5801 test, the most straightforward and objective indicator of accuracy is the risk of SJS/TEN with and without testing (format a). When both probabilities are provided, respondents can easily visualize the absolute magnitude of risk reduction, without the need to understand technical jargons or undertake additional calculations. Format b can potentially be misleading as it emphasizes on the relative level of risk reduction (over 99%), but neglects the fact that the incidence of allopurinol-induced SJS/TEN is low (0.2%) even without testing. Format c and d are not easily comprehensible by laypeople, and the need to understand technical jargons may increase the cognitive burden of making testing decision.

Other benefits of testing can result from the reduction in SJS/TEN risk, such as lower risk of dying, lower chance of having high medical expenditure to treat SJS. In addition, some literature suggests the "value of knowing" regarding the utility of testing, where even no treatment or preventive actions are involved after the test, knowing the test result has value.[149,174,175]

Barriers of taking actions

One type of barrier was cost. The HLA-B*5801 test currently costs S$375 in Singapore. Compared to allopurinol treatment cost of around $200 per year, test cost is high. Taking an expensive test for an inexpensive medicine may be a barrier to the uptake of genetic testing. The cost of long-term gout treatment depends on the genetic test results. Test positive patients require alternative drugs that do not induce SJS/TEN, but are significantly more expensive. Depending on the choice of second line drug, and the dosage, the medication cost can be twice to ten times the cost of allopurinol. Gout is a chronic condition, requiring long-term management. Switching to a more expensive medication may incur significant long-term cost.

Empirical evidence shows patients are sensitive to price when making medical decisions.[166,171,176] Various structural factors such as government subsidy and insurance reimbursement directly alter the out-of-pocket cost, which is the part of price that patients

pay. A study on the actual use of colorectal cancer screening services revealed that those with insurance coverage were more likely to attend screening.[151] Government subsidy and insurance coverage are therefore possible ways to remove the cost barriers of testing.

In addition to cost, there are other test-process related barriers. At the moment, HLA-B*5801 can only be done in centralized laboratories. Therefore the logistics is inconvenient, and the waiting time to receive test results is relatively long. In addition, patients will not receive urate lowering therapy before test result is received. An additional clinic visit may be required for test result pickup and prescription filing.

Ethical concerns for genetic testing has long been recognized.[177-179] For those genetic tests used to predict future disease risks, knowing the information may have negative impact because patients and family members worry about unfavorable results, especially for diseases without a cure or a prevention strategy.[166,174] The availability of genetic predisposition to insurers may lead to discrimination against the insured.[180,181] However, for pharmacogenetic testing, which has more defined clinical utility (ie. to guide drug selection, and dosage adjustment), empirical studies find relatively low level of ethical concerns. A phone survey of a sample of the U.S. public found that 90% of respondents were extremely or somewhat comfortable to share their pharmacogenetic test results with other doctors involved in their care management. 70% of respondents felt comfortable with incorporating their pharmacogenetic test results into their personal record. A survey targeting German patients revealed that only 27% of respondents were very or slightly worried about results sharing with insurance companies.[155]

Cues to action

A decision maker not only evaluates the benefits and harms of testing, but can also be influenced by the information cues. Medical decisions are not made in isolation. Even when information on treatment options is provided, patients usually seek other sources of information such as doctor's recommendation, media information, internet, or

opinions of family, friends, and the other peer patients. Among these cues, the research on doctor's recommendation and herd behaviour were extensively studied.

Doctor's recommendation

Studies have demonstrated the power of doctor's recommendation in influencing patients stated choices as well as actual behaviors on choosing treatment options, taking up screening tests or vaccinations.[166,182-187] Doctor's recommendation is one of the most important factors in patients decision making, and an experiment that randomly assigned recommendations led patients to choose an option that was obviously suboptimal.[188] Though patients are encouraged to make informed decisions on their own treatments, a survey on patients preferred role in medial decision making revealed that even though nearly all respondents preferred to know the different options, half prefer to leave the final decision to their physician.[189] The extent to which patient prefer decision making by physicians also vary by gender, education, and health status.[189]

In fact, physicians' preferences for pharmacogenetic testing are more extensively studied than patients' preferences for its importance in shaping behaviors. In general, physicians have positive believes that pharmacogenetic test may improve patient care by personalizing treatment for patients, and anticipate increased clinical usage.[155,190] However, even some recent studies revealed the lack of genetic testing knowledge and training among physicians. A survey of 260 US specialist and primary care physicians in 2010 identified that 40% to 72% of them had "no to minimal knowledge" on genetic topics, and were not certain how to incorporate genomic medicine into their practice.[191] Another national survey of a sample of US primary care physicians in 2011 showed that only 13% of responding physicians were comfortable ordering pharmacogenetic tests.[192] Therefore for successful implementation of test programs, physician education is crucial.

Different forms and strength of physician recommendation may have different impact. Among the studies reviewed, both general recommendations (recommend a behavior such as screening) and specific recommendations (recommend a specific test) improves the test uptake.[166,187] Stronger recommendations is associated with higher

uptake rate. [182] [186] A study to examine the relationship between strength of recommendation and HPV vaccination status revealed that, when the strength of doctor's recommendation was rated on a 1 to 5 scale, there is a 4-fold difference in the likelihood of vaccination between those receiving a strong recommendation and those receiving a weak recommendation.[182] Recommendations on the timing of test and location of test may also influence patients' decision.[184,186,187]

Patients consider physician recommendation important for many reasons, in a patient's survey by Gurmankin *et al*., the most common reasons of following the doctor's recommendations are: "physicians had important additional information", "physician had information about my risk that went beyond the data given in the question", "physicians know best" "I don't like having the responsibility of making my own medical decisions" "I don't trust myself to make the right decision".[188] Recommendation by physicians indicates the quality of a treatment option. Following doctor's recommendation may therefore be a decision heuristics that allows easy and fast decision making.[193,194] Though deviates from the "rational" decision making pathway based on logic and calculation, some empirical findings and economic theories have recognized the presence and advantages of decision heuristics.[193-195]

Herd behavior

Herd behavior has been recognized as another decision heuristic or shortcut. Herd behavior describes the trend that individuals' decisions tend to be influenced by what people around them are doing. Several related concepts are "following the herd" and "social conformity". It was first recognized by psychologist Soloman, and then widely observed in psychology, economics, consumer behaviors and finance.[194-199]. Though relatively fewer studies were conducted to understand patients' herd behavior in medical decisions, several studies have demonstrated the presence of herd behavior in fertility choices, and physician's prescription behaviors.[166,200-204]. One choice experiment by Hall *et al*. attempted to quantify the effect of providing information on other people's decision about genetic carrier screening on individual respondent's decision.[166] When informed

that “80% of people like you have been tested”, respondents were more likely to test, all else equal.

Following the herd is a simple decision heuristic, especially when health decision is difficult, and the optimal choice is unclear. Banerjee argues that in a sequential decision model, it is rational for decision makers to look at the decisions made the previous decision makers, as other decision makers may have information that is important. Moreover, he demonstrated that the optimizing strategy is to do what other people do, rather than using their information.[195] Carlsson developed an economic model of environmental conformity for the consumption of eco-friendly coffee. The key assumption is that individuals derive utility not only from consumption, but also from following certain social norms. Carlsson modeled the utility from eco-friendly coffee as the sum of direct utility from consumption and a self-image component.[205] The self-image of the individual can be negatively influenced by the difference between the product chosen by the individual and the social norm (eg: when 90% of people choose the eco-friendly coffee, consuming eco-friendly coffee is the norm).

Self-efficacy

Self-efficacy refers to one’s confidence in his or her ability to take the action and overcome the barriers. Studies suggest that decision are useful to the extent decision maker have the confidence to adequately implement the behaviour.[206] Self-efficacy is therefore important to determine the actual health behaviours.

Modifiable factors

In the Health Belief Model, another set of variables are the modifiable factors, which can influence individual’s perceived threat of disease, perceived benefit of action, and therefore the likelihood of action. Such factors include socio-demographic factors (eg: age, gender, ethnicity, employment status, income, housing type, education), and knowledge and experiences with gout, and genetics. Empirically, decision maker characteristics have been found to influence the attitudes towards genetic testing. Those

with higher income and education are found more likely to attend regular colorectal cancer screening.[151]

A summary of important variables

In summary, various determinants of testing decisions have been identified and summarized in Table 7. These factors were investigated in in-depth interviews with Singapore patients.

Table 7. Determinants of testing decisions

Concepts/ Domains	**Attributes/Factors**	**Related attributes**
Perceived threat	Risk of SJS/TEN	
	Severity of SJS/TEN	
Perceived benefit	SJS/TEN risk reduction (Some related factors are :	Accuracy and predictive value of test (sensitivity, specificity, false positive rate, false negative rate, positive predictive power, negative predictive power) Fatality reduction Cost saving
Barriers of testing	Cost of test	Availability of insurance reimbursement, government subsidy
	Cost of long-term gout treatment	
	Convenience of testing	
	Patients' privacy	
Cue to action	Doctor's recommendation	
	Herd behavior	
Self-efficacy	Self-efficacy	
Modifiable factors	Socio-demographic background	
	Knowledge and awareness of test	
	Knowledge of gout	

4.5 In-depth interview to gauge patients' opinions

After identifying the determining factors in the literature, in-depth interviews were conducted with 5 patients to understand laypeople's attitudes towards pharmacogenetic testing, and the decision making process. Individual interview was chosen over focused group, in order to understand each respondent's independent perception and valuation of genetic testing while minimizing the impact of peer respondents. The in-depth interview was aimed to achieve 3 objectives: 1) understand respondents' general perceptions and attitudes towards genetic testing; 2) verify whether respondents consider the pre-identified attributes to be important for their decision to adopt genetic testing; and 3) identify other important factors that were missing.

A structured interview guide was designed to guide the in-depth interview. (Appendix A) The guide included an introduction to gout and pharmacogenetic testing, a section on general preferences for genetic testing, and considerations on various test outcome features identified in the literature review. Then respondents were then asked to share their thoughts on the role of doctor's recommendation, and most common choice when making a testing decision. The perceptions and expectations about test service delivery process and use of genetics data were also elicited. Interviewer asked the guiding questions, and allowed respondents to share their opinions freely. Specific questions on the guide that were not answered by respondent in the previous step were asked again as a probe. Respondents were also given the opportunity to share other important factors that were not raised by interviewer.

Consistent with the literature, respondents were generally receptive to the idea of using a genetic test to reduce the risk of severe adverse drug reactions. Respondents considered the test outcome features (risk of SJS, test accuracy, cost of test, cost of long-term gout treatment) very important, while the service delivery process factors (test location, sample collection, results delivery) to be less important. Majority of respondents considered doctor's recommendation to be very important. Most respondents would

consider the choice of peer patients, but would not necessarily follow. No additional salient factors were raised by respondents. In addition, heterogeneity in preferences were observed across different respondents.

Findings from the in-depth interview, combined with the literature information, formed a pool of attributes, the effect of which would be quantified and further investigated using a discrete choice experiment (DCE).

4.6 Using discrete choice experiment (DCE) to study preferences for pharmacogenetic testing

Discrete choice experiment is a stated-preference method to quantify preferences using a series of choice questions.[50-52] When revealed preferences or actual market behaviors are not observable, such as when a market does not exist, or when a product is not yet available, stated preference method can provide useful insights on preferences by offering hypothetical choice sets. Discrete choice experiment is also referred to as choice-based conjoint analysis. The name “Conjoint analysis” arose from the key characteristics of this type of study that different features of products or services are “CONsidered JOINTly”.[53] Each feature is referred to as an attribute. And each choice alternative is composed of combinations of levels of each attribute. Compared to other stated-preference methods, such as contingency valuation, the key advantage of DCE is that it is better at measuring the preferences for each attribute level (the marginal value), the relative importance of various attributes, and the tradeoffs between different attributes.[54] DCE elicit preferences using choice questions, which is a more intuitive and realistic way of everyday decision making, compared to other methods such as rating, or ranking.[53]

First developed in marketing, later adopted by public and environmental economists, conjoint analysis and DCE have been increasingly used in health care in the recent decade. The preference of patients and other stakeholders regarding medical treatments, screening and preventive services, and health service delivery have been

used to inform clinical practices.[53,55-59] Recently, DCE has gained popularity in informing regulatory decisions. US FDA has published a draft guidance on the use of patients' preference information in 2015,[60] with a section on the methodology of DCE and its applications in weighing the benefit and risk of new drugs and devices.

Common attributes included in DCEs are health care outcome-related attributes (such as treatment efficacy, side effects, and survival), health care process-related attributes (such as waiting time, quality of care, mode of service, and type of health care professionals), cost attributes, and others. DCE allows the explicit quantification of tradeoffs individuals make between different attributes. The tradeoff between an attribute and the cost attribute provides estimates on the monetary value of the attribute level, or the willingness-to-pay (WTP). The DCE results have also been used to predict the choice probability or the uptake rate of a certain product or service.

As will be described in chapter 5, a DCE was conducted to understand patients' preferences for pharmacogenetic testing to reduce risk of severe adverse drug reactions prior to starting allopurinol in gout treatment in Singapore. Based on the literature and in-depth interview presented in this chapter, factors important for patients' testing decision making were included as attributes in the DCE. These factors include: the risk of SJS, the accuracy of genetic test, the test cost, the long-term treatment cost, doctor's recommendation and herd behavior. The objective is to examine the relative importance of these attribute, and quantify the tradeoffs patients made between different attributes. The WTP for genetic testing, and the test uptake rate were of interest. In addition, the impact of potential policies or test feature changes on test uptake was simulated to inform clinical practice and policy making. The hypotheses are: 1) Respondents prefer lower risk of SJS, lower cost of genetic test and long-term gout treatment; 2) There is preference heterogeneity across patients, in terms of relative importance of attributes and willingness-to-pay; 3) Information that an alternative is recommended by doctor leads to higher willingness-to-pay and higher uptake rate for this alternative, compared to when it is not the doctor recommended; and 4) Information that an alternative is the

most common choice results in higher willingness-to-pay and higher uptake rate for this alternative, compared to when it is not the most common choice.

Several factors were not studied in DCE. First, self-efficacy factors were not considered, as DCE only elicit stated preferences, but not actual behaviors. Second, test process variables (such as location of test, waiting time for test results, test results disclosure) were not included in the study, as respondents considered these factors to be less important. Third, those factors that are unlikely to change (such as incidence and mortality of SJS) were given as background information, instead of as attributes in DCE.

4.7 Conclusion

This chapter outlined the importance of understanding patients' preferences, and reviewed the theoretical and empirical literature on patients' preferences for genetic testing and its determinants. These leads to the formulation of specific research hypotheses to be tested in the DCE.

Chapter 5 High-risk Asian patients' preferences for pharmacogenetic testing to identify risk of severe adverse drug reaction in chronic gout treatment--A discrete choice experiment

5.1 Abstract

Aims

This study aims to investigate patients' preferences for using genetic testing to reduce the risk of a life-threatening adverse drug reaction named Stevens-Johnson syndrome (SJS). This study also explored the impact of doctor's recommendation and herd behavior on patients' decision making.

Methods

A discrete choice experiment was conducted in which 200 patients were asked to choose between 3 treatment alternatives that differed in six attributes: whether genetic test is involved, risk of developing SJS, cost of the test, cost of long-term gout treatment, doctor's recommendation, and the most common choice. Conditional logit, mixed logit, and latent class models were used to analyze the choice data. Relative importance of attributes, willingness-to-pay for risk reduction, and test uptake rate were estimated.

Results

The latent class model identified two distinct classes of patients. Most patients are risk averse, and had higher preference weights for level of risk reduction than for cost of test. Other patients are more cost conscious, and considered cost of test and long-term treatment more important than the level of risk reduction. Given the current available genetic test, the risk-averse class had higher willingness-to-pay (S$1,215) and predicted test uptake rate (98.3%) at a price of S$400 compared to the cost-conscious class (S$0, and 8.8%). Overall, our results predicted the test uptake rate to be 65.10% in Singapore. The study also revealed the strong impact of doctor's recommendation and moderate effect of herd behavior in shaping individuals' test decisions.

Conclusions

There is a potentially large demand for genetic tests that could reduce the risk of life-threatening ADRs. Physician recommendations and providing information on the choices of others are powerful influences on demand, even more so than moderate price reductions.

5.2 Introduction

Using a discrete choice experiment (DCE), this study aims to investigate patients' preferences for using genetic testing to reduce the risk of a life-threatening adverse drug reaction named Stevens-Johnson syndrome (SJS). Based on the literature review and in-depth interview described in the previous chapter, various test features and decision context information were included as attributes in the DCE. Based on DCE results, the willingness-to-pay for risk reduction and test uptake rate were estimated for various scenarios to inform clinical practice and policies. The specific aims and hypotheses are:

Aim 1: To quantify patients' preferences for various features of pharmacogenetic test.

Hypothesis 1.1: Respondents prefer lower risk of SJS, lower cost of genetic test and long-term gout treatment.

Hypothesis 1.2: The test uptake rate will be higher when a test can reduce the risk of SJS to a lower level, or when the cost of genetic test and long-term gout treatment is lower.

Hypothesis 1.3: Patients are willing to pay additional cost for a test-guided treatment strategy that results in lower risk of SJS.

Hypothesis 1.4: There is preference heterogeneity across patients, in terms of relative importance of attributes and willingness-to-pay. Some patients may consider the risk of SJS as the most important factor and have high willingness-to-pay for risk reduction, whereas others may care more about cost.

Aim 2: To quantify the extent to which information on doctor's recommendation can influence the likelihood of an alternative being chosen.

Hypothesis 2.1: Information that an alternative is recommended by doctor leads to higher willingness-to-pay and higher uptake rate for this alternative, compared to when it is not the doctor recommended.

Hypothesis2.3: Doctor's recommendation is more influential among women, elderly, and those with lower educational attainment.

Aim 3: To quantify the extent to which information on the most common choice can influence the likelihood of an alternative being chosen.

Hypothesis 3.1: Information that an alternative is the most common choice results in higher willingness-to-pay and higher uptake rate for this alternative, compared to when it is not the most common choice.

Hypothesis 3.3: Information on the most common choice is more influential among women, elderly, and those will lower educational attainment.

Hypothesis 3.4: When doctor's recommendation differs from the most common choice, doctor's recommendation is more influential on the final decision.

Aim 4: To forecast the impact of various hypothetical policies on test uptake rate.

Hypothesis 4: Providing information that a test is recommended by doctor is more effective in improving the test uptake rate compared to a strategy that lowers the cost of test or long-term gout treatment.

In addition to addressing the above research questions, this chapter also aims to provide a detailed description of the techniques and processes of conducting a DCE when decision context attributes are involved. Standard DCEs require attributes levels to vary independently in different choice alternatives within the same choice set. However, the presence of choice context requires the different choice alternatives to have correlated attribute levels, which adds to the complexity of study design. Alternative options are discussed to illustrate the process of evaluating and choosing the most appropriate method.

5.3 Methods

Conducting a DCE involves several key tasks: problem refinement and stimuli development, experimental design, survey instrument construction, data collection, and statistical analysis (Figure 7).[50] Stimuli development refers to the determination of attributes, levels, and choice question format. Experimental design is the process of systematically combining attribute levels to make choice alternatives and choice sets. The design process and methods have been reviewed in the literature.[50,51,207,208] Table 8 lists the questions to be addressed in each step. Importantly, study design is an iterative process. In-depth interview, cognitive interview and pre-testing are necessary to obtain respondents' feedback on the design, and suggest improvements on the earlier tasks. For instance, experimental design considerations and respondents feedback may require the modification of attribute levels. An untested design may fail to answer the research questions, and lead to biased preference estimates.

In this study, three iterations were undertaken for survey instrument design. The first iteration involved the identification of a preliminary list of attributes based on the literature review and in-depth interview. In the second iteration, attribute levels and choice question format were selected, and tested in cognitive interviews. In a cognitive interview, each participant was asked to answer specially constructed DCE questions, and "think aloud" to describe their decision making process and rationale to the interviewer.[176] Interviewer also directed questions to better understand the responses. Based on responses, attributes levels were fine-tuned, and choice format was revised so that respondents can understand the questions, and make trade-offs between various attributes and levels as intended. An experimental design and choice sets were generated at the end of the second iteration. The third iteration was a pre-test of the draft survey instrument before fielding to a large number of respondents.

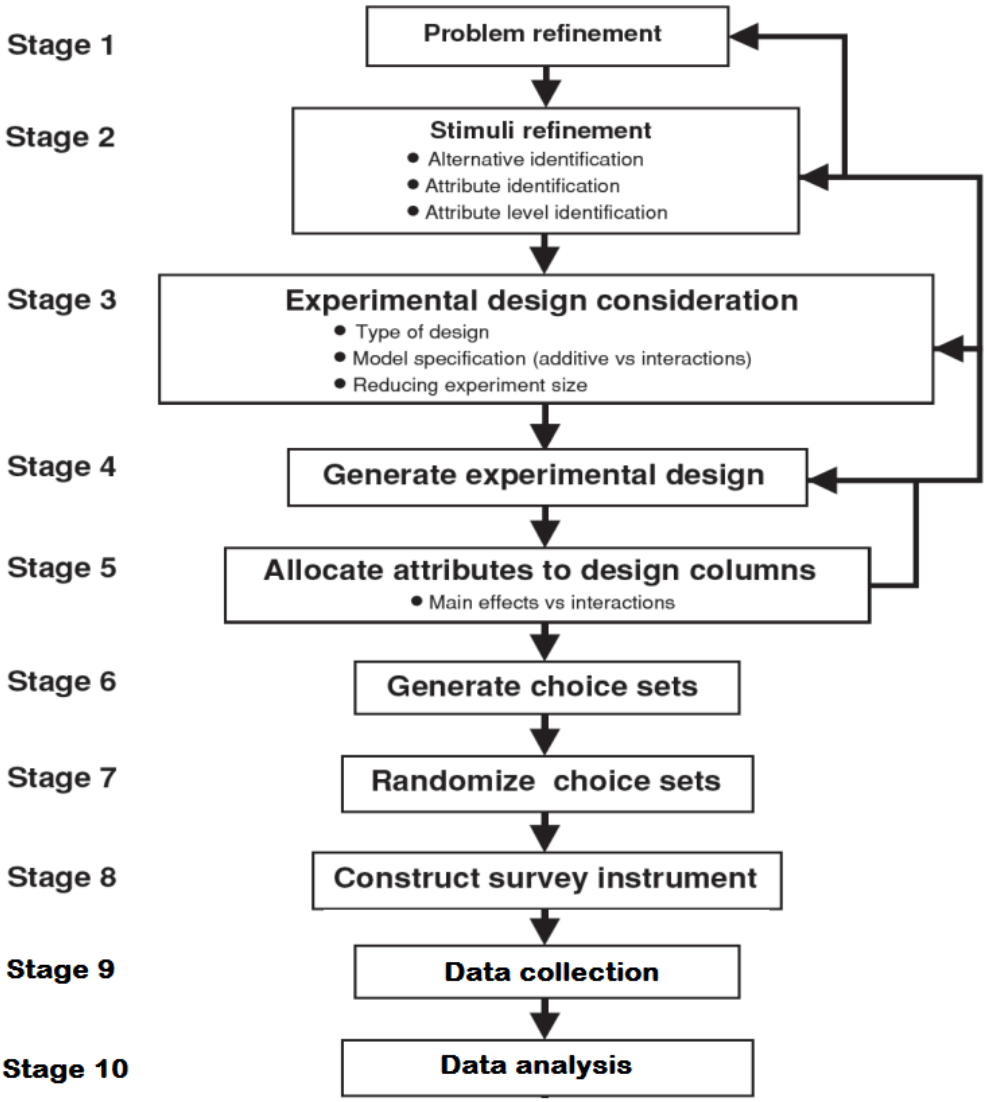


Figure 7. The process and key tasks of undertaking a discrete choice experiment. Modified from "Applied choice analysis: a primer." By Hensher, David A., John M. Rose, and William H. Greene. 2005. Cambridge University Press.[50]

Table 8. Checklist of factors to consider in undertaking and assessing the quality of a discrete choice experiment.

1. Conceptualizing the choice process	Was a choice rather than ranking, rating task used? What type of choice was used: binary response, pairs, multiple options? Was a generic or labelled choice used? Was an opt-out, neither or status quo option included? If a forced choice was used, was a justification provided? Was the task incentive compatible?
2. Attribute selection	How were they derived and validated? Was the number of attributes appropriate? Was the coverage appropriate? What form was used: generic or alternative specific? Was price included? If so, was an appropriate payment vehicle used? Was risk included? If so, was it appropriately communicated?
3. Level selection	How were they derived and validated? Was the number of levels per attribute appropriate? Was an appropriate range used? Were the levels evenly spaced?
4. Experimental design	What type of design was used? Full factorial? Fractional factorial? If fractional, which effects are identified: main effects; main effects + higher order interactions? How were the profiles generated and allocated to choice sets? What are the properties of the design? What is the efficiency of the design? Was identification checked (e.g. is the variance-co-variance matrix block diagonal)? Was the design blocked into versions? If so, how were choice sets allocated to versions? Were the resulting properties of the versions checked? Were respondents randomly allocated to versions? How many choice sets were considered per respondent? If some profiles were implausible – how was implausibility defined and how was it addressed?
5. Questionnaire design	Was an appropriate level of background and contextual information provided? Were the task instructions appropriate? Was the medium used to communicate attribute/level information (e.g. words, pictures, multi-media) appropriate?
6. Piloting	Was coverage of attributes and levels checked? Was understanding and complexity checked? Was the length and timing checked?
7. Population/study perspective	Appropriate for research question?
8. Sample and sample size	Were inclusion/exclusion criteria explicit? Was sample size appropriate for model estimation?
9. Data collection	What recruitment method was used? How were data collected (e.g. mail, personal interview, web survey)? What was the response rate? Were incentives used to enhance response rates?
10. Coding of data	Was coding explicitly discussed? Was the coding appropriate for effects to be estimated?
11. Econometric analysis	Were the estimation methods appropriate given experimental design and type of choice response? Was the functional form of the indirect utility functions appropriate given the experimental design? Were alternative specific constants included? Were sociodemographics and other co-variates included? Was goodness of fit considered?
12. Validity	Was internal or external validity investigated? Were answers for any respondents deleted and if so on what basis?
13. Interpretation	Was the interpretation appropriate given coding of data? Were results in line with *a priori* expectations?

From "Conducting discrete choice experiments to inform healthcare decision making." By Lancsar, Emily, and Jordan Louviere. Pharmacoeconomics 26, no. 8 (2008): 661-677.[51]

5.3.1 Problem refinement and stimuli development

This step involves the development of attributes, levels, and choice question format, the process of which is described in details in this section. The final set of attributes and levels are shown in Table 9. Two sample choice sets are displayed in Figure 8.

Table 9. Final attributes and levels studied in DCE

Attributes	Levels
The chance of getting the severe side effect	1 out of one million patients 1 out of 5,000 patients 1 out of 1,000 patients 1 out of 600 patients
Cost of one-time genetic test	S$20 S$200 S$400 S$1,000
Cost of gout medicines (over two years)	S$250 if test positive (2 in 10 chance), S$200 if test negative (8 in 10 chance) S$400 if test positive (2 in 10 chance), S$200 if test negative (8 in 10 chance) S$1,500 if test positive (2 in 10 chance), S$200 if test negative (8 in 10 chance) S$4,000 if test positive (2 in 10 chance), S$200 if test negative (8 in 10 chance)
Your doctor's recommendation	No information on doctor's recommendation An alternative is the doctor recommended alternative An alternative is not the doctor recommended alternative
Most common choice	No information on the most common choice An alternative is the most common choice An alternative is not the most common choice

Question 2: If you had to choose one of the treatment strategies below, which would you choose?

	Treatment A	Treatment B	Treatment C
Whether genetic testing is involved	Test	Test	No Test
The chance of getting the severe side effect	1 out of 600 patients	1 out of 5,000 patients	1 out of 500 patients
Cost of the one-time genetic test	S$400	S$20	$0
Cost of gout medicines (over two years)	**S$250** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$1,500** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$200** over two years
Your doctor's recommendation		Doctor recommended	
Most common choice	No information	No information	No information

Question: If these were the only 3 options available, which ONE would you choose? (Please tick √) ☐ ☐ ☐

Question 5: If you had to choose one of the treatment strategies below, which would you choose?

	Treatment A	Treatment B	Treatment C
Whether genetic testing is involved	Test	Test	No Test
The chance of getting the severe side effect	1 out of 1,000 patients	1 out of 5,000 patients	1 out of 500 patients
Cost of the one-time genetic test	S$20	S$400	$0
Cost of gout medicines (over two years)	**S$250** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$4,000** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$200** over two years
Your doctor's recommendation	Doctor recommended		
Most common choice	✔		

Question: If these were the only 3 options available, which ONE would you choose? (Please tick √) ☐ ☐ ☐

Figure 8. Sample DCE choice questions

5.3.1.1 Attributes development based on literature and in-depth interview

The determinants of genetic testing decisions identified based on the Health Belief Model and supported by empirical literature and in-depth interview in the previous chapter formed the initial pool of attributes:

Test feature attributes:

- Risk of developing SJS (with test-guided treatment)
- Cost of test
- Cost of long-term gout treatment

Choice context attributes:

- Doctor's recommendation
- Most common choice

5.3.1.2 Attribute level development and testing through cognitive interview

Determine Attribute levels

Once attributes were determined, the next step was to select levels for each attribute. Levels were quantified or unambiguously defined to avoid confusion and minimize variations in interpretation. Four criteria were considered for attributes level selection. Firstly, the observed or most realistic levels were included, in order to make predictions about real life behaviors. Secondly, policy relevant levels, or levels that would become realistic in the future were included, to improve the predictive power of the study to forecast impact of policies and future changes. Thirdly, a broad range of levels were included to explore the switching point, at which respondents may switch choices. When all levels of an attribute are considered very low or very high to a respondent, this attribute may dominate other attributes, or may be neglected, both of which are inefficient in collecting preference information. Including appropriate range of levels ensures that respondents actively evaluate the different attribute levels, and made trade-offs between attributes. Fourthly, the number of levels was set at 4 for risk and cost attributes, and 3 for doctor recommendation and most common choice

attributes. With more levels included, the more information regarding preferences for that attribute can be captured. However design size and cognitive burden will increase with number of levels.[50]

The attribute level extremes identified above were systematically tested in cognitive interviews to 1) fine-tune the attribute levels, especially the extreme ranges, so that respondents feel the presented attribute levels are relevant, and may change their decisions based on the different levels, and 2) to explore respondents' willingness to make trade-off between different attributes. In a cognitive interview, each participant was asked to answer specially constructed DCE questions, and "think aloud" to describe their decision making process and rationale to the interviewer.[176] The interviewer also directed questions to better understand the responses. Each specially constructed DCE choice set includes 2 hypothetical test alternatives, where two attributes were varied at one time, while fixing the other attributes at the middle levels, in order to examine the trade-offs made between any two attributes. The best level of an attribute was combined with the worst level of another attribute in one of the profiles, and vice versa for the other profile. (See Table 10 for the design of DCE choice sets for cognitive interview). A no test alternative was then added to each choice set as a fixed comparator.

Table 10. Cognitive interview DCE choice sets template to identify extreme ranges of attributes.

	Q1			Q2			Q3		
	Test A	**Test B**	**No test**	**Test A**	**Test B**	**No test**	**Test A**	**Test B**	**No test**
Risk of developing SJS	Best level	Worst level		Best level	Worst level		Middle level	Middle level	
Cost of test	Worst level	Best level	Fixed	Middle level	Middle level	Fixed	Best level	Worst level	Fixed
Cost of long-term gout treatment	Middle level	Middle level		Worst level	Best level		Worst level	Best level	
Doctor's recommendation	Middle level	Middle level		Middle level	Middle level		Middle level	Middle level	
Most common choice	Middle level	Middle level		Middle level	Middle level		Middle level	Middle level	

	Q4			Q5			Q6		
	Test A	**Test B**	**No test**	**Test A**	**Test B**	**No test**	**Test A**	**Test B**	**No test**
Risk of developing SJS	Best level	Worst level		Best level	Worst level		Middle level	Middle level	
Cost of test	Middle level	Middle level	Fixed	Middle level	Middle level	Fixed	Best level	Worst level	Fixed
Cost of long-term gout treatment	Middle level	Middle level		Middle level	Middle level		Middle level	Middle level	
Doctor's recommendation	Worst level	Best level		Middle level	Middle level		Worst level	Best level	
Most common choice	Middle level	Middle level		Worst level	Best level		Middle level	Middle level	

	Q7			Q8			Q9		
	Test A	**Test B**	**No test**	**Test A**	**Test B**	**No test**	**Test A**	**Test B**	**No test**
Risk of developing SJS	Middle level	Middle level		Middle level	Middle level		Middle level	Middle level	
Cost of test	Best level	Worst level	Fixed	Middle level	Middle level	Fixed	Middle level	Middle level	Fixed
Cost of long-term gout treatment	Middle level	Middle level		Best level	Worst level		Best level	Worst level	
Doctor's recommendation	Middle level	Middle level		Worst level	Best level		Middle level	Middle level	
Most common choice	Worst level	Best level		Middle level	Middle level		Worst level	Best level	

	Q10		
	Test A	**Test B**	**No test**
Risk of developing SJS	Middle level	Middle level	
Cost of test	Middle level	Middle level	Fixed
Cost of long-term gout treatment	Middle level	Middle level	
Doctor's recommendation	Best level	Worst level	
Most common choice	Worst level	Best level	

*In each choice set, only two attributes were varied, while all other attributes were fixed at the middle level. The two attributes varied were underlined.

A total of 50 diabetes patients were recruited from Singapore General Hospital (SGH) Diabetes Centre and the National University Hospital (NUH) Diabetes Clinic for cognitive interviews. Several rounds of cognitive interviews were conducted, with the attribute level extremes adjusted based on respondents' choices in the previous round of cognitive interview. For instance, when no respondent chose the no test alternative, it suggested that the highest risk of developing SJS associated with test alternatives should be increased to encourage trade-offs. If respondents always chose the lower risk alternative, regardless of cost, the highest cost level should be increased in order to identify the maximum willingness-to-pay. In the last round of cognitive interviews, tradeoffs were observed. Among respondents, very few made their responses always consistent with the better available level of one attribute (dominating on an attribute), indicating that with the current attribute level extremes, all attributes are important so that they make trade-offs between different attributes instead of only considering one attribute.

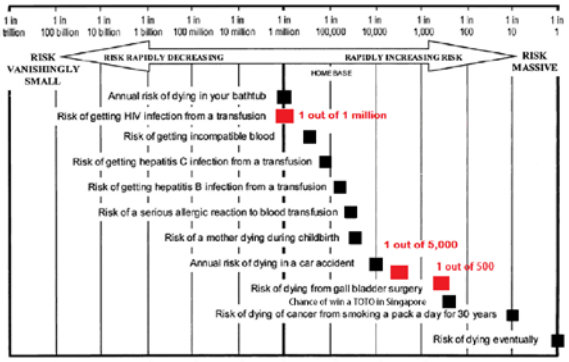
Determine the attribute level display format

In the cognitive interviews, the best framing and presentation format of attribute levels were also explored.

The risk of developing SJS is 0.2% without testing, and further reduced to almost 0 with testing-guided treatment. Such small probabilities that people do not often encounter in daily life are difficult to make sense of by respondents. In the literature, it is found that people are not good at understanding probability expressions, especially small probabilities.[169,173,209-211] Some common graphic displays tools such as grid, and dots do not work well for very small probabilities. We tested three possible formats of presenting the risk of SJS including the use of percentage, the use of frequency, and a graphic display with a Pailing scale (Table 11).[209,210] The frequency format (1 out of xxx patients) was found easy to understand and quantify, and was used in the final survey. Respondents reported that percentage expressions were not easy to imagine, and some respondents considered all levels to be very low when presented in

percentages. The Pailing perspective scale is a method to display the probability of an event relative to the probability of other events which people are more familiar with (such as the risk of dying off cancer, the chance of winning TOTO lottery, the chance of getting HIV infection from transfusion). In cognitive interviews, this was found to be time-consuming, and incurring significant cognitive burden to respondents.

Table 11. The risk display format tested in cognitive interviews

Display format	Percentage format	Frequency format (preferred)	Pailing perspective scale
Risk of getting SJS	0.2% of patients	1 out of 500 patients	1 in 500 

The cost of long-term gout treatment is an attribute with an uncertainty component. As genetic test results can aid the selection of drugs, the long-term gout treatment cost depends on the test results. We therefore displayed the gout treatment costs associated with positive and negative test results, as well the chance of testing positive. The gout treatment cost associated with negative test results (S$200 over 2 years) and the chance of test positive (20%) are fixed across different levels, and only the cost associated with positive test results was varied (S$250, S$400, S$1,500,and S$4,000). In the pre-testing, respondents could understand that gout treatment cost was uncertain at the time of making testing decisions, and the actual cost would depend on test results. We provided the cost of gout treatment in 2 years, as gout is a chronic condition with one episode of treatment lasting for over 2 years.

The display of doctor's recommendation was explored in the cognitive interviews. In real life scenarios, doctors often recommend one treatment from the available alternatives. In other cases, doctors may provide information on the treatment alternatives but make no clear recommendation, and encourage patients to make a decision based on his/her own preference. We therefore imposed the restriction that only one alternative can be recommended by doctor in each choice set, or no information on doctor's recommendation is provided. The doctor's recommendation attribute was framed as the information on doctor's recommendation. It differs from a real recommendation delivered by a doctor personally during a face-to-face consultation. The physical presence of doctor and the interactive nature of the recommendation will make an actual recommendation more salient and effective than providing information on doctor's recommendation in a survey questionnaire. To improve the saliency of the doctor's recommendation attribute, a flag shape label was used to indicate doctor's recommendation (Figure 8). A graphic display not only attracts respondents' attention, but also makes it easier to understand the recommendation.

To test for the presence of herd behavior, the choice of the herd can be described in quantitative or qualitative ways. Showing the percentage of respondents choosing each alternative gives precise information, however multiple levels may be required in the design to identify the percentage at which respondents will follow the herd. In addition, as several alternatives were offered in each choice set, respondents may undertake calculations with percentages, and confusion may arise if all percentages do not sum up to 100%. To simplify this attribute and avoid confusion, the levels were described qualitatively. An alternative can be "the most common choice" or "not the most common choice". In some choice sets, this attribute has the level "no information". A visual display was used, with the most common choice indicated by a tick mark (Figure 8).

Given the above considerations and findings, the final set of attributes and levels are shown in Table 9. The initial set of attributes and levels formulated without testing or revision are shown in Table 12. Comparing the two sets, significant changes in attribute framing and attribute levels were made to improve the survey.

Table 12. Initial set of attributes and levels before cognitive interview

Attributes	Levels
Your cost for the test	Free S$50 S$200 S$500
Your chance of developing SJS	0 (no chance) 1 SJS case in 50,000 users of allopurinol 1 SJS case in 5,000 users of allopurinol
Gout treatment costs	2 every 10 people will test positive. If you are one of these people, you will have to take a drug that will cost you SGD500 a year. If you do not test positive, you can safely use allopurinol 2 every 10 people will test positive. If you are one of these people, you will have to take a drug that will cost you SGD1,000 a year. If you do not test positive, you can safely use allopurinol 2 every 10 people will test positive. If you are one of these people, you will have to take a drug that will cost you SGD2,000 a year. If you do not test positive, you can safely use allopurinol
Your doctor's recommendation	You receive a doctor's recommendation on the genetic test You receive no recommendation on the genetic test
Herd behaviour	10% of people in your situation take the genetic test 90% of people in your situation take the genetic test

5.3.1.1 Determine DCE question format and test via cognitive interview

Besides fine-tuning the attributes and levels, there are several other objectives of cognitive interviews: 1) to explore the ability of respondents to understand the attributes and DCE questions, and determine the appropriate format of DCE question, 2) to understand the cognitive burden and difficulty level of the survey in the study population, and 3) to test and improve the wording of survey instrument.

Inclusion of an opt-out option in the choice set

DCE question aims to elicit a response on the preferred alternative within each choice set. However, it is possible that none of the test alternatives is preferred, even though one test is perceived better than the others. It is important to capture this type of non-demander preference when trying to make predictions about real life behaviors.[212] This is particularly relevant for this study as we sort to understand whether or not individual patients are willing to take a genetic test, in addition to estimating the preferences for test features. Without an opt-out option, the test uptake rate may be overestimated.[213] Three types of modifications can accommodate the non-demander behaviors: 1) having a no test alternative with all features displayed, 2) including "none" as an option in the response, or 3) adding a follow-up question after the preference question to verify whether the preferred option will be implemented when offered. After testing in cognitive interviews, the no test alternative was chosen. Display of all attributes levels for the no test alternative allows respondents to compare the cost and consequences of testing and no testing, and minimizes the discrepancies in individual beliefs about no test. To make the questions realistic, we constructed the no test alternative using the realistic attribute levels, and kept the risk and cost attributes of this alternative fixed in all DCE choice sets.

Number of alternatives in a choice set

Having more profiles in a question will increase the amount of information obtained from each question, however may increase the complexity of questions and the cognitive burden to respondents. Most DCE studies in health care include 2 or 3 alternatives in each choice set. In the cognitive interviews, both numbers were tested, and respondents had no difficulty in handling three alternatives. Furthermore, having 3 alternatives (2 test alternatives + 1 no test alternative) has advantages over 2 alternatives (1 test alternative+ 1 no test alternative) in reducing the labeling effect of testing. With only one test alternative and one no test alternative, respondents may

take the mental shortcut to always choose test or no test based on their prior belief about genetic testing instead of looking at the attribute levels.

Labeled vs. unlabeled alternatives

In addition to the attributes and levels, the label of alternatives also significantly influences the responses. In an unlabeled design, the alternatives are given generic names such “Alternative A”, “Alternative B” or “Treatment A” and “Treatment B”. In a labeled design, the name of alternative confers some information about the alternative, such as “Genetic test A”, “Genetic test B” and “No test”. Assigning informative labels to the profile will make it more realistic to respondents, which is likely to improve the power of DCE to predict real behaviors. However, the label has been shown to influence individual choices and reduce the attention respondents give to the attributes.[214] Both labeled and unlabeled designs were tested in the cognitive interviews.

When the three profiles were labeled as “Genetic test A”, “Genetic test B”, and “No test”, respondents were less likely to indicate the no test alternative as the most preferred, compared to the unlabeled design, where the alternatives were labeled “Alternative A”, “Alternative B” and “Alternative C”. This may be reasonable as the genetic test label confers information, and patients may have intrinsic preferences for taking a genetic test to reduce risk of life-threatening ADRs, regardless of the attribute levels. However, a small number of respondents mistakenly understood the no test alternative as having no gout treatment. In order to minimize the potential misunderstanding, the labels were revised to be “Treatment A” “Treatment B” and “Treatment C” to reassure respondents that gout treatment will be given in all three alternatives, with the difference being the involvement or absence of genetic testing prior to treatment. An additional attribute was introduced to indicate whether a treatment involves genetic testing to capture respondents’ intrinsic preferences for the label of genetic test.

Type of preference-eliciting questions

The most common type of preference-eliciting question requires respondents to indicate the most preferred alternative. Newer DCE studies has explored different types of questions, such as the best-worst type, which require respondents to report both the most preferred (best) and least preferred (worst) alternatives.[215] In a DCE study with 3 alternatives in each choice set, the best-worst type questions provide the complete preference ranking of an individual over different alternatives. The coefficient estimates using both the best and worst response have smaller standard errors than models estimated using responses on the best alternative only, demonstrating gains in statistical efficiency from the additional preference information gathered.[215,216] However, the cognitive processes and certainty of responses to the best and worst question are different, and there are controversies on the appropriate weights assigned to the best and worst questions.[215,217] In addition, there are concerns on the cognitive burden of asking 2 follow-up questions in each choice set. In pre-testing, best and worst types of questions were tested, and some confusion was observed, especially among those with lower education level. The switch between best and worst questions appeared to require a switch in the decision making pathway, and increased the cognitive burden. There is also trade-off between the number of follow-up questions in each choice set and the number of choice sets respondents can go through in a given amount of time. Therefore, respondents were only asked to choose the most preferred alternative in the final survey.

Number of DCE choice sets in the survey

In the DCE literature, a wide range of choice set numbers have been used. The optimal number of questions depends on the complexity of DCE questions, and the cognitive power of respondents. A study that compares a design of 5, 9, and 17 choice sets found that respondents exposed to 17 choice sets had higher response variance, suggesting a large number choice sets may increase cognitive burden.[218] Cognitive burden may leads to inattentive or inconsistent responses. Cognitive interviews reveal

that respondents can answer 10 questions with reasonably good attention and certainty. Even though a small number of respondents started to fatigue after 4-6 questions mainly because all DCE questions looked similar, they could re-gain focus with the encouragement of interviewers. 10 trade-off questions were included in the final survey.

5.3.2 Experimental design

Experimental design is the process of systematically generating a sample of choice sets which constitutes choice alternatives that are specific combinations of attributes and levels.[208] Experimental design should be tailored based on the research objective, specifications of attributes and levels, choice question format, as well as the analysis requirements.[208] According to the ISPOR Task Force on Conjoint Analysis, the good practice of experimental design requires researchers to evaluate alternative design approaches and justify the approach chosen.[207] In order to select and evaluate various design approaches, four aspects were considered. Johnson *et al.* highlighted two general objectives in experimental design: *model identification* and *efficiency*.[208] Louviere *et al.* discussed two additional design objectives: *reduce cognitive complexity* and *market realism*.[52] Model identification means independent and unbiased estimation of the desired form of effect parameters from the survey data, and is the most important design consideration. *Efficiency* refers to the statistical power of the design to estimate the effect parameters precisely with relatively small sample size. *Reducing cognitive complexity* requires researchers not to incur excessive cognitive burden on respondents, as cognitive burden may threaten the consistency and validity of responses. *Market realism* influences the power of the study to explain or predict real life behaviors. A perfect design may not exist. Often, the importance of the four objectives needs to be weighed and compromised to achieve a good balance.

5.3.2.1 Experimental design theories and approaches—a literature review

To generate an experimental design, several approaches are commonly used in the DCE literature.[208] Different approaches have different underlying algorithms and properties. In general, there are two classes of designs: full factorial design and fractional factorial design.[50,51,219] Full factorial design generates all possible combinations of levels from each attribute, and the main effects and interaction effects of all attribute levels can be estimated independently. However, full factorial design requires large number of questions if the study involves many attributes and levels, which are usually impractical. In this study, there are three attributes with four levels each, and two attributes with three levels each (denoted as $4^3 3^2$). A full factorial design would generate 4x4x4x3x3=576 different combinations. In contrast, fractional factorial designs select only a small fraction of possible combinations while ensuring the effects of interest can be estimated. Different approaches are used to select a fraction of combinations. Designs can be obtained from catalogues, software, or generated by hand.[51,208] The generated designs may differ in three key properties: *orthogonality, statistical efficiency* and *response efficiency*.

Orthogonality is a constraint that all attributes be statistically independent of each other (though conceptually attributes may be related), and zero correlations between attributes.[50] Orthogonality relates to the design objective of unbiased identification of parameters in statistical analysis. For example, in a study to understand patients' preferences for treatment effectiveness and adverse drug reactions, if the treatment that is more effective always results in lower rate of adverse drug reactions, researchers will not be able to distinguish the independent effect of effectiveness and adverse drug reactions on patients' preferences. *Balance* is a related property that requires each level of an attribute to appear equal number of times, and is a necessary condition for strict orthogonality. The designs that emphasize on orthogonality are referred to as orthogonal or near-orthogonal fractional factorial designs. These include orthogonal arrays (which can be obtained from manual catalogue), orthogonal main-effects plan (OMEP), and OMEP-based designs such as

the fold-over design, and designs generated by Sawtooth software.[208,220-222] These designs require zero or near zero correlation between attributes, and therefore guarantees the identification of main effects and sometimes interaction effects. However, small size orthogonal arrays may not be available for some number of attributes and levels. Moreover, orthogonal designs cannot incorporate constraints on dominance or implausible combinations. For instance, some random combinations may be implausible or dominated (where all levels of one alternative are unambiguously better than the levels of another alternative). This is likely to occur when the attribute levels are naturally ordered.[208] This type of combinations does not reveal information on preferences as the better choice is obvious regardless of preference if people are rational. Imposing restrictions to avoid implausible combinations may improve statistical efficiency and reduce potential confusion among respondents.

Statistical efficiency refers to the minimization of confidence intervals around parameter estimates in a choice model for a given sample size.[208] A statistically less efficient design may be compensated by a large sample size to obtain rather small confidence intervals.[208] However, when the intended sample size is small, statistical efficiency is crucial. Optimal fractional factorial designs emphasize on statistical efficiency at the expense of orthogonality. D-efficiency and D-optimality are commonly used efficiency criteria to measure, generate and compare the efficiency of designs.[186] Design approaches that focus on efficiency include the SAS macros using D-efficiency, Street and Burgess' cyclic design, Sandor and Wedel's Bayesian design, and Bliemer's design.[216,219,223,224]

Besides statistical efficiency, there is another type of efficiency referred to as *response efficiency*, which is about the measurement errors resulting from poor quality response. This property relates to the objective of minimizing cognitive burden. When DCE questions are complex or ambiguous, or when a large number of DCE questions are included, respondents may fatigue and pay less attention to the questions, or even

take mental shortcuts that deviate from utility maximization. The inconsistency in responses may result in bigger variance in estimates. Louviere *et al.* demonstrated that an increase in statistical efficiency was always associated with a decrease in response consistency.[225]

A good design requires a balance of orthogonality, statistical efficiency and response efficiency. In practice, there are trade-offs that researchers have to make between these three considerations. Statistical efficiency can be increased by asking a large number of difficult trade-off questions with no implausible combinations, and no level overlaps. However these will violate strict orthogonality by imposing correlations, and incur significant cognitive burden which will threaten response efficiency. Empirically some design properties such as *D-efficiency, correlation, balance, and overlap* can be indicators of orthogonality and efficiency, and should be checked after design is generated.

5.3.2.2 Generate experimental design using D-efficiency criteria in SAS

In this study, the D-efficiency measure was used to generate a fractional factorial design. D-efficiency minimizes the joint confidence sphere around the complete set of estimated model parameters, that is, maximizes the statistical efficiency.[208] The advantage of D-efficiency approach is the flexibility to incorporate restrictions while maximizing statistical efficiency. D-efficiency design was generated in SAS software based on the algorithm described by Kuhfeld.[219] After the design was generated, properties including orthogonality, balance and overlap were checked to ensure sufficient identification of parameters. Pre-testing was conducted to ensure response efficiency. A flow chart illustrates the process of experimental design in SAS (Figure 9).

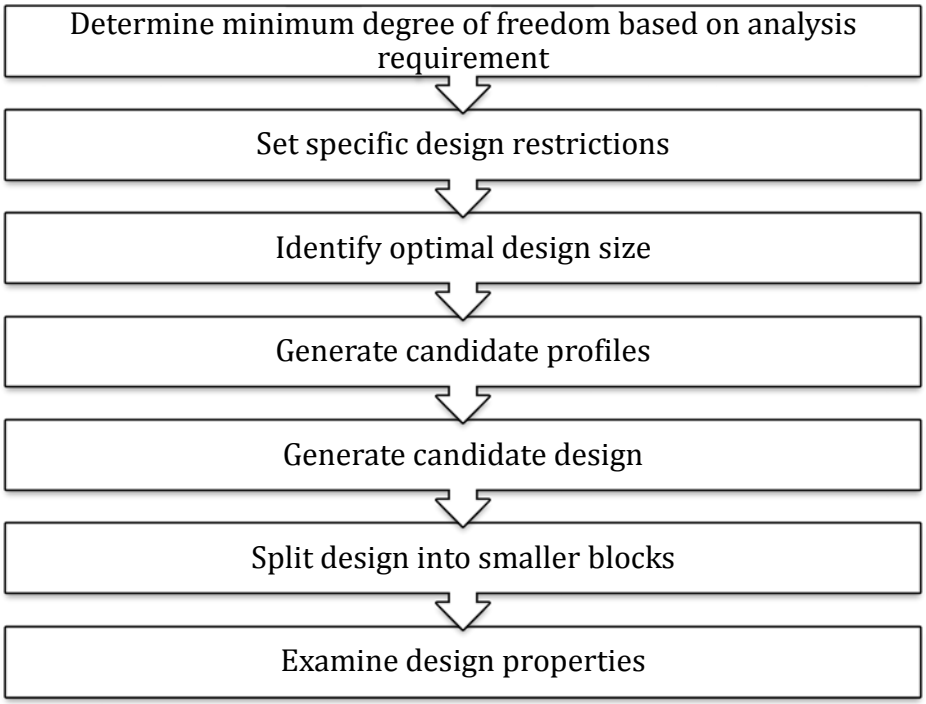


Figure 9. Process of experimental design in SAS

Determine minimal degree of freedom required

The first step of design is to determine the minimum design size or the degree of freedom required based on the analysis plan. We want to estimate the main effect of each attribute. Cost of test will be estimated as a linear variable, and the linearity of utility in this attribute was tested. All other variables will be treated as categorical and the effect of each level will be estimated separately. The presence of no test option requires one more degree of freedom to estimate the alternative-specific constant for no test. In addition, two more degrees of freedom are required to differentiate the effect of doctor recommendation and herd information on test alternative and no test alternative. In total, 14 parameters need to be estimated, which implies that the most parsimonious design needs to contain 14 DCE choice sets to ensure model identification (Table 13).

Table 13. Minimum degree of freedom required for analysis

Attribute	Number of levels	Parameters to be estimated/degree of freedom required
Cost of test	4	1
Risk of SJS	4	3
Cost of gout treatment	4	3
Alternative specific constant for no test	1	1
Doctor's recommendation (main effect+ interaction with no test)	3	2+1=3
Most common choice (main effect+ interaction with no test)	3	2+1=3
Total		14

Special design considerations for doctor's recommendation and the most common choice

To make the choice questions realistic, correlations between alternatives were imposed on doctor's recommendation and most common choice attributes, so that within any choice set, at most one alternative can be labeled "doctor recommended" or "most common choice". That means, when one of the three alternatives is the doctor recommended, the other two were not recommended by definition. As a result, doctor's recommendation and most common choice attributes do not vary freely across alternatives within the same choice set, and are specific to each choice set to form a choice context. Standard DCE designs do not easily accommodate the within-choice set correlations in attribute levels.

Three possible design solutions were considered. The first solution is to include only one test alternative in a choice set, and treat choice context attribute as a normal attribute. However, it will reduce the amount of information obtained from each question, and may require a larger number of questions. The second solution is to have multiple versions of questionnaires that contain the same set of DCE questions that only differ in the choice context.[226] This again requires larger sample sizes. We

adopted a third solution, in which we designed the level of doctor's recommendation and most common choice attributes for the two test alternatives and allow the levels to vary independently across alternatives. The level for the no test alternative was inferred and displayed based on the correlation. See Figure 10 for an illustration of the design output and level modifications. In brief, when there are 2 levels of doctor's recommendation (doctor recommended, not the recommended), there are 4 possible scenarios in a 2-alternative choice set: 1) A is recommended, 2) B is recommended, 3) A and B are both recommended, and 4) neither A nor B is recommended. For scenario 1, 2 and 4, the level for the no test alternative can be easily imputed based on the restriction that at only one alternative can be recommended. For scenario 3, it violates our restriction, and the levels can be replaced by "no information" in all alternatives. SAS algorithm minimizes level overlap, which is the chance of the two alternatives sharing the same level (scenario 3, and 4) is low. To ensure the four scenarios occur in equal frequencies, the two levels were duplicated, and 4 levels (doctor recommended, not the recommended, doctor recommended, not the recommended) were used in design. The most common choice attribute was designed in the same way. Notably, modifying the levels after the generation of design may alter the design properties. So some important properties (such as efficiency, orthogonality, balance, and overlap) of the final design were evaluated subsequently.

Design

	Profile A	Profile B
1	Recommended	Not the recommended
2	Not the recommended	Recommended
3	Recommended	Recommended
4	Not the recommended	Not the recommended

Manipulate the levels

Profile A	Profile B	Profile C (no test)
Recommended	Not the recommended	Not the recommended
Not the recommended	Recommended	Not the recommended
No info	No info	No info
Not the recommended	Not the recommended	Recommended

Figure 10. Possible design scenarios and level modifications

5.3.2.3 Generate experimental design in SAS

The above design considerations require a design for 5 attributes with 4 levels each. Interaction term was specified between the doctor's recommendation and most common choice attributes. Among the design sizes suggested by SAS, 32 was chosen as the final design size. All attributes have 4 levels, and the interaction term has 16 levels. 32 is dividable by both 4 and 16, and therefore are likely to results in good level balance, which is necessary for orthogonality and efficiency. The design was generated using the %***mktex,*** %***choiceff,*** and %***mktblock*** autocall SAS macros. In brief, 20,000 alternatives were constructed using the attribute levels. 32 choice sets with 2 alternatives in each choice set were then generated using the 20,000 alternatives based on D-efficiency criteria, with restrictions to exclude dominant-pair choice sets where one alternative unambiguously dominate the other. The 32 choice sets were partitioned into 4 blocks of 8 questions, so that each respondent does not need to answer all questions. The levels for doctor's recommendation and most common choice attributes were manipulated as described previously to form the final design. Final design was included in Appendix B.

5.3.2.4 Examine design properties

The final design was examined in terms of correlations (orthogonality), level balance, cross-level balance, and overlap, all of which are important for parameter identification and design efficiency. In brief, no serious correlations between different attributes and attribute levels were detected. Attribute levels were roughly balanced, that is, different levels of the same attribute appeared roughly equal number of times. The frequency of level combinations between any 2 levels of different attributes were roughly balanced (cross-level balance). The frequency of overlap where two alternatives within the same choice set share the same level for a certain attribute was low.

5.3.2.5 Survey validity test

To assess the reliability of responses, we incorporated two internal validity questions to examine respondents' attention and understanding.(see Figure 11) Based on cognitive interview feedback, some respondents had difficulty quantifying the small probabilities in the risk attribute. Therefore the first validity test was placed before DCE questions, and required respondents to identify the scenario indicating higher risk. Respondents who failed to identify the higher risk scenario was given additional explanations on probability expressions before moving on to DCE choice questions. The second test was a "dominant-pair" test in DCE format, in which the two test alternatives share the same level for all attributes, except for the risk attribute where one alternative results in lower risk than the other. In this test, utility maximizers should always prefer the lower risk alternative, regardless of preferences. Respondents who prefer the high risk alternative are likely to be inattentive or misunderstand the risk attribute.

Question B2	**Comparing the two scenarios below, which indicates higher risk?** ☐ 1 out of 500 patients get the severe side effect ☐ 1 out of 1,000 patients get the severe side effect

<u>Question 1</u>: If you had to choose one of the treatment strategies below, which would you choose?

	Treatment A	**Treatment B**	**Treatment C**
Whether genetic testing is involved	Test	Test	No Test
The chance of getting the severe side effect	1 out of one million patients	1 out of 600 patients	1 out of 500 patients
Cost of the one-time genetic test	S$200	S$200	$0
Cost of gout medicines (over two years)	**S$400** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$400** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$200** over two years
Your doctor's recommendation	No information	No information	No information
Most common choice	No information	No information	No information
Question: If these were the only 3 options available, <u>which ONE would you choose?</u> (Please tick √)	☐	☐	☐

Figure 11. Validity test questions

5.3.3 Survey design

The survey questionnaire can be found in Appendix C. Questionnaire starts with two screener questions to verify the eligibility of respondents. An introduction section then briefly educates respondents on gout, side effect of gout treatment and genetic testing, so that respondents have the essential knowledge to answer choice questions, even if they do not have prior knowledge on the topic. The third section defines the attributes and possible levels that respondents will see in the DCE questions. Before introducing the cost attributes, a budget reminder is included to remind respondents to think about the impact of a certain amount of money on their daily life. There is a literature on the hypothetical bias of the DCE approach, which is mainly due to the fact that respondents only state their preference, without actually paying to receive the preferred service or product. The willingness-to-pay estimated from DCE may be higher than that in real life. A budget reminder in cheap-talk format may encourage respondents to think about cost carefully, which will improve the predictive power of DCE.[176] After introducing each attribute, one to two warm-up questions are included to understand respondents' perceptions and attitude towards the attribute levels and to encourage active thinking. Another purpose of the warm-up question is to provide a check point and attract respondents' attention to each attribute, as it was observed in cognitive interviews that respondents had the tendency to focus on questions and skip trunks of reading. The DCE section then starts with short instructions and precautions on common mistakes that respondents should avoid. An example DCE question is provided with explanations on how respondents should interpret the question. Nine DCE questions follows, with the first one being validity test question, and 8 questions from the experimental design. The survey questionnaire ends with questions on respondents' background, including demographics, medical history related to gout, and socio-economic status.

The survey was pre-tested (n=10) to ensure that respondents have no difficulty answering questions, and the questionnaire doesn't incur too much cognitive burden. Based on pre-testing the survey was revised and simplified. For instance, some technical jargons such as "HLA-B*5801 testing", "adverse drug reactions", which respondents have difficulty understanding are replaced with simple language such as "genetic testing", and "side effects of medicines".

A proportion of Singaporean do not speak or read English, the majority of which can speak and read Mandarin. To also gauge the preference of this group, the survey questionnaire and informed consent were translated into Mandarin, and accuracy was verified via a back translation by a different researcher.

5.3.4 Sample size calculation

Orme's rule of thumb was used to determine the minimum acceptable sample size for DCE[53]

$$\frac{nta}{c} \geq 500$$

where n is the minimum sample size, t is the number of DCE tasks, a is the number of choice alternatives per task, and c is the maximum number of attribute levels. In our study, t=8, a=3, c=4. The minimum sample size required is therefore 84. The actual sample size was set at 200, which allows accurate estimation of all attribute levels, and additional analyses.

5.3.5 Sampling and survey fielding

Gout and diabetes are recognized features of metabolic syndrome.[227,228] Gout is a risk factor for diabetes,[229] and diabetes patients have higher risk for gout.[230] Diabetes patients therefore have higher chance of requiring chronic gout treatment with allopurinol, and facing the genetic testing decision in the future. We surveyed a convenient sample of 200 diabetes patients from the Singapore General Hospital (SGH) Diabetes Centre and the National University Hospital (NUH) Diabetes Clinic,

which are the specialist diabetes clinics in two of the largest government's restructured general hospitals in Singapore. The inclusion criteria include being a Singapore citizen or permanent resident, having a diagnosis of diabetes, between the ages of 21 and 80 years. Those who have a diagnosis of gout and have been treated with urate-lowering therapy, and those with limited mental capacity were excluded from study.

We sampled diabetes patients instead of gout patients for two reasons. Firstly, the genetic testing decision is only relevant for those patients who require chronic gout treatment with allopurinol, but have not initiated allopurinol. Based on current knowledge patients who have taken allopurinol but did not develop SJS within the first two months are unlikely to develop SJS in the future, and do not require genetic testing.[120] Therefore, a significant proportion of chronic gout patients are not eligible for our study. Secondly, most gout patients are managed in the primary care setting by general practitioners (GPs) and family physicians in the government's polyclinics and private clinics. GPs usually initiate allopurinol for gout patients and manage the symptoms, and only refer complex cases such as non-response and severe adverse reactions to rheumatologists in the specialist clinics in hospitals. There are eighteen government's polyclinics, and over 2,000 GPs in private clinics. Gout patients therefore seek care in diverse locations. The number of gout patients treated by each doctor and clinic is small, making it operationally challenging to sample.

To recruit respondents, trained interviewers approached patients in the waiting room of the diabetes clinic, verified their eligibility using the inclusion and exclusion criteria, and asked for their willingness to proceed with the survey after reading the information sheet. Informed consent was obtained from each respondent. The study received ethical approval from the National Healthcare Group's Domain Specific Review Board (DSRB), and Singhealth Centralised Institutional Review Board (CIRB).

Each respondent was asked to complete a paper version of the survey instrument, with the help an interviewer to explain the information on the survey instrument and clarify doubts. There were four equivalent versions of survey

instruments each contain one of the four blocks of DCE questions. To minimize version effect and ensure balanced number of each version, we used a block randomization method to randomly assign a survey version to each respondent. Respondents were allocated to blocks of 8, and within each block, two copies of each version of questionnaires were answered. Five interviewers conducted the survey interviews. Interviewers were trained to facilitate the interview and clarify doubts based on a standardized script. Interviewers were also instructed to be neutral and not to express their own opinions on the topic. The presence of potential interviewer effect was tested in data analysis. In addition to respondents' responses to questions, comments from respondents were also documented by interviewers.

5.3.6 Data analysis

5.3.6.1 Analysis of dominance preferences

In discrete choice experiments, respondents are encouraged to make trade-offs between attributes. However, respondents may be unwilling to trade (non-compensatory decision making) and have strong preferences that deviate from this assumption. Lancaster defined a scenario "dominance" as "A characteristic is dominant within some group of characteristics, in some set of situations, if the consumer always prefers a collection with more of the dominant characteristic, whatever the amounts of the other characteristics."[231] Empirically, if a respondent chose the alternative with the best available level of an attribute in all choice sets, the respondent was considered to have a dominant preference for that attribute.[232] Dominant preferences for each attribute including the test label were analyzed. The proportion of respondents with dominance preferences was calculated. Socio-demographic predictors of dominant preference for each specific attribute were also analyzed using logistic regressions.

DCE analysis models

Random utility model framework

The theoretical framework of analyzing choices is McFadden's random utility model (RUM).[233,234] Each respondent faces a choice among j alternatives, repeated under s scenarios or choice situations. The utility that individual n derives from alternative j in scenario s can be decomposed into a systematic component V_{nsj} and a stochastic component ε_{nsj} :

$$U_{nsj} = V_{nsj} + \varepsilon_{nsj} \qquad (1)$$

The analyst do not observe U_{nsj}, but may observe some characteristics of the alternatives X_{nsj} and characteristics of decision maker Z_n, which determine the representative utility V_{nsj}:

$$V_{nsj}=V(X_{nsj}, Z_n) \qquad (2)$$

Analyst assumes decision makers to be utility maximizers, and only chooses alternative i when $U_{nsi} > U_{nsj}\ \forall j \neq i$. After assigning a joint density of the random component ε_{nsj}, the choice probability that decision maker n chooses alternative i in scenario s can be expressed as:

$$\begin{aligned} P_{nsi} &= Prob\left(U_{nsi} > U_{nsj}\ \forall j \neq i\right) \\ &= Prob\left(V_{nsi} + \varepsilon_{nsi} > V_{nsj} + \varepsilon_{nsj}\ \forall j \neq i\right) \\ &= Prob\left(\varepsilon_{nsj} - \varepsilon_{nsi} < V_{nsi} - V_{nsj}\ \forall j \neq i\right) \qquad (3) \end{aligned}$$

Depending on the specification of the density of ε_{nsj}, equation (3) may or may not have a closed form. A logistic specification will give closed form solutions. Several commonly used logistic model include conditional logit model (CLM), mixed logit (MXL) model, and latent class logit model (LCM).[235,236]

Conditional logit model (CLM)[1]

V_{nsj} can be specified as the following:

$$V_{nsj}=X'_{nsj}\beta + Z'_n\gamma \qquad (4)$$

X_{nsj} is a matrix of alternative characteristics, $Z_n$is a vector of decision maker characteristics, and β and γ are vectors of coefficients.

Assuming $\varepsilon_{nsj}$to be identically and independently distributed (IID) as extreme value, this results in the conditional logit specification.[237] The choice probability is the integral of $P_{nsj}|\varepsilon_{nsj}$ over all values of ε_{nsj}. The solution to this integral is the probability of individual i choosing alternative j in scenario s:

$$P_{nsj} = \frac{\exp(X'_{nsj}\beta+Z'_n\gamma)}{\sum_{j=1}^{J}\exp(X'_{nsh}\beta+Z'_n\gamma)} \qquad (5)$$

This model is easy to estimate using maximum likelihood method.[237] However, one limitation of this model is that it cannot account for preference heterogeneity among different individuals.

Mixed logit model (MXL)[2]

Heterogeneity among individuals is usually expected due to differences in tastes and decision making processes. Therefore different individuals may value and weight attributes and levels differently. Mixed logit model is a more general specification that allows the coefficients to differ across individuals.

$$V_{nsj}=X'_{nsj}\beta_n + Z'_n\gamma \qquad (6)$$

β_n is a Kx1 vector of coefficients for attribute levels for individual n, and $\beta_n \sim N_K(\bar{\beta}, V)$ ie. $\beta_{n1},\beta_{n2,\ldots},\beta_{nK,}$ follow a multivariate normal distribution. Now each coefficient β_{nk}

[1] Conditional logit model for discrete choice analysis is also referred to as multinomial logit (MNL) model, or mix conditional logit model in the literature.

[2] Mixed logit model is also referred to as random parameter logit model or random effect logit model. Mixed logit model are abbreviated as MXL or MLM.

follows a distribution, with the mean $\bar{\beta}_k$ representing the mean parameter for the population.

The mixed logit choice probability is given by:

$$P_{nsj} = \int \frac{\exp(X'_{nsj}\beta + Z'_n\gamma)}{\sum_{j=1}^{J}\exp(X'_{nsh}\beta + Z'_n\gamma)} f(\beta|\theta)d\beta \quad (7)$$

Where $f(\beta|\theta)$ is the density function of β.

Mixed logit model can be estimated using maximum simulated likelihood (MSL) method.[238,239]

Latent class logit model (LCM)

In the mixed logit model, the distributions of coefficients are continuous. A discrete distribution of coefficients will lead to a latent class model. Latent class model assumes that individual behaviors depend on observable attributes and latent heterogeneity which are unobservable. In latent class models, individuals are implicitly sorted into different classes, however analyst does not know which class a particular individual belongs to.[236] A latent class model estimates a different set of coefficients for each class. The probability of individual n, whose is a member of class q, choosing alternative j in choice set s is given by:

$$P_{nsj|q} = \frac{\exp(X'_{nsj}\beta_q)}{\sum_{j=1}^{J}\exp(X'_{nsh}\beta_q)} \quad (8)$$

β_q is a vector of coefficient for class q. The probability of individual n being in class q can be specified as:

$$H_{nq} = \frac{\exp(Z'_n\gamma_q)}{\sum_{q=1}^{Q}\exp(Z'_n\gamma_q)} \quad (9)$$

where γ_q is a vector of coefficient for class q.

Latent class model can be estimated using maximum likelihood method, and the optimal number of classes can be selected based on AIC, BIC criteria.[236]

A summary of the three models

Under the framework of random utility model, various logit models are commonly used to analyze DCE data. The choice of model depends on the data characteristics. CLM which is the most basic model for DCE analysis was used to analyze the data, and the observed characteristics that influence choice probabilities were also identified. However, unobserved heterogeneity was not accounted for in CLM. To account for unobserved heterogeneity or other sources of unobserved variability, MXL and LCM were used to obtain more accurate estimates, and make predictions. The difference in MXL and LCM lies in the assumption of underlying distribution. In MXL model, respondents were assumed to come from the same underlying distribution, whereas in LCM, there were distinct distributions of preferences, and preferences for each class were estimated.

5.3.6.2 Variable specification and coding

The commonly used coding approaches for attribute levels in DCE are linear, dummy, and effect codes. For attribute levels that are categorical, dummy and effect codes both allows the estimation for each attribute level separately. While dummy coding uses only 0 and 1, effect coding uses 0, 1, and -1. See Table 14 for the effect coding template. The advantage of effect coding is that none of the levels has a coding of all 0's, as a result, none of the levels will be confounded with the grand mean (ie. the constant term in utility function).

Table 14. Effect coding template for categorical variables with 2 to 5 levels

	Coding Variable 1	Coding Variable 2	Coding Variable 3	Coding Variable 4
Level 1	1			
Level 2	-1			
Level 1	1	0		
Level 2	0	1		
Level 3	-1	-1		
Level 1	1	0	0	
Level 2	0	1	0	
Level 3	0	0	1	
Level 4	-1	-1	-1	
Level 1	1	0	0	0
Level 2	0	1	0	0
Level 3	0	0	1	0
Level 4	0	0	0	1
Level 5	-1	-1	-1	-1

Variable specification and coding used in analysis are shown in Table 15. Cost of test variable was assumed to be linear in utility function, for simplicity in willingness-to-pay estimates. It was also treated as categorical variable and effect coded to explore the linearity of this cost variable. Risk of developing SJS and cost of gout treatment were effect coded, as they were not expected to be linear. Doctor's recommendation was coded using two dummy variables, one indicating whether information on doctor's recommendation was available in a choice set, the other indicating whether an alternative was recommended by doctor. Most common choice was coded in a similar way using two dummy variables. All socio-demographic variables were coded using dummy variables. For categorical socio-demographic variables with more than two categories, some categories were combined, and collapsed into two categories for regression analyses (Table 16).

Table 15. Variable specification and coding type

Attributes	Levels	Variable in analysis	Variable coding type
Alternative type			
Whether genetic testing is involved	No Yes	D_{notest}	Dummy-coded
Alternative specific attributes (Specific to each test alternative)			
Cost of test	0 20 200 400 1000	cost	Linear
Risk of adverse side effect	0 1 in 50,000 1 in 5,000 1 in 1,000	risk0 risk1 risk2 (omitted)	Effect-coded
Cost of gout treatment	400 1000 2000 4000	Drug1 Drug2 Drug3 (Omitted)	Effect-coded
General/Context attributes (Specific to both test alternatives in a choice set)			
Doctor's recommendation	No information on doctor's recommendation Doctor recommended Not the doctor recommended	noinforec drrec (omitted)	Dummy-coded
Herd behavior	No information on herd behavior Most common choice Not the most common choice	noinfoherd herd (Omitted)	Dummy-coded

Table 16. Decision maker characteristics and coding

Continuous variables

Age=age in years

Income, linear=Monthly household income (in $1,000)[a]

Dummy variables

Gender, female=1 if female

Ethnicity, minority=1 if non-Chinese

Gout, hadgout=1 if had a diagnosis of gout

Hypertension, hypertension=1 if had hypertension

Self-reported health, healthy=1 if health status is quite good or very good

Housing type, housingbig=1 if HDB 5 room or private housing

Education level, eduhigh=1 if JC/Poly or above

Working status, working=1 if full-time/part-time/self-employed

[a] Household income was measured as categorical, and linearized assuming the average income of each category equaled the mean of the upper and lower bound of that category.

5.3.6.3 Utility function

There are four types of explanatory variables in the utility function: 1) alternative specific constant that accounts for the type of alternative (test involved vs. no test involved); 2) test feature attributes (cost of test, risk, cost of gout treatment) that vary across test alternatives; 3) context attributes (doctor's recommendation, and information on the most common choice), which vary across different choice sets; and 4) socio-demographic variables that may influence individual taste, and decision making process. The specification of these four types of variables in the utility function is as following.

The test alternative and no test alternative each gives some intrinsic utility associated with these two types of alternatives. As only the difference in utility from various alternatives matters for a decision, the intrinsic utility associated with test alternatives is normalized to 0, and the intrinsic utility associated with no test alternative is represented by an alternative specific constant (β_{ALS}). Three attributes (cost of test, risk and cost of gout treatment) always have fixed levels in the no test alternative throughout the survey, and the effect of them will be accounted for by β_{ALS}. There are several related assumptions. 1) The utility associated with the risk and costs of the no test option do not vary across choice sets, even though the contrast of risk and cost levels between testing options and no test option vary across choice sets. This is likely to hold, and is a common practice when including a fixed comparator (such as none option or status quo) in the choices sets. 2) Decision maker characteristics influence the testing decision by influencing the intrinsic utility associated with no test alternative, and therefore the tendency of an individual to choose no test alternative. The effect of socio-demographic variables is estimated as an interaction term with the dummy variable for no test alternative.

Utility derived from a no test alternative is:

$$U_{no\ test} = D_{no\ test}\beta_{ALS} + drrec\beta_1 + D_{no\ test} * drrec\beta_2 + D_{no\ test} * noinforec\beta_3 + herd\beta_4 + D_{no\ test} * herd\beta_5 + D_{no\ test} * noinfoherd\beta_6 + (Z * D_{no\ test})'\gamma + \varepsilon \quad (10)$$

where $D_{no\ test}$ is a dummy variable for no test alternative, β_{ALS} being the alternative specific constant for no test, $drrec$ is a dummy for being the doctor recommended alternative, and $noinforec$ is the dummy for having no information on doctor's recommendation. *herd* is a dummy for being the most common choice, and *noinforherd* is a dummy if there is no information on the most common choice. Z represents decision maker characteristics. $\beta_1\ to\ \beta_6$ are the utility weights of corresponding attribute levels, and γ reflects the impact of decision maker characteristics on utility. ε is the random error term.

Utility of a test alternative

Utility of a test alternative is specified as the following:

$$U_{test} = cost\beta_1 + risk1\beta_2 + risk2\beta_3 + risk3\beta_4 + drug1\beta_5 + drug2\beta_6 + drug3\beta_7 + drrec\beta_8 + herd\beta_9 + \varepsilon \quad (11)$$

where cost, risk1, risk2, risk3, drug1, drug2, drug3, drrec, herd describe test characteristics as specified in Table 15, and ε is the random error term.

Utility of any alternative is:

Based on equation (10) (11), the general utility function used in estimation can be written as:

$$U = D_{no\ test}\beta_{ALS} + cost\beta_1 + risk1\beta_2 + risk2\beta_3 + risk3\beta_4 + drug1\beta_5 + drug2\beta_6 + drug3\beta_7 + drrec\beta_8 + D_{no\ test} * drrec\beta_9 + D_{no\ test} * noinforec\beta_{10} + herd\beta_{11} + D_{no\ test} * herd\beta_{12} + D_{no\ test} * noinfoherd\beta_{13} + (Z * D_{no\ test})'\gamma + \varepsilon \quad (12)$$

5.3.6.4 Model estimation

Five models were used in the analyses:

1) CLM1: Conditional logistic regression without control variables
2) CLM2: Conditional logistic regression with control variables
3) MXL1: Mixed logistic regression without control variables, and all attributes except test cost were set random
4) MXL2: Mixed logistic regression with control variables, all attributes except test cost were set random
5) LCM1: latent class model

Conditional logit model, mixed logit model and latent class models were conducted using Stata version MP11, with and without controlling for decision maker characteristics.

Conditional logit model was estimated using the *clogit* function. The user written *mixlogit* function was used to estimate the mixed logit model, with all attributes except cost of test specified as random. Cost of test is assumed to be fixed a priori, as is the common practice in mixed logit estimates [240,241]. Specifying prices to be random will give rise to problems in willingness-to-pay estimation due to scale heterogeneity. *mixlogit* fits the model based on maximum simulated likelihood.[242] Latent class model was fitted using the user-written commands *lclogit* and *lclogitml*. *lclogit* uses an expectation-maximization algorithm for estimation.[243,244] Two latent classes were specified, and control variables that were significant in the conditional logit model were included in the fractional multinomial logit model of class membership. The optimal number of latent classes were determined based on the best model fit. The model with 2 latent classes was found superior to the simpler one-class model according to the AIC and BIC goodness-of-fit statistics. Models with more than 2 latent classes failed to converge due to the high number of parameters to estimate relative to the sample size.

5.3.6.5 Preference weights and attribute importance

Coefficients of categorical variables from three regression models were plotted to illustrate the relative importance of attributes and attribute levels, or preference weights of each attribute level. Attribute levels with larger preference weights were preferred to those with smaller preference weights. A greater distance between the best and worst level of an attribute indicates higher importance of the attribute.

5.3.6.6 WTP estimates

As the primary motivation to take genetic test is to reduce the risk of developing SJS, an indicator of interest is the willingness-to-pay for various levels of risk reduction. Assuming the cost of test to be linear, willingness-to-pay for risk reduction (K2 to K1) can be calculated using the following formula:

$$WTP = -\frac{(\beta_{K1} - \beta_{K2})}{\beta_{cost}}$$

where β_{cost} is the coefficient of test cost attribute, β_{K1} and β_{K2} are the coefficient of risk level K1 and K2.

The marginal willingness-to-pay for an alternative when it is doctor recommended was calculated as:

$$WTP = -\frac{\beta_K}{\beta_{cost}}$$

Where β_K is the coefficient of the attribute level “doctor recommended”.

Similarly, the marginal willingness-to-pay for an alternative when it is the most common choice was calculated.

5.3.6.7 Uptake rate prediction

At the system level, uptake rate can help to visualize the effect of policies or changes in test features. Hypothetical choice sets with one test alternative and one no

test alternative was constructed, and the uptake rate of the test was predicted for different scenarios.

Using utility weights estimated from the models, utility scores associated with different alternatives can be calculated from the utility function, which can be used to predict the uptake probability for any hypothetical test and scenarios. For instance, in a hypothetical choice set with a test alternative (T) and a no test alternative (N), the uptake probability of the test alternative is

$$P(T) = \frac{\exp(V(T))}{\exp\big(V(T)\big) + \exp(V(N))}$$

5.4 Results

5.4.1 Sample characteristics

A total of 205 Singaporean diabetes patients were recruited, among whom 199 completed the survey questionnaire. 10 respondents were excluded from analysis due to prior long-term gout treatment with allopurinol. 189 respondents were included in the final analysis. Respondents' characteristics are summarized in Table 17. The average age was 57.1 years (95% CI: 55.3 to 59.0). Respondent were mostly male (65.6%), Chinese (61.4%), and currently working (55%). 35.4% had completed Junior College (JC)/diploma or university education, and 51.9% stayed in HDB 5 room or private properties. A small percentage (5.8%) of respondents had gout or hyperuricemia, but did not receive urate-lowering therapy. A significant percentage of respondents had experiences with serious adverse drug reactions (13.8%).

Table 17. Characteristics of 189 respondents

Variables	N (%)
Age	57
Male	124 (66%)
Had Gout or hyperuricemia	11 (6%)
Ethnicity	
Chinese	116 (61%)
Malay	19 (10%)
Indian	50 (27)
Other	4 (2%)
Hypertension	111 (59%)
Highest education attained	
No formal education	3 (2%)
Primary	22 (12%)
Secondary	97 (51%)
JC/polytechnic/diploma	36 (19%)
University and above	31 (16%)
Housing type	
HDB (1-2 room)	19 (10%)
HDB (3 room)	23 (12%)
HDB (4 room)	48 (25%)
HDB (5 room and above)	58 (31%)
Condominium/Private flat	20 (11%)
Bungalow/semi-detached/terrace house	20 (11%)
Self-rated health status	
Very good	6 (3%)
Quite good	75 (30%)
Neither good nor poor	77 (41%)
Quite poor	29 (15%)
Very poor	2 (1%)
Experiences of severe adverse drug reaction	26 (14%)
Employment status	
Full-time employed	91 (48%)
Part-time employed	13 (7%)
Self-employed	12 (6%)
Homemaker	7 (4%)
Retired	55 (29%)
Unemployed	11 (6%)
Household income	
S$0-1,500	34 (18%)
S$1,500-3,000	50 (27%)
S$3,000-5,000	30 (16%)
S$5,000-8,000	25 (14%)
S$8,000-10,000	18 (10%)
Above S$10,000	28 (15%)

5.4.2 Analysis of warm-up questions

Attitudes towards SJS risk, cost of test, and cost of long-term gout treatment

When the risk of developing severe adverse drug reaction was fixed at 1 in 500 (0.2%), 55% of respondents felt at risk, and the rest were not worried about it. If a hypothetical test could reduce risk of SJS, but cost S$400, 52% of respondents expressed willingness to take the test (definitely would or probably would). When told that the long-term gout treatment cost for positive test result (2 in 10 chance) was S$2,000 in two years, 57% of respondents expressed willingness to take the test and receive test-guided treatment (definitely would or probably would).

Attitudes towards doctor's recommendation and most common choice

Most respondents reported doctor's recommendation to be influential on their decision making. 51 % of respondents definitely would consider doctor's recommendation, and 33% probably would consider. On the other hand, 15% of respondents probably or definitely would not consider doctor's recommendation, and preferred to make independent decisions. 49 % of respondents considered information on the most common choice to be influential on their decision, whereas 51% of respondents would not consider this piece of information.

5.4.3 Validity test

In the first validity test, 89.4% of respondent managed to identify that "1 out 500 patients gets the severe side effect" indicated higher risk than "1 out of 1,000 patients gets the severe side effect" at the first attempt, and another 6.9% answered it correctly in the second attempt, both of which were considered to have good understanding of probabilities. In the second validity test, which was a "dominant-pair" test DCE question, 13.2% preferred the higher risk test alternative in the first attempt, but corrected the answer in the second attempt. Another 2.1% of respondents indicated preference for the higher risk test alternative more than once, indicating confusion or lack of understanding of the probability or the DCE question.

Failure of validity tests was defined as giving the wrong answer more than once in both validity test questions, and 0.5% of respondents failed the validity tests. These respondents were not excluded from final analysis because 1) excluding these respondents did not significantly change the regression results; and 2) respondents who failed initially may understand questions correctly in later parts of DCE, as they learned more about the survey questions.

5.4.4 Non-demanders for genetic testing

Non-demanders refer to those respondents who prefer not to receive test services when offered, regardless of attribute levels.[213] Among 189 respondents, 16 (9%) chose the no test alternative in all choice sets, referred to as non-demanders. Note that no test always resulted in lowest costs; therefore these non-demanders may have strong preferences for low cost, or for the no test label. 91 (48%) respondents chose no test in at least one choice sets. Out of 1,512 choice observations (8 from each respondent), no test was preferred in 317 (21%) observations, On the other hand, 98 (52%) never preferred the no test alternative.

Logistic regression analyses showed that respondents of older age were more likely to always prefer no test, whereas being non-Chinese ethnicity, with JC/poly or above education, currently working were less likely to be non-demanders for genetic test ($p<0.05$ for all factors mentioned above).

5.4.5 Dominant Preferences

22 (12%) respondents were dominant on the risk of developing SJS, and always preferred the alternative with lower risk of SJS, regardless of costs, doctor's recommendation or most common choice. Those experienced serious adverse drug reactions, with JC/poly or above education, were more likely to dominant on the risk of getting SJS ($p<0.1$ for all factors mentioned above). Dominance on costs is not discussed here, as it cannot be disentangled from the preference for no test label.

21 (11%) of respondents were dominant on doctor's recommendation, and always chose the doctor recommended alternative when information was available. Malay or Indian ethnicity was a predictor of this dominance preference ($p<0.05$). In contrast, only 3 (2%) of respondents were dominant on most common choice, and always chose the alternative that was labeled the most common choice.

Both non-demanders and dominant preferences deviate from standard assumption that individual make tradeoffs between different attributes in discrete choice experiments. The non-tradeoff may confound the estimated preference weights estimates. However, the percentage of non-demanders and dominant preferences was relatively low in the sample (34% in total), and was therefore not a major concern.

5.4.6 Results from logit models

5.4.6.1 Model fits

Table 18 compares the goodness-of-fit of different logit models. MXL and LCM models which allow for unobserved preference heterogeneity, significantly improved the fit compared to CLM, as indicated by the increase in log-likelihood, and decrease in AIC and BIC. Controlling for decision maker characteristics in both CLM and MXL improved the fit of models, even though more parameters need to be estimated. However these metrics cannot be used to compare MXL and LCM, as the models were not nested, and the base model was different.[236]

Table 18. Goodness-of-fit of models

	CLM1	CLM2	MXL1	MXL2	LCM
Model Type	Conditional logit	Conditional logit	Mixed logit	Mixed logit	Latent class (2 classes)
Decision maker characteristics	-	Controlled	-	Controlled	Controlled
Log-likelihood	-1391	-1307	-1035	-1002	-1144
Pseudo R^2 [a]	0.16	0.21	-	-	-
AIC[b]	2811	2662	2124	2073	2360
BIC[c]	2901	2816	2297	2291	2592

[a]Pseudo R^2 is defined as 1-LL/LL0, where LL is the simulated log-likelihood function evaluated at the estimated parameters, while LL0 is the value of a log-likelihood function for a base model that only contains a non-random alternative-specific constant.

[b]AIC=-2(LL-M) where M is the number of parameters

[c]BIC=-2LL+MlnN where N is the number of observations.

5.4.6.2 Model estimates from CLM and MXL models

Estimates from CLM, MXL, and LCM are shown in Table 19. Note that the estimates from different models are not directly comparable due to scale differences[52].[166,236] In MXL, the estimates are normalized relative to the extreme value part of the error term, which is the net of the error components introduced by the random coefficients. In CLM, the error term captures both sources of error, and therefore it will have a larger variance. Therefore estimates from CLM are expected to be smaller than those in MXL, which is consistent with our observation that in Table 19, coefficients from MXL are larger in magnitude than CLM coefficients. Nevertheless, signs from the models can be compared.

Table 19. Estimates from CLM and MXL models

	Conditional logit		Mixed logit, all coefficients random except cost of test			
	CLM1	CLM2	MXL1		MXL2	
	Mean	Mean	Mean	SD	Mean	SD
ALS_{notest}	-0.70**	0.50	-2.77**	4.88**	-0.75	4.79**
Cost of test (in $1,000)	-0.46**	-0.49**	-0.83**		-0.90**	
Risk of SJS: 1 in one million	0.45**	0.47**	0.77**	1.19**	0.87**	1.24**
Risk of SJS: 1 in 5,000	0.23**	0.23**	0.35**	0.19	0.31**	0.13
Risk of SJS: 1 in 1,000	-0.15*	-0.16*	-0.22*	0.12	-0.21**	0.12
Risk of SJS: 1 in 600	-0.54**	-0.55**	-0.90**	1.26**	-0.97**	1.26**
Cost of treatment: $250	0.37**	0.38**	0.68**	0.72**	0.75**	0.69**
Cost of treatment: $400	0.41**	0.43**	0.65**	0.58**	0.55**	0.82**
Cost of treatment: $1,500	-0.03	-0.04	-0.11	0.07	-0.08	0.04
Cost of treatment: $4,000	-0.75**	-0.76**	-1.22**	1.23**	-1.22**	1.47**
Doctor recommended	0.80**	0.81**	1.23**	1.6**	1.3**	1.53**
Doctor recommended*notest	-0.04	0.12	0.99*	1.22*	1.37**	1.83**
No recommendation available	0.09	0.09**	0.22	0.18	0.39	0.14
Most common choice	0.31**	0.30**	0.52**	0.56**	0.50**	0.74**
Most common choice * notest	-0.16	-0.09	0.45	0.36	0.70	0.38
No Most common choice info available	0.14	0.14	-0.09	0.04	-0.06	0.12
Control variables * no test						
Female gender		0.07				
Age		0.00				
Ethnic minority		-0.82**			-0.46	
Big housing		-0.50**			-3.59**	
High education level		-0.66**			-0.67	
Household income (in $1,000)		0.04*			0.23**	
Currently working		-1.32**			-3.98**	
Had diagnosis of gout		0.62*			1.24	
Had severe ADRs		-0.76**			-1.24*	
Self-reported to be healthy		0.02				

** p<0.01, *p<0.05

Intrinsic preferences for no test

The negative alternative-specific intercept ALS_{notest} in Model CLM1 and MXL1 implies that no test resulted in disutility, and patients had intrinsic preferences for taking a test to reduce risk of severe adverse drug reactions. However, there is considerable heterogeneity in the preference for no test in the sample, evidenced by the significance and big magnitude of standard deviation estimate in mixed logit model MXL1. ALS_{notest} from model CLM2 and MXL2 cannot be interpreted alone, as all decision maker characteristics were estimated as interaction terms with no test alternative, and assumed to influence the utility individuals derived from the no test alternative. Model MXL2 shows that respondents staying in big or private housing, currently working, or had experienced severe ADRs had higher disutility from no test, and therefore are more likely to take the genetic test. Unexpectedly, respondents with higher household income were more likely to choose no test. Yet this finding could be confounded by housing type and working status.

Preferences towards test features

Consistent with hypotheses, respondents were more likely to test when the test cost was lower, when the test-guided treatment results in lower risk of developing SJS, and when long-term gout treatment cost was lower. SD estimates from MXL models revealed significant diversity in the way people value these attribute levels.

The effect of information on doctor's recommendation and most common choice

Doctor's recommendation on a test alternative significantly improved the likelihood of that alternative being chosen. Interestingly, the interaction term of doctor's recommendation with no test alternative was significant and had a positive sign, suggesting that when the no test alternative was recommended by doctor, the increase in likelihood of it being chosen was even more than when a test alternative was recommended. The significance of standard deviation estimates suggested considerable individual differences in the valuing of doctor's recommendation in the sample. On the other hand, as expected, when no recommendation was available, there was no impact on respondents' choices.

Information on the most common choice also had positive impact on the likelihood of an alternative being chosen, even though the effect was much smaller than that of doctor's recommendation. The effect did not seem to be different for test and no test alternatives. Significant individual heterogeneity was evidenced by the big standard deviation estimates.

5.4.6.3 Attribute importance based on MXL model

The relative importance of various attributes can be inferred from the estimated coefficients. For dummy attributes, larger magnitude of the coefficient indicates the importance of the attribute. ALS_{notest} was largest in magnitude among all coefficients, suggesting respondents had very strong preferences to avoid no test alternative in general. Doctor's recommendation also had big coefficient, suggesting it was very influential on respondents' preferences. On the other hand, information on most common choice was less influential. For effect-coded attributes, a greater difference between coefficients of the best and worst level of attribute indicates a greater significance of that attribute in influencing decisions, within the range provided in this study. To compare the relative importance of three test feature attributes, another mixed logit model (referred to as MXL1c) which is a modification of model MXL1 in which the cost of test was also specified as categorical. The preference weights of three test feature attributes were plotted in Figure 12 to better compare the relative importance of these attributes. From Figure 12, long-term gout treatment cost and the risk of developing SJS in test-guided treatment were more important than the cost of test. Respondents were not sensitive about the cost of test within the range of $200 to $400. Similarly, respondents were not sensitive to the drug cost for test positive individuals as long as it was lower than $400. However when the cost was increased to $4,000 in two years, it significantly discouraged testing. The four risk levels are well segregated, and resulted in a wide range of preference weights, indicating respondents were risk-conscious. A 1 in 1 million chance of developing SJS in test-guided treatment (ie. a very accurate test) results in the highest probability of test uptake, increase in risk (due to reduced accuracy of test) reduces probability of test uptake.

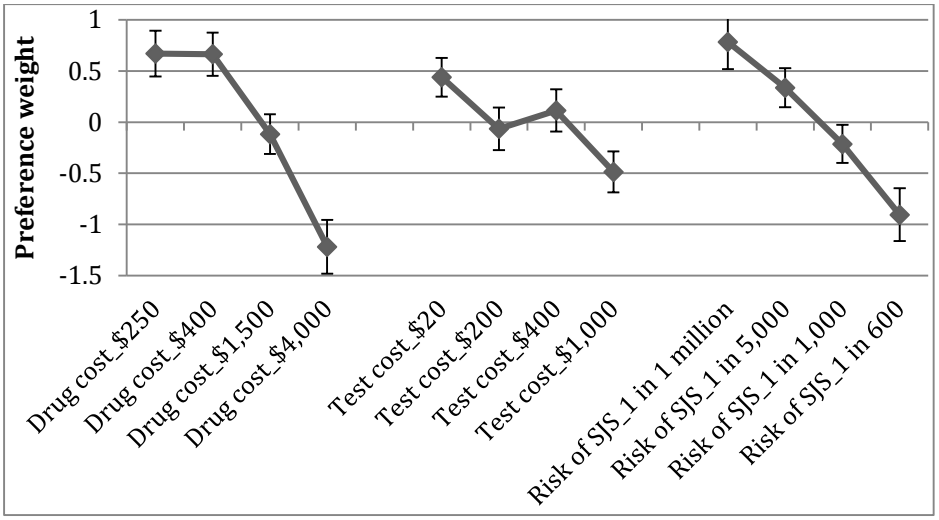


Figure 12. Preference weights of various test feature attributes. Error bars indicate 95% confidence intervals.

The linearity of test cost attribute was also explored using estimates from mixed logit model MXL1c. As shown in Figure 13, preference weights were roughly linear in test cost, though not perfect.

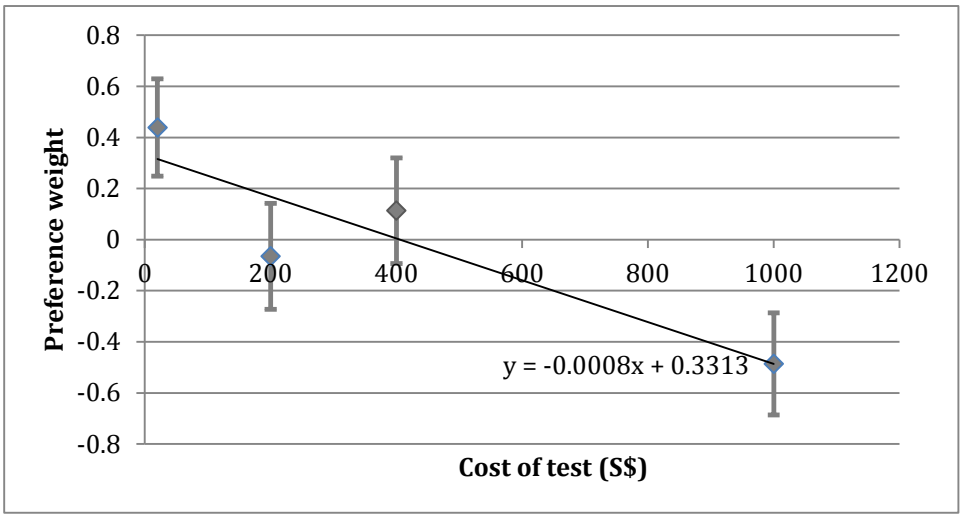


Figure 13. Preference weights for cost of test.

5.4.6.4 Model estimates from LCM models

Latent class model with 2 classes are shown in Table 20. The optimal number of classes was selected based on AIC BIC criteria. The 3-class model failed to converge, likely because the sample size did not have sufficient power for 3 classes.

Table 20. Estimates from CLM and MXL models

	Latent class 1 (cost conscious) 37%	Latent class 2 (risk averse) 63%
ALS_{notest}	1.04**	-4.95**
Cost of test (in $1,000)	-1.26**	-0.37**
Risk of SJS: 1 in one million	-0.29	0.75**
Risk of SJS: 1 in 5,000	0.09	0.24**
Risk of SJS: 1 in 1,000	0.03	-0.22**
Risk of SJS: 1 in 600	0.17	-0.77**
Cost of treatment: $250	0.26	0.44**
Cost of treatment: $400	0.67**	0.46**
Cost of treatment: $1,500	-0.13	-0.10
Cost of treatment: $4,000	-0.79**	-0.80**
Doctor recommended	1.30**	0.69**
Doctor recommended*notest	0.07	1.01
No recommendation available	0.18	0.50
Most common choice	0.81**	0.28**
Most common choice * notest	-1.02*	2.21**
No Most common choice info available	-0.47	0.18

** p<0.01, *p<0.05

The two classes generated from the latent class model are named as “risk averse” class, and “cost conscious” class based on their preference weights for different attributes (Figure 14 and Table 20). 37% of respondents fall in the “cost conscious” class, whereas the rest 63% belong to the “risk averse” class. The “risk averse” class has a big negative ALS_{notest}, indicating the disutility results from not testing. The range of preference weights for risk of SJS is wider than that of test cost, implying that this group is more concerned about the risk of developing SJS, and cost has relatively small impact. On the contrary, the “cost conscious” class had a wide range of preference weights for cost of test, implying their

decision is sensitive to test cost. Though no test gave disutility, the magnitude of disutility is moderate. In this class, none of the risk levels turned out to be significant, implying that this class was more averse to high cost than risk of developing SJS. Both groups considered cost of long-term gout treatment important, and high treatment cost reduces the probability of test in both classes. Doctor's recommendation and most common choice affect both classes of respondents.

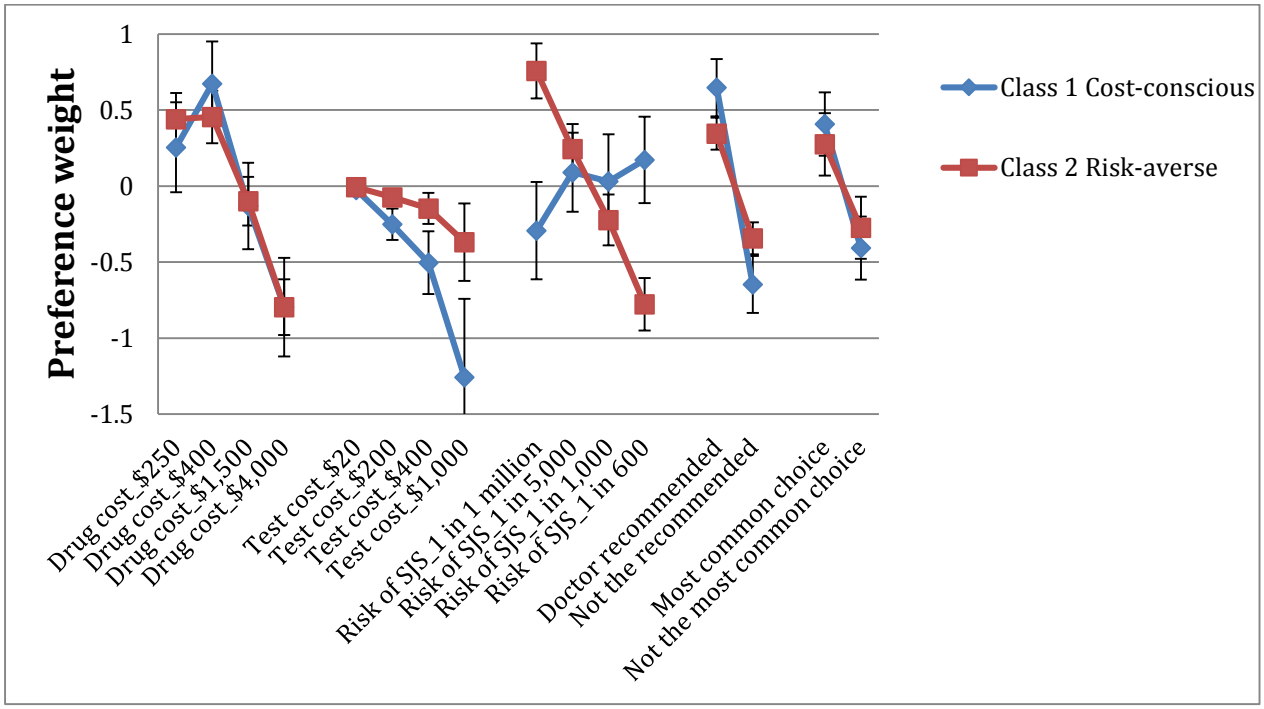


Figure 14. Preference weights in latent class model

The association of class membership and individual characteristics are shown in Table 21. Most factors did not appear to determine class membership, except that individuals who were working were less likely to be in the cost conscious class ($p<0.1$).

Table 21. Class membership and individual characteristics

	Coefficient	P value
Stay in big housing	-0.36	0.30
Ethnic minority	-0.28	0.43
Severe ADR	-0.36	0.46
Currently working	-0.65	0.06
High education level	-0.55	0.16
Monthly household income	0.02	0.66
Had diagnosis of gout	0.74	0.26

5.4.6.5 WTP estimates from MXL and LCM models

Though coefficients from different models are not directly comparable, the willingness-to-pay estimates are. The willingness-to-pay for various attribute level improvements is listed in Table 22.

The current HLA-B*5801 test can reduce the risk of developing SJS from 1 out of 500 to below 1 out of 1 million. MXL model predicts that on average, respondents were willing to pay close to $2,000 more to do this genetic test compared to a test that can only reduce the risk to 1 in 600. The presence of information on doctor's recommendation and most common choice influenced individual's willingness-to-pay for genetic test. When a test was recommended by the doctor, on average respondents were willing to pay S$1,474 more for the test (95% CI: 817, 2,131), compared to when the test was not the doctor recommended option. Similarly, when a test was labeled the most common choice, respondents were willing to pay S$623 more for the test (95% CI: 249, 997), compare to when it was not the most common choice.

Estimates from LCM shows distinct preferences across the two classes. In class 1 (cost-conscious class), the WTP for risk reduction was negative and statistically not significant (from zero). In class 2 (risk-averse class), on average, respondents would be willing to pay S$1,215 to reduce the risk of developing SJS to 1 out of a million from 1 out of

500. Both classes had high WTP for testing when test was recommended by doctor (S$1,032 and S$548 for class 1 and class 2), or if it is the most common choice (S$647 and S$219 for class 1 and class 2). Class1 had higher WTP than class 2, indicating the information was more influential among class1 members. WTP estimates from LCM are lower than in MXL, especially in the cost-conscious class. Considering the non-demanders observed in the sample, the WTPs estimated using LCM were more realistic.

Table 22. Willingness-to-pay for attribute improvements (in S$)

		WTP (95%CI)		
		MXL1	**CLM- cost conscious class**	**CLM- risk averse class**
Risk reduction	1 out of 600 → 1 out of 1,000	818 (351, 1284)	-109 (-491,274)[NS]	438 (161, 716)
	1 out of 600 → 1 out of 5,000	1,504 (835, 2173)	-62 (-406, 280)[NS]	809 (415, 1,202)
	1 out of 600 → 1 out of 1 million	1,999 (1,153, 2,846)	-368 (-808,73)[NS]	1,215 (662, 1,768)
Information on doctor's recommendation and herd behavior	Not the recommended→Doctor recommended	1,474 (817, 2,131)	1,032 (535, 1,530)	548 (265, 830)
	Not the most common choice→most common choice	623 (249, 997)	647 (211, 1,081)	219 (34, 405)

[NS] Not significantly different from 0.

5.4.6.1 Test uptake rate estimates from MXL and LCM models

The uptake rate was forecasted for eight clinically relevant or policy relevant scenarios. This provides insights on the impact of various policies or clinical practices on uptake rate. When only one test alternative and one no test alternative was offered, the predicted uptake rates were shown in Table 23.

The uptake rate predicted from mixed logit model are high in all scenarios, and the change in uptake rate across different scenarios were small. The overall uptake rate predicted by latent class model are lower than that predicted by mixed logit model in all scenarios. Given the non-demander behaviors observed, the uptake rate predicted by latent class model may be more realistic. The uptake rate in class 2 were above 95% in all scenarios, and the changes in uptake were small when test features or information changed; whereas in class 1, uptake rate differed significantly across scenarios. When the long-term cost of gout treatment was reduced from $4,000 to $400, uptake rate was significantly increased in class 1 (29.4% vs 8.8%, p=0.0001), suggesting that the use of cheaper generic drugs for those who test positive can significantly increase the uptake of the test. On the other hand, when the cost of test was subsidized by 75%, the uptake rate improved, but the effect was small (12.3% vs 8.8%, p=0.0162). For patients who are more cost-conscious, one alternative may be a less accurate but cheaper test. When available, the cheaper alternative is more preferred by class 1 (16.2% vs 8.8%,p=0.0409). Compared to the small effect of cost strategies on test uptake, information strategies had bigger impacts on the uptake rate in class 1. When the current test is recommended by doctor, the predicted uptake rate increases to 29.7% from 8.8% ($p<0.0001$). When the current test is the most common choice, the uptake rate was predicted to increase by 3%. When the test was recommended by doctor and was the most common choice, the uptake rate was 37.3%, suggesting a synergistic effect of the information. When the test is recommended by the doctor, but the most common choice is not to test, the uptake is still higher than without information (24.5%

vs 8.8%, p=0.0157), suggesting the impact of doctor's recommendation is much bigger than that of herd behavior.

Table 23. Test uptake rate in various hypothetical scenarios

Hypothetical scenarios		A **A realistic test (base case)**	B **Cheaper alternative drug**	C **75% subsidy on test cost**	D **A less accurate test, that is 75% cheaper**	E **A clinical guideline that requires doctors to recommend test**	F **Provide information that test is most common choice**	G **Provide information that test is doctor recommended and the most common choice**	H **Provide information that doctor recommend test, but test is not the most common choice**
Attribute levels	Chance of getting the severe side effect (with test-guided treatment)	1 out of 1 million patients	1 out of 1 million patients	1 out of 1 million patients	1 out of 1,000 patients	1 out of 1 million patients	1 out of 1 million patients	1 out of 1 million patients	1 out of 1 million patients
	Cost of one-test genetic test	S$400	S$400	S$100	S$100	S$400	S$400	S$400	S$400
	Cost of gout medicines (over 2 years)	S$4,000	S$400	S$4,000	S$4,000	S$4,000	S$4,000	S$4,000	S$4,000
	Your doctor's recommendation	No information	No information	No information	No information	Test	No information	Test	Test
	Most common choice	No information	No information	No information	No information	No information	Test	Test	No test
Test uptake rate	**Test uptake rate (MXL estimates)**	86.4%	97.6%	89.1%	75.3%	96.4%	90.7%	97.7%	90.4%
	Test uptake rate (LCM overall estimates)[a]	65.10%	73.49%	66.52%	66.46%	73.60%	66.62%	76.55%	68.84%
	Test uptake rate (LCM class 1 estimates)	8.8%	29.4%	12.3%	16.2%	29.7%	11.9%	37.3%	24.5%
	Test uptake rate (LCM class 2 estimates)	98.3%	99.5%	98.5%	96.1%	99.5%	98.9%	99.7%	95.0%

[a]Overall test uptake rate was calculated as uptake in class1 * class 1 share+ uptake in class2 * class 2 share

5.5 Discussions

5.5.1 Patients' attitude towards pharmacogenetic testing to reduce risk of severe ADR

This is the first study to use a discrete choice experiment to quantify patients' preferences for using pharmacogenetic testing to reduce severe ADRs. Our study revealed that the majority of patients were willing to adopt risk-mitigation strategies such as genetic testing. Our uptake rate prediction shows that given the current available test and treatment (test cost=S$400, SJS risk=1 out of 1 million, gout treatment cost=S$4,000 in two years, no information on doctor's recommendation and most common choice), 65% of the eligible patients were willing to test and receive test-guided treatment (Table 16). When attribute levels were varied to form hypothetical tests, 92% of respondents preferred test to no test in at least one of the 8 scenarios offered. 52% always preferred testing to no testing in all 8 scenarios.

In the research and development of diagnostic tools, test sensitivity, specificity, false negative, false positive are the most important features researchers consider. Clinicians and researchers are usually concerned about high false negative or false positive rates, as these false results may lead to inadequate or redundant treatment, both may have adverse consequences. The current available HLA-B*5801 test can reduces risk to below one in one million, however, it has high false positive rate, specifically, 20% of patients may test positive and require more expensive gout treatment, when in fact over 95% of test positive patients would not develop SJS even if taking allopurinol. Cost-effectiveness analysis showed that the extra cost due to the high false negative rate of test compromised the test to be non-cost-effective at the population level. However, we showed here, 65% of all respondents are willing to take the test to reduce risk of severe ADR, even though high cost may be incurred.

In fact, qualitative remarks from patients showed that the genetic test might have positive "value of information", regardless of test results. A positive test results

may help to reduce the risk of SJS, whereas a negative result gives the confidence and assurance that the patient will not develop SJS. Patients may take the test for the peace of mind, and derive utility from both positive and negative test results. The same may be true for physicians. Allopurinol ADR has led some physicians to reduce allopurinol prescription and switch to the safer but more expensive new drug febuxostat. Genetic testing may lead to more confident use of allopurinol, which may result in cost-saving for the health system.

5.5.2 Tradeoffs between test features and heterogeneity in preferences

Advancement of technologies may alter test characteristics in the near future. For instance, advancement in genetic testing technology may improve the accuracy and predictive power of testing, while reducing the testing cost. Low cost alternative risk mitigation strategies may be developed. The cost of long-term gout treatment may be lower when the drug patent expires. It is useful to understand how potential changes influence patients' testing decisions. We investigated the tradeoffs patients made between various test features (cost of test, cost of long-term gout treatment, and the SJS risk associated with test-guided treatment), estimated the willingness-to-pay for risk reduction, and predicted the stated uptake rate.

Latent class analyses revealed two distinct classes of decision makers that assign different preference weights to various features. One class considered cost of test and long-term treatment more important than the level of risk, so this class was described as the cost-conscious class. The other class had high preference weights for risk level, and relatively small weights for the cost of test. This class was named the risk-averse class. Consistently, the risk averse class derived disutility from no test, and therefore more likely to test, whereas the cost-conscious class derive positive utility from no test, likely due to the fact that no test incurs the lowest costs in the short term and in the long term. As expected, the risk averse class has higher uptake rate than

the cost-conscious class, when offered the same test (98% vs 65%, p<0.01). The uptake rate and willingness to pay among risk averse patients depends heavily on risk level. Contrarily, no significance was detected for the cost-conscious group. These implies that the risk averse class are willing to take a more expensive test that is more accurate, and the cost conscious class may be willing to adopt a cheaper and less accurate risk mitigation strategy. A feasible alternative strategy to genetic testing is a safety monitoring program that aims to identify early signs of SJS and reduce the severity and mortality associated with SJS. This cheaper alternative strategy is more preferred to testing among the cost conscious group. Interestingly, the class membership was not significantly correlated with socio-demographic factors, suggesting the belief may not be determined by their ability to pay.

The long-term treatment cost attribute has an uncertainty component as treatment is selected based on test results. To shed some lights on individual's interpretation about this attribute, the estimated willingness-to-pay was compared to the expected value of treatment cost. When the treatment cost for test positive individuals was reduced from S$4,000 to S$400, the expected payment decreased from by $720 (from $960 to $240). However, WTP analysis shows that in both classes, the WTP in one-time test cost increased more than the decrease in expected value ($1,166, and $3,380 respectively for the two classes). This implies respondents do not do expected value calculation. Rather, they were willing to pay extra to avoid the risk of extremely high cost. Even though each individual has a small chance (20%) of requiring the expensive treatment, a high long-term treatment cost significantly discourages testing.

5.5.3 Impact of doctor and herd

Genetic testing decision in real clinical settings not only depends on the features of testing. The choice context in which the test options are offered, and the way testing information is delivered also matters. We found physician recommendation

to be very important in decision making. 8% of respondents always preferred the doctor recommended option, even when the recommended option had high risk, or was expensive. This is consistent with other studies that found strong impact of doctor's recommendation on patient choices.[166,182-187] Random information on doctor's recommendation appears to be influential in both cost-conscious class and risk-averse class. When a test was recommended by doctor, respondents were willing to pay more for the test compared to when the test was not recommended. This suggests respondents infer about the quality and value of test based on doctor's recommendation. The uptake rate prediction showed consistent results. Doctor's recommendation appeared to be more influential on ethnic minorities. Compare to doctor's recommendation, herd behavior had smaller impact on patients' testing decisions. Our findings confirmed the presence of herd behavior in patients' genetic testing decisions, even though small in magnitude. Our findings are consistent with the limited literature in medical decision making. In an online survey on women's breast cancer treatment choices, when information on social norm suggested chemotherapy to be popular, women showed higher interest in chemotherapy.[245] Hall *et al.* found small but significant impact of information that "80% of people like you have been tested for Tay Sachs Disease (TSD)" on the stated uptake of TSD screening test.[166] Yang *et al.* studied prescription of antipsychotic drugs by physicians, and found that the prescription could be influenced by their colleagues in the same hospital. Though the peer effect was small, it was more influential on the prescription of new drugs.[204]

To many patients, recommendation by physicians or being the most common choice indicates the quality of a treatment option.[195] Following doctor's recommendation or the most common choice may therefore be a decision heuristics that allows easy and fast decision making.[193,194] Some decision making theories suggests the objective of decision making is not to maximize utility, but to simplify decision making.[169,246] Though these decision heuristics may deviate from the "rational" decision making pathway based on logic and calculation, some empirical findings and

economic theories have recognized the ecological rationality and advantage of decision heuristics.[193-195] For instance, Banerjee argues that in a sequential decision model, it is rational for decision makers to look at the decisions made by the previous decision makers, as other decision makers may have information that is important. Furthermore, he demonstrated that the optimizing strategy is to do what other people do, rather than using their information.[195]

5.5.4 Implication on genetic testing practices and policies in Singapore

This study provides timely information for various stakeholders such as test providers, physicians, and regulators. We observed a demand for routine screening in primary care and hospital settings. Alternative risk-mitigation tools that are cheaper but less accurate (such as a less accurate test or a safety monitoring program) may also be desirable to some patients. Our study also highlighted three important aspects of physician practices regarding genetic testing. First, information on ADR risk and genetic tests is perceived important by patients, so it's desirable for doctors to educate and effectively communicate with patients on ADR risk, potential risk-mitigation strategies, as well as cost and long-term consequences of genetic testing. Secondly, doctor's opinion and recommendation play a very important role in shaping patient's decision. So it is important for doctors to consider patient preferences, and recognize the heterogeneity in patients preferences when making recommendations, in addition to professional judgment of the risk and benefit profile. Third, some patients had preferences for cheaper long-term gout treatment, and are willing to bear with higher risk of SJS for lower treatment cost, which implies bypassing allopurinol and switching to febuxostat for risk considerations may not be desirable. From the regulator perspective, the value of genetic test was recognized by patients, however, a mandate is not justifiable as a significant proportion of patients prefer not to test, and the test cost is currently paid out of pocket. Subsidizing may be a feasible way to improve

uptake, however a group of price elastic individuals should be targeted, instead of the whole population. Cost-effectiveness and equity are also important considerations when determining subsidies.

5.5.5 Implication on health promotion strategies in Singapore

Various types of policy tools are commonly used for health promotion, including education campaigns and information provision, making guidelines and mandates, and financial incentives. Evaluating policy impact in real life settings require considerable effort and resources, and may be subject to confounding. A comparison of various policy tools in the same setting is even more difficult. Moreover, it is desirable to have information on the policy effects ex ante. This study demonstrates a feasible and less resource demanding way to evaluate and compare various policy effects. Moreover, the latent class analysis suggests there may not be a one size fits all policy strategy. Different subgroups may require different policy solutions to achieve the most impact. It is therefore important to understand heterogeneity in target population behaviors and preferences before designing and implementing a health promotion policy.

Providing detailed information on the benefits and costs of genetic testing in comparison to no test result in significant rate of stated test uptake compared to no detailed information. This suggests that for services individuals are not familiar with, information provision may increase uptake.

Based on stated preferences in this study, providing information on physician's recommendation is likely to be an effective and low-cost health promotion strategy. Efforts to alter physician behaviors such as recommendation by regulators or clinical practice guideline may achieve even stronger effects, as the physical presence of physician to deliver the recommendation achieves extra persuasive power than information on physician's recommendation. Providing information on the most common practice may nudge individuals towards following the herd. Even though less effective than physician's recommendation, it is a low-cost information strategy. Both

types of information strategies appeared to influence most individuals. In our study, providing information on doctor's recommendation and herd behavior increased the stated uptake rate more than a 75% price subsidy, suggesting behavior economics strategies may be more effective and less costly than financial strategies, and deserve more consideration in health promotion policies.

In contrast, financial incentives may not have intended effectiveness, especially among price-inelastic individuals. In our study, 62.9% of respondents are classified as risk averse, among which the test uptake rate was over 95% even without financial incentives. Additional subsidies in test cost did not significantly increase the uptake. On the other hand, 37.1% of respondents were cost-conscious, and a 75% subsidy on test cost would increase the uptake from 8.8% to 12.3%, whereas a 90% decrease in long-term gout treatment cost would improve uptake to 29.4%. Comparing the magnitude of effects, financial incentives were not as effective as physician's recommendation.

5.5.6 Comparing econometric models for DCE data

When comparing econometric models for choice analysis, different domains need to be considered. Statistical measures of model fit, such as likelihood ratio, pseudo R^2 are useful, but these values cannot be used to compare MXL and LCM models if the two models are not nested.[236] The estimates from different models are not directly comparable due to scale differences, but behavioral outputs such as WTP estimates and choice probability prediction (stated uptake rate) are useful indicators of model appropriateness, as some information or prior knowledge on behaviors are available.

All evidence suggests heterogeneity in preferences for genetic testing services. In analysis of non-demanders, 8.5% of respondents always preferred not to test. On the contrary, 51.85% always preferred to test. These groups were making different tradeoffs. MXL model shows the estimates had big standard deviation, suggestion large variations across individuals. A latent class analysis identified two classes with

distinct choice patterns and preferences weights. Based on the findings, among the three models used, those account for heterogeneity are more appropriate than CLM.

Mixed logit models allow parameters to be random and follow a continuous distribution, and provide information on the population average parameter. Latent class models accounts for heterogeneity by allowing parameters to follow discrete distributions. Both models provide insights on preferences heterogeneity. When predicting the uptake rate, LCM appeared to give more reasonable predictions than MXL. MXL predicts higher uptake rates (ranges from 75.3% to 97.7%) than LCM (ranges from 65.1% to 73.6%) for all scenarios. Analysis of non-demanders showed that 8.5% of respondents always preferred not to test in all DCE questions, regardless of attribute levels. Analyses of warm-up questions also revealed the proportion of respondent willing to consider genetic testing was between 50% to 60%, when each attribute was considered separately. Even though various preference-eliciting methods may yield different responses, LCM predictions were more consistent with other findings. The inflated prediction from MXL could arise for two reasons: 1) The coefficient estimates were population averages, and extreme preferences were not well accounted for; and 2) for those who dominated on a certain attribute, no tradeoffs were made between different attributes, which may confound the model estimates.

5.5.7 Strengths of study

The study has many strengths. First, we included context variables in the DCE, which is closer to the real clinical setting, and the predictive value of model is better. Secondly, a no test option was included to allow for opting out, which is more realistic. In addition, the features of no test alternative were displayed. The precise definition of no test minimized the prior individual beliefs or knowledge about the no test. Thirdly, we used mixed logit and latent class models in analyses, which allow the analysis of heterogeneity in preferences. Fourthly, simulation of policy effects was conducted, which provide forecasts on the possible policy impact. Fifthly, we used a budget

reminder to minimize hypothetical bias. Choosing a test does not require respondents to pay, and it is possible that decision may change when individuals need to make real payment. The fact that the DCE is hypothetical scenarios and choices can potentially lead to bias, such as over estimation of willingness to pay. To minimize hypothetical bias, a budget reminder was incorporated to remind respondents to think carefully about the opportunity of the test cost and the impact of test cost on their budget. Sixthly, various validity tests were included to test respondents' attention and understanding of DCE questions.

5.5.8 Limitations of study

The study has several limitations. Firstly, DCE as a stated preference method has intrinsic limitations that people's stated preferences may differ from actual behaviors. However, several studies evaluating the external validity of DCE has found consistency in DCE results and actual behavior.[247,248] The DCE questions have been made as realistic as possible in this study by including choice context variables and budget reminder. More research are required to further examine the external validity of DCE, in relation to actual behavior, and other stated preference methods. Secondly, due to practical considerations, the study was conducted among diabetes patients who were at higher risk of gout than the general public. Patients' preferences may change with time and disease experiences. However, as comprehensive background information provided, and the socio-demographic features were similar, we do not expect differences in preferences between diabetes and gout patients. Thirdly, willingness-to-pay for risk reduction may be related to the ability to pay, or wealth. However, wealth is difficult to measure. We used controlled for household income and housing type as proxies for wealth in the analysis to account for potential effect of ability to pay on preferences. Fourthly, respondents were recruited using a convenience sampling method, and may be subject to selection bias. For instance,

patients with no interest in ADRs or genetic testing, and those with low cognitive capacity may refuse to participate in the study. Nonetheless, we compared the sample characteristics with the population demographics, and did not detect significant differences that will threaten the generalizability of our study.

5.6 Conclusions

Using a discrete choice experiment, this study quantified patients' preferences for using pharmacogenetic testing to reduce severe ADRs. The study identified substantial heterogeneity across individuals. Most patients are risk averse, and had higher preference weights for level of risk reduction than for cost of test. This group of patients have higher willingness-to-pay for genetic testing. Other patients are more cost conscious, and considered cost of test and long-term treatment more important than the level of risk reduction. Overall, our results predicted the test uptake rate to be 65% in Singapore. The study also revealed the strong impact of doctor's recommendation and moderate effect of herd behavior in shaping individuals' test decisions. These information strategies can be effective and inexpensive health promotion tools.

Chapter 6 Conclusions and future directions

6.1 Conclusions

In this thesis, I introduced the challenges in the adoption of pharmacogenetic testing in clinical practice to reduce risk of life-threatening adverse drug reaction, and described two economic evaluation methods that can inform the decision making at health system level and individual level on whether genetic testing should be done.

In chapter 2, I presented a cost-effectiveness analysis of HLA-B*1502 genetic testing prior to carbamazepine treatment in epilepsy treatment. Results suggest that in a life time, testing is highly cost-effective for Singaporean Chinese, Malays, but not for the Indians. As there are several effective alternative anti-epileptic drugs, avoiding carbamazepine in HLA-B*1502 patients can reduce risk of SJS and associated mortality and morbidities, but will not worsen seizure control. In addition, our model implies that HLA-B*1502 testing is more likely to be cost-effective in populations with high HLA-B*1502 frequency and high incidence of CBZ/PHT-induced SJS/TEN, such as various southern eastern Asian countries. From a policy perspective, our results imply that HLA-B*1502 is a high value service in Singapore. Following this study, genetic testing for HLA-B*1502 prior to carbamazepine treatment has been recommended in Singapore, and testing services were made available in several tertiary hospitals. The reduction in SJS case reports has demonstrated the effectiveness of HLA-B*1502 testing.

In chapter 3, I described a cost-effectiveness analysis of HLA-B*5801 genetic testing prior to allopurinol treatment in chronic gout management. When evaluated over a life time, genetic testing and avoiding allopurinol in testing positive patients is not cost-effective. In fact, it reduced the total QALYs, while incurring higher cost. This is because test positive patients (18.5%) would have fewer alternative treatment options, and thus worse gout outcomes, while SJS/TEN would be avoided in only 1.5% of patients. This shows that genetic testing does not necessarily improve

QALYs at population level, though risk of SJS is lowered, especially when the choice of alternative drugs are limited. Instead, a strategy that combines genetic testing and safety monitoring may become cost-effective under certain circumstances, and achieves a balance between risk mitigation and gout management outcomes. From a policy perspective, mandating HLA-B*5801 testing is not desirable.

In chapter 4, I reviewed the factors determining patients' preferences for taking genetic test to reduce risk of life-threatening adverse drug reactions. Determinants of patients' testing decision include test features (including cost of test, risk of SJS, and cost of long-term gout treatment) and decision context information (including information on doctor's recommendation, and the most common choice made by others). The impact of these factors were systematically tested and quantified in a discrete choice experiment descried in Chapter 5. Though HLA-B*5801 genetic testing is not cost-effective as shown in chapter 3, my results from discrete choice experiment shows that a significant proportion of patients are willing to test to reduce the risk of SJS. This group of patients are less sensitive to test cost and treatment costs, and have high willingness-to-pay for risk reduction. On the contrary, the other patients are more cost-conscious and only willing to test when the test can significantly reduce risk of SJS, or when the treatment cost is low. Given the current available test, the predicted test uptake rate is 65.1% among Singaporean patients. The effect of choice context factors on patient's decision making was also explored. Doctor's recommendation is the single most effective factor in improving test uptake rate (by 8.5%). On the other hand, labelling test as the most common choice slightly increased the uptake rate of test, suggesting herd behaviour is not as strong as doctor's recommendation. From a service provision perspective, the study results suggest there is a strong demand for genetic testing. From a policy perspective, mandating testing is likely to induce a welfare loss among those who prefer not to test. From a health promotion perspective, the study identified doctor's

recommendation and information on other people's choice to be effective and low-cost strategies to encourage healthy behaviours.

In conclusion, the thesis has demonstrated how economic evaluations can inform the decision on genetic testing adoption at the health system level and individual patient level. Understanding the economic value of health services, and patient's preferences may improve the cost-effectiveness and efficiency of health service delivery.

6.2 Future directions

6.2.1 Cost-effectiveness analysis

Cost-effectiveness models are static, and analyze the most likely scenario at the time of study. Cost-effectiveness results will change when input parameters are altered, such as the discovery of a new drug, the changes in drug prices related to patent or demand factors. It is worthwhile to re-evaluate the cost-effectiveness results when the context significantly changes.

Genetic testing technology is rapidly evolving. In the near future, it may become possible to conveniently test thousands of genes or the whole genome at relatively low costs. If true, genes that can predict individual's drug responses may be tested all together at birth, or at the first time a drug with known genetic risk factor is prescribed. The genetic profiles could even be incorporated into individual's electronic medical records. These will completely change the marginal cost and marginal benefit of genetic testing. Future studies could explore the costs and effectiveness of a combined testing of all the genes known to associate with adverse drug reactions, or all drug-related genes.

6.2.2 Discrete choice experiments

Discrete choice experiments as a useful tool to quantify patients' preferences, has increasing been used in health care. However, the external validity of DCEs in health care are rarely tested. Often times, it is empirically unclear how well the predicted behaviors are consistent with real behaviors. In fact, for most DCEs, a market to test the external validity does not exist at the time of study. It is important to find opportunities to compare the DCE with actual choices. For instance, when HLA-B*5801 becomes available in Singapore in the future, individual patients choices and the overall test uptake rate could be compared with the predictions. Nevertheless, any discrepancies should be interpreted with caution, as the context of DCE may be different from real life in many ways.

Our study suggests providing information on doctor's recommendation and herd behavior may be as effective as or more effective than traditional policy intervention strategies, such as cost subsidies. Further studies should be conducted to systematically evaluate the impact of various forms and strengths of doctor's recommendations, and herd information in nudging patients' choices.

References

1. Collins FS. A Brief Primer on Genetic Testing. Paper presented at: World Economic Forum, National Human Genome Research Institute,(http://www. genome. gov/10506784)2003.
2. Gillis N, Innocenti F. Evidence required to demonstrate clinical utility of pharmacogenetic testing: the debate continues. *Clin. Pharmacol. Ther.* 2014;96(6):655-657.
3. Roses AD. Pharmacogenetics and the practice of medicine. *Nature.* 2000;405(6788):857-865.
4. FDA. Table of Pharmacogenomic Biomarkers in Drug Labeling. http://www.fda.gov/drugs/scienceresearch/researchareas/pharmacogenetics/ucm083378.htm.
5. Frueh FW, Amur S, Mummaneni P, et al. Pharmacogenomic biomarker information in drug labels approved by the United States food and drug administration: prevalence of related drug use. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy.* 2008;28(8):992-998.
6. Hamburg MA, Collins FS. The path to personalized medicine. *N. Engl. J. Med.* 2010;363(4):301-304.
7. Savonarola A, Palmirotta R, Guadagni F, Silvestris F. Pharmacogenetics and pharmacogenomics: role of mutational analysis in anti-cancer targeted therapy. *The pharmacogenomics journal.* 2012;12(4):277-286.
8. Harris L, Fritsche H, Mennel R, et al. American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *J. Clin. Oncol.* 2007;25(33):5287-5312.
9. Allegra CJ, Jessup JM, Somerfield MR, et al. American Society of Clinical Oncology provisional clinical opinion: testing for KRAS gene mutations in patients with metastatic colorectal carcinoma to predict response to anti–epidermal growth factor receptor monoclonal antibody therapy. *J. Clin. Oncol.* 2009;27(12):2091-2096.
10. Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. *JAMA.* 1998;279(15):1200-1205.
11. Pirmohamed M, James S, Meakin S, et al. *Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients.* Vol 3292004.
12. Need AC, Motulsky AG, Goldstein DB. Priorities and standards in pharmacogenetic research. *Nat. Genet.* 2005;37(7):671-681.
13. HSA. http://www.hsa.gov.sg/content/dam/HSA/HPRG/Safety_Alerts_Product_Recalls_Enforcement/Adverse_Drug_Reaction_News/2009/ADR_News_Mar2009_Vol.11_No.1.pdf. 2009; http://www.hsa.gov.sg/content/dam/HSA/HPRG/Safety_Alerts_Product_Recalls_Enforcement/Adverse_Drug_Reaction_News/2009/ADR_News_Mar2009_Vol.11_No.1.pdf.
14. Thong BY-H, Leong K-P, Tang C-Y, Chng H-H. Drug allergy in a general hospital: results of a novel prospective inpatient reporting system. *Annals of Allergy, Asthma & Immunology.* 2003;90(3):342-347.
15. Roujeau JC, Stern RS. Severe adverse cutaneous reactions to drugs. *N. Engl. J. Med.* 1994;331(19):1272-1285.
16. Roujeau JC, Chosidow O, Saiag P, Guillaume JC. Toxic epidermal necrolysis (Lyell syndrome). *J. Am. Acad. Dermatol.* 1990;23(6 Pt 1):1039-1058.

17. Roujeau J-C, Kelly JP, Naldi L, et al. Medication Use and the Risk of Stevens–Johnson Syndrome or Toxic Epidermal Necrolysis. *N. Engl. J. Med.* 1995;333(24):1600-1608.
18. Mockenhaupt M, Viboud C, Dunant A, et al. Stevens–Johnson syndrome and toxic epidermal necrolysis: assessment of medication risks with emphasis on recently marketed drugs. The EuroSCAR-study. *J. Invest. Dermatol.* 2008;128(1):35-44.
19. Halevy S, Ghislain P-D, Mockenhaupt M, et al. Allopurinol is the most common cause of Stevens-Johnson syndrome and toxic epidermal necrolysis in Europe and Israel. *J. Am. Acad. Dermatol.* 2008;58(1):25-32.
20. Devi K, George S, Criton S, Suja V, Sridevi P. Carbamazepine-The commonest cause of toxic epidermal necrolysis and Stevens-Johnson syndrome: A study of 7 years. *Indian Journal of Dermatology, Venereology, and Leprology.* 2005;71(5):325.
21. Chung W-H, Hung S-I, Hong H-S, et al. Medical genetics: A marker for Stevens-Johnson syndrome. *Nature.* 2004;428(6982):486-486.
22. Hung S-I, Chung W-H, Jee S-H, et al. Genetic susceptibility to carbamazepine-induced cutaneous adverse drug reactions. *Pharmacogenet. Genomics.* 2006;16(4):297-306.
23. Man CB, Kwan P, Baum L, et al. Association between HLA‐B* 1502 allele and antiepileptic drug‐induced cutaneous reactions in Han Chinese. *Epilepsia.* 2007;48(5):1015-1018.
24. Wang Q, Zhou J-q, Zhou L-m, et al. Association between HLA-B* 1502 allele and carbamazepine-induced severe cutaneous adverse reactions in Han people of southern China mainland. *Seizure.* 2011;20(6):446-448.
25. Wu X, Hu F, An D, et al. Association between carbamazepine-induced cutaneous adverse drug reactions and the HLA-B* 1502 allele among patients in central China. *Epilepsy Behav.* 2010;19(3):405-408.
26. Zhang Y, Wang J, Zhao L-M, et al. Strong association between HLA-B* 1502 and carbamazepine-induced Stevens-Johnson syndrome and toxic epidermal necrolysis in mainland Han Chinese patients. *Eur. J. Clin. Pharmacol.* 2011;67(9):885-887.
27. Locharernkul C, Loplumlert J, Limotai C, et al. Carbamazepine and phenytoin induced Stevens‐Johnson syndrome is associated with HLA‐B* 1502 allele in Thai population. *Epilepsia.* 2008;49(12):2087-2091.
28. Tassaneeyakul W, Tiamkao S, Jantararoungtong T, et al. Association between HLA‐B* 1502 and carbamazepine‐induced severe cutaneous adverse drug reactions in a Thai population. *Epilepsia.* 2010;51(5):926-930.
29. Chang CC, Too CL, Murad S, Hussein SH. Association of HLA‐B* 1502 allele with carbamazepine‐induced toxic epidermal necrolysis and Stevens‐Johnson syndrome in the multi‐ethnic Malaysian population. *Int. J. Dermatol.* 2011;50(2):221-224.
30. Then S-M, Rani ZZM, Raymond AA, Ratnaningrum S, Jamal R. Frequency of the HLA-B* 1502 allele contributing to carbamazepine-induced hypersensitivity reactions in a cohort of Malaysian epilepsy patients. 2011.
31. Toh D, Tan L, Aw D, et al. Building pharmacogenetics into a pharmacovigilance program in Singapore: using serious skin rash as a pilot study. *The pharmacogenomics journal.* 2014;14(4):316-321.
32. Chong KW, Chan DW, Cheung YB, et al. Association of carbamazepine-induced severe cutaneous drug reactions and HLA-B* 1502 allele status, and dose and treatment duration in paediatric neurology patients in Singapore. *Arch. Dis. Child.* 2013:archdischild-2013-304767.
33. Kim S-H, Lee KW, Song W-J, et al. Carbamazepine-induced severe cutaneous adverse reactions and HLA genotypes in Koreans. *Epilepsy Res.* 2011;97(1):190-197.

34. Mehta TY, Prajapati LM, Mittal B, et al. Association of HLA-B* 1502 allele and carbamazepine-induced Stevens-Johnson syndrome among Indians. *Indian Journal of Dermatology, Venereology, and Leprology.* 2009;75(6):579.
35. Kaniwa N, Saito Y, Aihara M, et al. HLA-B locus in Japanese patients with anti-epileptics and allopurinol-related Stevens-Johnson syndrome and toxic epidermal necrolysis. 2008.
36. Alfirevic A, Jorgensen AL, Williamson PR, Chadwick DW, Park BK, Pirmohamed M. HLA-B locus in Caucasian patients with carbamazepine hypersensitivity. 2006.
37. Lonjou C, Thomas L, Borot N, et al. A marker for Stevens-Johnson syndrome…: ethnicity matters. *The pharmacogenomics journal.* 2006;6(4):265-268.
38. Hung S-I, Chung W-H, Liou L-B, et al. HLA-B* 5801 allele as a genetic marker for severe cutaneous adverse reactions caused by allopurinol. *Proc. Natl. Acad. Sci. U. S. A.* 2005;102(11):4134-4139.
39. Chiu M, Hu M, Ng M, et al. Association between HLA‐B* 58: 01 allele and severe cutaneous adverse reactions with allopurinol in Han Chinese in Hong Kong. *Br. J. Dermatol.* 2012;167(1):44-49.
40. Cao Z-h, Wei Z-y, Zhu Q-y, et al. HLA-B* 58: 01 allele is associated with augmented risk for both mild and severe cutaneous adverse reactions induced by allopurinol in Han Chinese. *Pharmacogenomics.* 2012;13(10):1193-1201.
41. Kang H-R, Jee YK, Kim Y-S, et al. Positive and negative associations of HLA class I alleles with allopurinol-induced SCARs in Koreans. *Pharmacogenet. Genomics.* 2011;21(5):303-307.
42. Jung JW, Song WJ, Kim YS, et al. HLA-B58 can help the clinical decision on starting allopurinol in patients with chronic renal insufficiency. *Nephrol. Dial. Transplant.* 2011;26(11):3567-3572.
43. Tassaneeyakul W, Jantararoungtong T, Chen P, et al. Strong association between HLA-B*5801 and allopurinol-induced Stevens-Johnson syndrome and toxic epidermal necrolysis in a Thai population. *Pharmacogenetics and genomics.* 2009;19(9):704-709.
44. Tohkin M, Kaniwa N, Saito Y, et al. A whole-genome association study of major determinants for allopurinol-related Stevens–Johnson syndrome and toxic epidermal necrolysis in Japanese patients. *The pharmacogenomics journal.* 2013;13(1):60-69.
45. Jarjour S, Barrette M, Normand V, Rouleau JL, Dubé M-P, de Denus S. Genetic markers associated with cutaneous adverse drug reactions to allopurinol: a systematic review. 2015.
46. Tutton R. Pharmacogenomic biomarkers in drug labels: what do they tell us? *Pharmacogenomics.* 2014;15(3):297-304.
47. Sanderson S, Zimmern R, Kroese M, Higgins J, Patch C, Emery J. How can the evaluation of genetic tests be enhanced? Lessons learned from the ACCE framework and evaluating genetic tests in the United Kingdom. *Genet. Med.* 2005;7(7):495-500.
48. Grossman I. Routine pharmacogenetic testing in clinical practice: dream or reality? *Pharmacogenomics.* 2007;8(10):1449-1459.
49. Devlin N, Parkin D. Does NICE have a cost‐effectiveness threshold and what other factors influence its decisions? A binary choice analysis. *Health Econ.* 2004;13(5):437-452.
50. Hensher DA, Rose JM, Greene WH. *Applied choice analysis: a primer.* Cambridge University Press; 2005.
51. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making. *Pharmacoeconomics.* 2008;26(8):661-677.

52. Louviere JJ, Hensher DA, Swait JD. *Stated choice methods: analysis and applications.* Cambridge University Press; 2000.
53. Orme BK. Getting started with conjoint analysis. *Madison, WI: Research Publishers LLC.* 2006.
54. Hanley N, MacMillan D, Wright RE, et al. Contingent valuation versus choice experiments: estimating the benefits of environmentally sensitive areas in Scotland. *Journal of agricultural economics.* 1998;49(1):1-15.
55. Bridges J. Stated preference methods in health care evaluation: an emerging methodological paradigm in health economics. *Applied health economics and health policy.* 2003;2(4):213-224.
56. Bridges JFP, Hauber AB, Marshall D, et al. Conjoint analysis applications in health—a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. *Value in Health.* 2011;14(4):403-413.
57. Ryan M, Gerard K, Amaya-Amaya M. *Using discrete choice experiments to value health and health care.* Vol 11: Springer Science & Business Media; 2007.
58. de Bekker‐Grob EW, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. *Health Econ.* 2012;21(2):145-172.
59. Hanley N, Wright RE, Adamowicz V. Using choice experiments to value the environment. *Environmental and resource economics.* 1998;11(3-4):413-428.
60. Speight J, Barendse SM. FDA guidance on patient reported outcomes. *BMJ.* 2010;340.
61. WHO. Epilepsy Fact sheet N°999. 2001; http://www.who.int/mediacentre/factsheets/fs999/en/.
62. Mac TL, Tran DS, Quet F, Odermatt P, Preux PM, Tan CT. Epidemiology, aetiology, and clinical management of epilepsy in Asia: a systematic review. *Lancet Neurol.* 2007;6(6):533-543.
63. Heaney DC, Shorvon SD, Sander JW. An economic appraisal of carbamazepine, lamotrigine, phenytoin and valproate as initial treatment in adults with newly diagnosed epilepsy. *Epilepsia.* 1998;39 Suppl 3:S19-25.
64. Roujeau JC, Stern RS. Severe adverse cutaneous reactions to drugs. *N Engl J Med.* 1994;331(19):1272-1285.
65. Yap F, Wahiduzzaman M, Pubalan M. Stevens-Johnson syndrome (SJS) and Toxic epidermal necrolysis (TEN) in Sarawak: A four years' review. *Egyptian Dermatology Online Journal.* 2008;4(1):1-13.
66. Hung SI, Chung WH, Liu ZS, et al. Common risk allele in aromatic antiepileptic-drug induced Stevens-Johnson syndrome and toxic epidermal necrolysis in Han Chinese. *Pharmacogenomics.* 2010;11(3):349-356.
67. Chung WH, Hung SI, Hong HS, et al. Medical genetics: a marker for Stevens-Johnson syndrome. *Nature.* 2004;428(6982):486.
68. Lim KS, Kwan P, Tan CT. Association of HLA-B* 1502 allele and carbamazepine-induced severe adverse cutaneous drug reaction among Asians, a review. *Neurology Asia.* 2008;13(6):15-21.
69. FDA. Information for Healthcare Professionals: Dangerous or Even Fatal Skin Reactions - Carbamazepine (marketed as Carbatrol, Equetro, Tegretol, and generics) (accessible at www.fda.gov/Drugs/DrugSafety/PostmarketdrugsafetyinformationforPatientsandProviders/ucm124718.htm). . 2007.
70. Authority SHS. Serious adverse skin reactions associated with carbamazepine. *http://www.hsa.gov.sg/publish/hsaportal/en/health_products_regulation/safety_information/adr_bulletin.html.*2009 March (Volume 11 Number 1)

71. Chang CC, Too CL, Murad S, Hussein SH. Association of HLA-B*1502 allele with carbamazepine-induced toxic epidermal necrolysis and Stevens-Johnson syndrome in the multi-ethnic Malaysian population. *Int J Dermatol.* 2011;50(2):221-224.
72. Tudur Smith C, Marson AG, Clough HE, Williamson PR. Carbamazepine versus phenytoin monotherapy for epilepsy. *Cochrane Database Syst Rev.* 2002(2):CD001911.
73. Lim SH, Tan EK, Chen C. Pattern of anti-epileptic drug usage in a tertiary referral hospital in Singapore. *Neurol J Southeast Asia.* 1997;2:77-85.
74. Marson AG, Williamson PR, Hutton JL, Clough HE, Chadwick DW. Carbamazepine versus valproate monotherapy for epilepsy. *Cochrane Database Syst Rev.* 2000(3):CD001030.
75. Glauser T, Ben-Menachem E, Bourgeois B, et al. ILAE treatment guidelines: evidence-based analysis of antiepileptic drug efficacy and effectiveness as initial monotherapy for epileptic seizures and syndromes. *Epilepsia.* 2006;47(7):1094-1120.
76. Mattson RH, Cramer JA, Collins JF. A comparison of valproate with carbamazepine for the treatment of complex partial seizures and secondarily generalized tonic-clonic seizures in adults. The Department of Veterans Affairs Epilepsy Cooperative Study No. 264 Group. *N. Engl. J. Med.* 1992;327(11):765-771.
77. Marson AG, Williamson PR, Clough H, Hutton JL, Chadwick DW. Carbamazepine versus valproate monotherapy for epilepsy: a meta-analysis. *Epilepsia.* 2002;43(5):505-513.
78. Privitera MD, Brodie MJ, Mattson RH, Chadwick DW, Neto W, Wang S. Topiramate, carbamazepine and valproate monotherapy: double-blind comparison in newly diagnosed epilepsy. *Acta Neurol. Scand.* 2003;107(3):165-175.
79. Gaitatzis A, Johnson AL, Chadwick DW, Shorvon SD, Sander JW. Life expectancy in people with newly diagnosed epilepsy. *Brain.* 2004;127(Pt 11):2427-2432.
80. Selai C, Bannister D, Trimble M. Antiepileptic drugs and the regulation of mood and quality of life (QOL): the evidence from epilepsy. *Epilepsia.* 2005;46 Suppl 4:50-57.
81. Remak E, Hutton J, Price M, Peeters K, Adriaenssen I. A Markov model of treatment of newly diagnosed epilepsy in the UK. An initial assessment of cost-effectiveness of topiramate. *Eur J Health Econ.* 2003;4(4):271-278.
82. Sanchez JL, Pereperez SB, Bastida JL, Martinez MM. Cost-utility analysis applied to the treatment of burn patients in a specialized center. *Arch. Surg.* 2007;142(1):50-57; discussion 57.
83. Williams F, Meenagh A, Darke C, et al. Analysis of the distribution of HLA-B alleles in populations from five continents. *Hum. Immunol.* 2001;62(6):645-650.
84. Chen P, Lin JJ, Lu CS, et al. Carbamazepine-induced toxic effects and HLA-B*1502 screening in Taiwan. *N Engl J Med.* 2011;364(12):1126-1133.
85. Singapore Department of Statistics. Census of population 2010. Statistical release 1: Demographic Characteristics E, Language and Religion. [online]. Available at: www.singstat.gov.sg/pubn/popn/c2010sr1/cop2010sr1.pdf. Accessed June 18, 2011.
86. Grosse SD. Assessing cost-effectiveness in healthcare: history of the $50,000 per QALY threshold. *Expert Rev. Pharmacoecon. Outcomes Res.* 2008;8(2):165-178.
87. Gonzalez-Galarza FF, Christmas S, Middleton D, Jones AR. Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations. *Nucleic Acids Res.* 2011;39(Database issue):D913-919.
88. HSA. Serious adverse skin reactions associated with carbamazepine 2009.
89. Chen Z, Liew D, Kwan P. Effects of a HLA-B* 15: 02 screening policy on antiepileptic drug use and severe skin reactions. *Neurology.* 2014;83(22):2077-2084.
90. HSA. 29 Aug 2013: Recommendations for HLA-B*1502 genotype testing prior to initiation of carbamazepine in new patients. 2013;

http://www.hsa.gov.sg/content/hsa/en/Health_Products_Regulation/Safety_Information_and_Product_Recalls/Product_Safety_Alerts/2013/recommendations_for.html.
91. Health Sciences Authority Singapore. 31 Dec 2013: HLA-B*1502 genotype testing: Towards safer use of carbamazepine. 2013; http://www.hsa.gov.sg/content/hsa/en/Health_Products_Regulation/Safety_Information_and_Product_Recalls/Product_Safety_Alerts/2013/hla-b_1502_genotype.html.
92. Rho YH, Zhu Y, Choi HK. The epidemiology of uric acid and fructose. *Semin. Nephrol.* 2011;31(5):410-419.
93. McAdams DeMarco MA, Maynard JW, Baer AN, et al. Diuretic use, increased serum urate levels, and risk of incident gout in a population-based study of adults with hypertension: The Atherosclerosis Risk in Communities cohort study. *Arthritis Rheum.* 2012;64(1):121-129.
94. Khanna D, Fitzgerald JD, Khanna PP, et al. 2012 American College of Rheumatology guidelines for management of gout. Part 1: Systematic nonpharmacologic and pharmacologic therapeutic approaches to hyperuricemia. *Arthritis Care Res.* 2012;64(10):1431-1446.
95. Singh JA, Strand V. Gout is associated with more comorbidities, poorer health-related quality of life and higher healthcare utilisation in US veterans. *Ann. Rheum. Dis.* 2008;67(9):1310-1316.
96. Palla I, Fusco F, Tani C, Baldini C, Mosca M, Turchetti G. The economic impact of gout: a systematic literature review. *Clin. Exp. Rheumatol.* 2012;30(4 Suppl 73):S145.
97. Zhang W, Doherty M, Bardin T, et al. EULAR evidence based recommendations for gout. Part II: Management. Report of a task force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). *Ann. Rheum. Dis.* 2006;65(10):1312-1324.
98. Teng GG, Ang L-W, Saag KG, Yu MC, Yuan J-M, Koh W-P. Mortality due to coronary heart disease and kidney disease among middle-aged and elderly men and women with gout in the Singapore Chinese Health Study. *Ann. Rheum. Dis.* 2012;71(6):924-928.
99. Khanna D, Fitzgerald JD, Khanna PP, et al. 2012 American College of Rheumatology guidelines for management of gout. Part 1: Systematic nonpharmacologic and pharmacologic therapeutic approaches to hyperuricemia. *Arthritis Care Res.* 2012;64(10):1431-1446.
100. HSA. Serious skin reactions associated with allopurinol 2009.
101. Lin MS, Dai YS, Pwu RF, Chen YH, Chang NC. Risk estimates for drugs suspected of being associated with Stevens‐Johnson syndrome and toxic epidermal necrolysis: a case‐control study. *Intern. Med. J.* 2005;35(3):188-190.
102. Kaniwa N, Saito Y, Aihara M, et al. HLA-B locus in Japanese patients with anti-epileptics and allopurinol-related Stevens–Johnson syndrome and toxic epidermal necrolysis. *Pharmacogenomics.* 2008;9(11):1617-1622.
103. Halevy S, Ghislain PD, Mockenhaupt M, et al. Allopurinol is the most common cause of Stevens-Johnson syndrome and toxic epidermal necrolysis in Europe and Israel. *J. Am. Acad. Dermatol.* 2008;58(1):25-32.
104. Mahar PD, Wasiak J, Hii B, et al. A systematic review of the management and outcome of toxic epidermal necrolysis treated in burns centres. *Burns.* 2014.
105. Somkrua R, Eickman EE, Saokaew S, Lohitnavy M, Chaiyakunapruk N. Association of HLA-B*5801 allele and allopurinol-induced Stevens Johnson syndrome and toxic epidermal necrolysis: a systematic review and meta-analysis. *BMC medical genetics.* 2011;12:118.

106. Hung S-I, Chung W-H, Liou L-B, et al. HLA-B*5801 allele as a genetic marker for severe cutaneous adverse reactions caused by allopurinol. *Proc. Natl. Acad. Sci. U. S. A.* 2005;102(11):4134-4139.
107. Zineh I, Mummaneni P, Lyndly J, et al. Allopurinol pharmacogenetics: assessment of potential clinical usefulness. *Pharmacogenomics.* 2011;12(12):1741-1749.
108. European Medicine Agency CfMPfHU, Pharmacovigilance Working Party. Allopurinol - Risk of skin reactions assoicated with HLA-B*5801 allele. *PhVWP monthly report on safety concerns, guidelines and general matters.* 2012;July 2012(1207).
109. Taiwan Food and Drug Administration. Drug Safety News 2009.08.28. Available from: http://www.fda.gov.tw/tc/siteListContent.aspx?sid=1571&id=7616&chk=1197ef3b-eda3-4a7b-95e0-7c6e7013d32e&param=pn%3D69%26sid%3D1571#.VGtc8vmUd9g. 2009. Accessed 4 March 2015.
110. dbMHC database. Available from: http://www.ncbi.nlm.nih.gov/projects/gv/mhc/ihwg.cgi. Accessed 4 March 2015.
111. Garcia-Doval I, LeCleach L, Bocquet H, Otero X, Roujeau J. Toxic epidermal necrolysis and stevens-johnson syndrome: Does early withdrawal of causative drugs decrease the risk of death? *Arch. Dermatol.* 2000;136(3):323-327.
112. Firestein GS, Kelley W, Ruddy S, Harris E, Sledge C. Textbook of rheumatology. *Textbook of rheumatology.* 1997.
113. Statistics SDo. Life expectancy at birth. . http://www.singstat.gov.sg/statistics/visualising_data/chart/Life_Expectancy_At_Birth.html
114. Jordan KM, Cameron JS, Snaith M, et al. British Society for Rheumatology and British Health Professionals in Rheumatology guideline for the management of gout. *Rheumatology (Oxford).* 2007;46(8):1372-1374.
115. Sarawate CA, Brewer KK, Yang W, et al. Gout Medication Treatment Patterns and Adherence to Standards of Care From a Managed Care Perspective. *Mayo Clin. Proc.* 2006;81(7):925-934.
116. Reinders MK, Haagsma C, Jansen TLTA, et al. A randomised controlled trial on the efficacy and tolerability with dose escalation of allopurinol 300–600 mg/day versus benzbromarone 100–200 mg/day in patients with gout. *Ann. Rheum. Dis.* 2009;68(6):892-897.
117. Terkeltaub R. Update on gout: new therapeutic strategies and options. *Nature Reviews Rheumatology.* 2010;6(1):30-38.
118. Reinders MK, van Roon EN, Jansen TLTA, et al. Efficacy and tolerability of urate-lowering drugs in gout: a randomised controlled trial of benzbromarone versus probenecid after failure of allopurinol. *Ann. Rheum. Dis.* 2009;68(1):51-56.
119. Ramasamy SN, Korb-Wells CS, Kannangara DR, et al. Allopurinol Hypersensitivity: A Systematic Review of All Published Cases, 1950–2012. *Drug Saf.* 2013;36(10):953-980.
120. Stamp LK, Taylor WJ, Jones PB, et al. Starting dose is a risk factor for allopurinol hypersensitivity syndrome: A proposed safe starting dose of allopurinol. *Arthritis Rheum.* 2012;64(8):2529-2536.
121. Yip L, Thong B, Lim J, et al. Ocular manifestations and complications of Stevens–Johnson syndrome and toxic epidermal necrolysis: an Asian series. *Allergy.* 2007;62(5):527-531.
122. Gueudry J, Roujeau J-C, Binaghi M, Soubrane G, Muraine M. Risk factors for the development of ocular complications of Stevens-Johnson syndrome and toxic epidermal necrolysis. *Arch. Dermatol.* 2009;145(2):157-162.

123. Dong D, Sung C, Finkelstein EA. Cost-effectiveness of HLA-B* 1502 genotyping in adult patients with newly diagnosed epilepsy in Singapore. *Neurology.* 2012;79(12):1259-1267.
124. Waduthantri S, Yong SS, Tan CH, et al. Cost of Dry Eye Treatment in an Asian Clinic Setting. *PLoS One.* 2012;7(6):e37711.
125. Beard S, Scheele B, Nuki G, Pearson I. Cost-effectiveness of febuxostat in chronic gout. *Eur J Health Econ.* 2013:1-11.
126. Schiffman RM, Walt JG, Jacobsen G, Doyle JJ, Lebovics G, Sumner W. Utility assessment among patients with dry eye disease. *Ophthalmology.* 2003;110(7):1412-1419.
127. Schumacher HR, Becker MA, Wortmann RL, et al. Effects of febuxostat versus allopurinol and placebo in reducing serum urate in subjects with hyperuricemia and gout: A 28‐week, phase III, randomized, double‐blind, parallel‐group trial. *Arthritis Care Res.* 2008;59(11):1540-1548.
128. Becker MA, Schumacher Jr HR, Wortmann RL, et al. Febuxostat compared with allopurinol in patients with hyperuricemia and gout. *N. Engl. J. Med.* 2005;353(23):2450-2461.
129. Reinders M, Haagsma C, Jansen TTA, et al. A randomised controlled trial on the efficacy and tolerability with dose escalation of allopurinol 300–600 mg/day versus benzbromarone 100–200 mg/day in patients with gout. *Ann. Rheum. Dis.* 2009;68(6):892-897.
130. Reinders M, Van Roon E, Jansen TTA, et al. Efficacy and tolerability of urate-lowering drugs in gout: a randomised controlled trial of benzbromarone versus probenecid after failure of allopurinol. *Ann. Rheum. Dis.* 2009;68(1):51-56.
131. Pui K, Gow PJ, Dalbeth N. Efficacy and tolerability of probenecid as urate-lowering therapy in gout; clinical experience in high-prevalence population. *The Journal of rheumatology.* 2013;40(6):872-876.
132. Reinders MK, van Roon EN, Jansen TL, et al. Efficacy and tolerability of urate-lowering drugs in gout: a randomised controlled trial of benzbromarone versus probenecid after failure of allopurinol. *Annals of the rheumatic diseases.* 2009;68(1):51-56.
133. Pillai NE, Okada Y, Saw W-Y, et al. Predicting HLA alleles from high-resolution SNP data in three Southeast Asian populations. *Hum. Mol. Genet.* 2014:ddu149.
134. Tassaneeyakul W, Jantararoungtong T, Chen P, et al. Strong association between HLA-B* 5801 and allopurinol-induced Stevens-Johnson syndrome and toxic epidermal necrolysis in a Thai population. *Pharmacogenet. Genomics.* 2009;19(9):704-709.
135. Thong BY. Stevens-Johnson syndrome / toxic epidermal necrolysis: an Asia-Pacific perspective. *Asia Pacific allergy.* 2013;3(4):215-223.
136. Robinson PC, Merriman TR, Herbison P, Highton J. Hospital admissions associated with gout and their comorbidities in New Zealand and England 1999–2009. *Rheumatology (Oxford).* 2013;52(1):118-126.
137. Monetary Authority of Singapore (2014) Foreign exchange rates. Available at: https://secure.mas.gov.sg/msb/ExchangeRates.aspx. Accessed May 6, 2014.
138. McGhan WF, Al M, Doshi JA, Kamae I, Marx SE, Rindress D. The ISPOR Good Practices for Quality Improvement of Cost-Effectiveness Research Task Force Report. *Value Health.* 2009;12(8):1086-1099.
139. Administration USFaD. Safety Labeling Changes Approved by FDA Centre for Drug Evaluation and Research (CDER). *http://www.fda.gov/Safety/MedWatch/SafetyInformation/ucm243770.htm.* November 2012.

140. Saokaew S, Tassaneeyakul W, Maenthaisong R, Chaiyakunapruk N. Cost-Effectiveness Analysis of HLA-B* 5801 Testing in Preventing Allopurinol-Induced SJS/TEN in Thai Population. *PLoS One.* 2014;9(4):e94294.
141. Ramasamy SN, Korb-Wells CS, Kannangara DR, et al. Allopurinol hypersensitivity: a systematic review of all published cases, 1950-2012. *Drug Saf.* 2013;36(10):953-980.
142. Singer JZ, Wallace SL. The allopurinol hypersensitivity syndrome. Unnecessary morbidity and mortality. *Arthritis Rheum.* 1986;29(1):82-87.
143. Kuo CF, Grainge MJ, Mallen C, Zhang W, Doherty M. Rising burden of gout in the UK but continuing suboptimal management: a nationwide population study. *Ann. Rheum. Dis.* 2014.
144. Roddy E, Zhang W, Doherty M. Concordance of the management of chronic gout in a UK primary-care population with the EULAR gout recommendations. *Ann. Rheum. Dis.* 2007;66(10):1311-1315.
145. Doherty M, Jansen TL, Nuki G, et al. Gout: why is this curable disease so seldom cured? *Ann. Rheum. Dis.* 2012;71(11):1765-1770.
146. Von Neumann J, Morgenstern O. *Theory of games and economic behavior.* Princeton university press; 2007.
147. Kahneman D, Tversky A. Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society.* 1979:263-291.
148. Tversky A, Kahneman D. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty.* 1992;5(4):297-323.
149. Foster MW, Mulvihill JJ, Sharp RR. Evaluating the utility of personal genomic information. *Genet. Med.* 2009;11(8):570-574.
150. Bennette CS, Trinidad SB, Fullerton SM, et al. Return of incidental findings in genomic medicine: measuring what patients value[mdash]development of an instrument to measure preferences for information from next-generation testing (IMPRINT). *Genet. Med.* 2013;15(11):873-881.
151. Liang S-Y, Phillips KA, Nagamine M, Ladabaum U, Haas JS. Rates and Predictors of Colorectal Cancer Screening. *Prev. Chronic Dis.* 2006;3(4):A117.
152. Smith RA, Cokkinides V, Brawley OW. Cancer screening in the United States, 2009: A review of current American Cancer Society guidelines and issues in cancer screening. *CA Cancer J. Clin.* 2009;59(1):27-41.
153. Ho MP, Gonzalez JM, Lerner HP, et al. Incorporating patient-preference evidence into regulatory decision making. *Surg. Endosc.* 2015:1-10.
154. FDA. Patient Preference Information – Submission, Review in PMAs, HDE Applications, and De Novo Requests, and Inclusion in Device Labeling.Draft Guidance for Industry, Food and Drug Administration Staff, and Other Stakeholders. 2015.
155. Rogausch A, Prause D, Schallenberg A, Brockmöller J, Himmel W. Patients' and physicians' perspectives on pharmacogenetic testing. 2006.
156. Haga SB, Tindall G, O'Daniel JM. Public perspectives about pharmacogenetic testing and managing ancillary findings. *Genetic testing and molecular biomarkers.* 2012;16(3):193-197.
157. Haga SB, O'Daniel JM, Tindall GM, Lipkus IR, Agans R. Survey of US public attitudes toward pharmacogenetic testing. *The pharmacogenomics journal.* 2012;12(3):197-204.
158. Payne K, Fargher EA, Roberts SA, et al. Valuing pharmacogenetic testing services: A comparison of patients' and health care professionals' preferences. *Value Health.* 2011;14(1):121-134.
159. Herbild L, Bech M, Gyrd‐Hansen D. Estimating the Danish populations' preferences for pharmacogenetic testing using a discrete choice experiment. The case of treating depression. *Value Health.* 2009;12(4):560-567.

160. Hochbaum GM. *Public participation in medical screening programs: a socio-psychological study.* US Department of Health, Education, and Welfare, Public Health Service, Bureau of State Services, Division of Special Health Services, Tuberculosis Program; 1958.
161. Rosenstock IM. Historical origins of the health belief model. *Health Education & Behavior.* 1974;2(4):328-335.
162. Andersen R. A behavioral model of families' use of health services. *Research Ser.* 1968(25).
163. Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? *J. Health Soc. Behav.* 1995:1-10.
164. Ajzen I. The theory of planned behavior. *Organ. Behav. Hum. Decis. Process.* 1991;50(2):179-211.
165. Gerend MA, Shepherd JE. Predicting human papillomavirus vaccine uptake in young adult women: Comparing the health belief model and theory of planned behavior. *Ann. Behav. Med.* 2012;44(2):171-180.
166. Hall J, Fiebig DG, King MT, Hossain I, Louviere JJ. What influences participation in genetic carrier testing?: Results from a discrete choice experiment. *J. Health Econ.* 2006;25(3):520-537.
167. Rattanavipapong W, Koopitakkajorn T, Praditsitthikorn N, Mahasirimongkol S, Teerawattananon Y. Economic evaluation of HLA‐B* 15: 02 screening for carbamazepine‐induced severe adverse drug reactions in Thailand. *Epilepsia.* 2013;54(9):1628-1638.
168. Camerer CF, Kunreuther H. Decision processes for low probability events: Policy implications. *J. Policy Anal. Manage.* 1989;8(4):565-592.
169. Reyna VF, Brainerd CJ. Fuzzy‐trace theory and framing effects in choice: Gist extraction, truncation, and conversion. *Journal of Behavioral Decision Making.* 1991;4(4):249-262.
170. Menon G. HEALTH RISK PERCEPTIONS AND CONSUMER PSYCHOLOGY.
171. Knight SJ, Mohamed AF, Marshall DA, Ladabaum U, Phillips KA, Walsh JME. Value of Genetic Testing for Hereditary Colorectal Cancer in a Probability-Based US Online Sample. *Med. Decis. Making.* 2015.
172. Marshall DA, Johnson FR, Phillips KA, Marshall JK, Thabane L, Kulin NA. Measuring Patient Preferences for Colorectal Cancer Screening Using a Choice-Format Survey. *Value Health.* 2007;10(5):415-430.
173. Edwards A, Elwyn G. Understanding risk and lessons for clinical risk communication about treatment preferences. *Qual. Health Care.* 2001;10(suppl 1):i9-i13.
174. Berwick DM, Weinstein MC. What do patients value? Willingness to pay for ultrasound in normal pregnancy. *Med. Care.* 1985:881-893.
175. Grosse SD, Kalman L, Khoury MJ. Evaluation of the validity and utility of genetic testing for rare diseases. *Rare Diseases Epidemiology*: Springer; 2010:115-131.
176. Özdemir S, Johnson FR, Hauber AB. Hypothetical bias, cheap talk, and stated willingness to pay for health care. *J. Health Econ.* 2009;28(4):894-901.
177. Pediatrics AAo. Ethical issues with genetic testing in pediatrics. 2001.
178. Brierley KL, Blouch E, Cogswell W, et al. Adverse events in cancer genetic testing: medical, ethical, legal, and financial implications. *The Cancer Journal.* 2012;18(4):303-309.
179. Christenhusz GM, Devriendt K, Dierickx K. To tell or not to tell? A systematic review of ethical reflections on incidental findings arising in genetics contexts. *Eur. J. Hum. Genet.* 2013;21(3):248-255.

180. Matloff ET, Shappell H, Brierley K, Bernhardt BA, McKinnon W, Peshkin BN. What would you do? Specialists' perspectives on cancer genetic testing, prophylactic surgery, and insurance discrimination. *J. Clin. Oncol.* 2000;18(12):2484-2492.
181. Gostin L. Genetic discrimination: the use of genetically based diagnostic and prognostic tests by employers and insurers. *Am. JL & Med.* 1991;17:109.
182. Rosenthal SL, Weiss TW, Zimet GD, Ma L, Good MB, Vichnin MD. Predictors of HPV vaccine uptake among women aged 19–26: Importance of a physician's recommendation. *Vaccine.* 2011;29(5):890-895.
183. Wydeven KV, Kwan A, Hardan AY, Bernstein JA. Underutilization of genetics services for autism: the importance of parental awareness and provider recommendation. *Journal of genetic counseling.* 2012;21(6):803-813.
184. Johar M, Fiebig DG, Haas M, Viney R. Using repeated choice experiments to evaluate the impact of policy changes on cervical screening. *Applied Economics.* 2012;45(14):1845-1855.
185. King MT, Hall J, Lancsar E, et al. Patient preferences for managing asthma: results from a discrete choice experiment. *Health Econ.* 2007;16(7):703-717.
186. Hall J, Kenny P, King M, Louviere J, Viney R, Yeoh A. Using stated preference discrete choice modelling to evaluate the introduction of varicella vaccination. *Health Econ.* 2002;11(5):457-465.
187. Fiebig DG, Haas M, Hossain I, Street DJ, Viney R. Decisions about Pap tests: what influences women and providers? *Soc. Sci. Med.* 2009;68(10):1766-1774.
188. Gurmankin AD, Baron J, Hershey JC, Ubel PA. The Role of Physicians' Recommendations in Medical Treatment Decisions. *Med. Decis. Making.* 2002;22(3):262-271.
189. Levinson W, Kao A, Kuby A, Thisted RA. Not all patients want to participate in decision making. *J. Gen. Intern. Med.* 2005;20(6):531-535.
190. Fargher EA, Eddy C, Newman W, et al. Patients' and healthcare professionals' views on pharmacogenetic testing and its future delivery in the NHS. *Pharmacogenomics.* 2007;8(11):1511-1519.
191. Selkirk CG, Weissman SM, Anderson A, Hulick PJ. Physicians' preparedness for integration of genomic and pharmacogenetic testing into practice within a major healthcare system. *Genetic testing and molecular biomarkers.* 2013;17(3):219-225.
192. Haga SB, Burke W, Ginsburg GS, Mills R, Agans R. Primary care physicians' knowledge of and experience with pharmacogenetic testing. *Clin. Genet.* 2012;82(4):388-394.
193. Gigerenzer G, Gaissmaier W. Heuristic decision making. *Annu. Rev. Psychol.* 2011;62:451-482.
194. Tversky A, Kahneman D. Judgment under uncertainty: Heuristics and biases. *Science.* 1974;185(4157):1124-1131.
195. Banerjee AV. A simple model of herd behavior. *The Quarterly Journal of Economics.* 1992:797-817.
196. Asch SE. Effects of group pressure upon the modification and distortion of judgments. *Groups, Leadership, and Men. S.* 1951:222-236.
197. Bikhchandani S, Sharma S. Herd behavior in financial markets. *IMF Staff papers.* 2000:279-310.
198. Chen Y-F. Herd behavior in purchasing books online. *Comput. Human Behav.* 2008;24(5):1977-1992.
199. Rook L. An economic psychological approach to herd behavior. *Journal of Economic Issues.* 2006:75-95.
200. Watkins SC. From local to national communities: The transformation of demographic regimes in Western Europe, 1870-1960. *Population and Development Review.* 1990:241-272.

201. Spence D. Herd mentality. *BMJ.* 2011;343.
202. Fred HL. Elephant medicine revisited. *Tex. Heart Inst. J.* 2008;35(4):385.
203. Brody JR, Kern SE. Stagnation and herd mentality in the biomedical sciences. *Cancer Biol. Ther.* 2004;3(9):903-910.
204. Yang M, Lien H-M, Chou S-Y. IS THERE A PHYSICIAN PEER EFFECT? EVIDENCE FROM NEW DRUG PRESCRIPTIONS. *Econ. Inq.* 2014;52(1):116-137.
205. Carlsson F, García J, Löfgren Å. Conformity and the Demand for Environmental Goods. *Environ Resource Econ.* 2010;47(3):407-421.
206. Glanz K, Rimer BK, Viswanath K. *Health behavior and health education: theory, research, and practice.* John Wiley & Sons; 2008.
207. Bridges JF, Hauber AB, Marshall D, et al. Conjoint analysis applications in health—a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. *Value Health.* 2011;14(4):403-413.
208. Reed Johnson F, Lancsar E, Marshall D, et al. Constructing Experimental Designs for Discrete-Choice Experiments: Report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. *Value Health.* 2013;16(1):3-13.
209. Paling J. Strategies to help patients understand risks. *BMJ: British Medical Journal.* 2003;327(7417):745.
210. Caponecchia C. *Strategies for the effective communication of probabilities.* The University of New South Wales;2007.
211. Lipkus IM. ONLINE COLOR_Numeric, Verbal, and Visual Formats of Conveying Health Risks: Suggested Best Practices and Future Recommendations. *Med. Decis. Making.* 2007.
212. Ryan M, Skåtun D. Modelling non-demanders in choice experiments. *Health Econ.* 2004;13(4):397-402.
213. Ryan M, Skåtun D. Modelling non‐demanders in choice experiments. *Health Econ.* 2004;13(4):397-402.
214. Bekker‐Grob D, Esther W, Hol L, et al. Labeled versus unlabeled discrete choice experiments in health economics: an application to colorectal cancer screening. *Value Health.* 2010;13(2):315-323.
215. Lancsar E, Louviere J, Donaldson C, Currie G, Burgess L. Best worst discrete choice experiments in health: Methods and an application. *Soc. Sci. Med.* 2013;76(0):74-82.
216. Sandor Z, Wedel M. Designing conjoint choice experiments using managers' prior beliefs. *Journal of Marketing Research.* 2001;38(4):430-444.
217. Shafir E. Choosing versus rejecting: Why some options are both better and worse than others. *Mem. Cognit.* 1993;21(4):546-556.
218. Bech M, Kjaer T, Lauridsen J. Does the number of choice sets matter? Results from a web survey applying a discrete choice experiment. *Health Econ.* 2011;20(3):273-286.
219. Kuhfeld WF. Marketing research methods in SAS. *Experimental Design, Choice, Conjoint, and Graphical Techniques. Cary, NC, SAS-Institute TS-722.* 2005.
220. Kuhfeld WF. Orthogonal Arrays. http://support.sas.com/techsup/technote/ts723.html.
221. Sloane NJA. A Library of Orthogonal Arrays. http://neilsloane.com/oadir/.
222. Huber J. Conjoint analysis: how we got here and where we are (An Update). Paper presented at: Sawtooth Software Conference2005.
223. Street DJ, Burgess L, Louviere JJ. Quick and easy choice sets: Constructing optimal and nearly optimal stated choice experiments. *International Journal of Research in Marketing.* 2005;22(4):459-470.
224. Rose JM, Bliemer MC, Hensher DA, Collins AT. Designing efficient stated choice experiments in the presence of reference alternatives. *Transportation Research Part B: Methodological.* 2008;42(4):395-406.

225. Louviere JJ, Islam T, Wasi N, Street D, Burgess L. Designing discrete choice experiments: Do optimal designs come at a price? *Journal of Consumer Research.* 2008;35(2):360-375.
226. Carlsson F, García JH, Löfgren Å. Conformity and the demand for environmental goods. *Environ Resource Econ.* 2010;47(3):407-421.
227. Puig JG, Martínez MA. Hyperuricemia, gout and the metabolic syndrome. *Curr. Opin. Rheumatol.* 2008;20(2):187-191.
228. Doherty M. New insights into the epidemiology of gout. *Rheumatology (Oxford).* 2009;48(suppl 2):ii2-ii8.
229. Mikuls TR, Farrar JT, Bilker WB, Fernandes S, Schumacher HR, Saag KG. Gout epidemiology: results from the UK General Practice Research Database, 1990–1999. *Ann. Rheum. Dis.* 2005;64(2):267-272.
230. Suppiah R, Dissanayake A, Dalbeth N. High prevalence of gout in patients with Type 2 diabetes: male sex, renal impairment, and diuretic use are major risk factors. *The New Zealand medical journal.* 2008;121(1283):43-50.
231. Lancaster K. Operationally relevant characteristics in the theory of consumer behavior. *Essays in honor of Lord Robbins. Weidenfels and Nicholson, London.* 1972.
232. Scott A. Identifying and analysing dominant preferences in discrete choice experiments: an application in health care. *J. Econ. Psychol.* 2002;23(3):383-398.
233. McFadden D. Econometric models for probabilistic choice among products. *Journal of Business.* 1980:S13-S29.
234. McFadden D. *Econometric models of probabilistic choice.* Boston: Mit Press Cambridge, MA; 1981.
235. Shen J. Latent class model or mixed logit model? A comparison by transport mode choice data. *Applied Economics.* 2009;41(22):2915-2924.
236. Greene WH, Hensher DA. A latent class model for discrete choice analysis: contrasts with mixed logit. *Transportation Research Part B: Methodological.* 2003;37(8):681-698.
237. McFadden D. Conditional logit analysis of qualitative choice behavior. 1973.
238. Hole AR. Estimating mixed logit models using maximum simulated likelihood. *The Stata Journal.* 2007;7(3):388-401.
239. Train KE. *Discrete choice methods with simulation.* Cambridge university press; 2009.
240. Revelt D, Train K. Mixed logit with repeated choices: households' choices of appliance efficiency level. *Rev. Econ. Stat.* 1998;80(4):647-657.
241. Layton DF, Brown G. Heterogeneous preferences regarding global climate change. *Rev. Econ. Stat.* 2000;82(4):616-624.
242. Train K. Discrete choice methods with simulation. 2003. *Cambridge University Press.*24:25-26.
243. Train KE. EM algorithms for nonparametric estimation of mixing distributions. *Journal of Choice Modelling.* 2008;1(1):40-69.
244. Bhat CR. An endogenous segmentation mode choice model with an application to intercity travel. *Transportation science.* 1997;31(1):34-48.
245. Zikmund-Fisher BJ, Windschitl PD, Exe N, Ubel PA. 'I'll do what they did": Social norm information and cancer treatment decisions. *Patient Educ. Couns.* 2011;85(2):225-229.
246. Schwenk CR. The cognitive perspective on strategic decision making. *Journal of management studies.* 1988;25(1):41-55.
247. Telser H, Zweifel P. Validity of discrete-choice experiments evidence for health risk reduction. *Applied Economics.* 2007;39(1):69-78.
248. Mark TL, Swait J. Using stated preference and revealed preference modeling to evaluate prescribing decisions. *Health Econ.* 2004;13(6):563-573.

Appendix A. In-depth interview guide

In-depth Interview Guide

Key questions to be asked during the interview
Understand the awareness

Section	Topic	Time (minutes)
1	Introduction	10
2	Preferences for genetic testing	15
3	Role of doctor's recommendation	10
4	Role of majority choice	10
5	Service delivery	5
6	Interview experience and suggestions	5
Total time		55

Part 1: Introduction

Duration	Discussion
5 mins	**Approach potential participant** **Self introduction** (Hi, My name is Dong Di, I am a PhD student at the National University of Singapore, Graduate Medical School [show student card]. As part of my PhD research, I am conducting a survey to understand Singaporean's preferences for genetic testing; hoping that doctors and policy makers better understand the preferences of the general public, and hope future medical services and policies can be consistent with people's preferences. I wonder whether you can spare 10-20 minutes time to answer some questions and share your view on genetic testing with me? We will have 20 dollar NTUC voucher as token of appreciation) [**eg of questions**: Whether people know it, do they think it beneficial, why and why not take it? how much people are willing to pay?] [**If participant do not know genetic testing**: It's ok, actually a lot of people do not know it, some heard of it but do not know the details, I will introduce the details] [**If participant no time for today**: Do you think another time works for you? If so, I can schedule an appointment in my school]
1 mins	**Ask for permission to record (**Do you think it ok if I audio record our conversation, so I won't miss the points we discussed. Information is confidential, I will not release to someone outside our study team. There is no right or wrong answers, be open minded.) *If environment crowded and noisy, do not record, just take note. **Find a place if necessary**
	Moderator to greet and welcome participant (Thanks for agreeing to do this interview)
2 mins	**Introduce genetic testing** (Before we get start, I'd like to ask you whether you have heard of/done genetic testing. Basically, it's done like this. (Draw some blood, and analyze your genes in the laboratory)

	Introduce the type of genetic testing in this study (There are many uses of genetic testing, for instance prenatal genetic screen for genetic abnormalities, cancer marker screening, genetic screening to optimize drug dosing. The one we will **focus** on today is to use genetic testing to prevent life-threatening adverse drug reactions.
10 mins	Gout (痛风) is a common chronic rheumatic disease in Singapore that affects 4% of elderly people in Singapore. Patients have severe pain on their hot and swollen toes and joints. It's painful that patients can't stand. **ASK** Do you know someone who got this? Patients with recurrent gout are usually treated with a medicine called allopurinol, which they take every day. However, a small number of people will have life-threatening drug allergy ith allopurinol. **Ask** Do you know drug allergy? Stevens Johnson Syndrome (SJS) is a life-threatening drug allergy. Patients with SJS will have severe rash all over the body, the skin may detach. Eyes and mucosa may get severe inflammation. It's extremely painful, comparable to a severe burn. It costs $5000 to $20000 to treat, and 20% patients may die. Genetic test (HLA-B*5801) can identify those people with high risk for allopurinol allergy, and doctor can prescribe another medicine without risk, but more expensive. **Ask** What do you think of this genetic test?

Part 2: Preferences for genetic testing

Duration	Discussion
15mins	**<u>Topic: How do you decide whether or not to do a genetic test before initiating allopurinol if you need allopurinol?</u>**
	Hypothetical scenario: Suppose you are in the clinic where your doctor says you need to take allopurinol to manage your chronic gout. The doctor tells you that this drug is generally well tolerated, and effective. However there is a small chance of life-threatening adverse drug reaction. There is a genetic test that can tell you whether you are at risk of this adverse reaction, though the test is not 100% accurate. You need to pay some amount for the test, and you may need to take more expensive medicines if the test says you are at risk. Now you need to make a decision on whether or not you want to do the test.
	1. What do you think of the genetic test? Worth doing? Why?
	2. What factors will you consider when you make this decision? Probe question: Is cost/likelihood of adverse reaction/accuracy/drug cost/how your sample is collected and handled important?

	3. What other information would you like to know besides the information given above?
	4. What do you think is the most important factor?
	5. How much are you willing to pay for such a test (without considering whether it's feasible)?
	6. What's the highest amount you can accept?
	7. What's the maximum risk level that you can tolerate?
	8. Will changes in test features increase your likelihood of taking the test?

Part 3: Role of doctor's recommendation

Duration	Discussion
10mins	**Topic: How is doctor recommendation influencing your decision?**
	9. Is doctor's recommendation important for testing decision?
	10. Suppose a doctor gives you information on the test features, and his recommendation, are you going to follow his recommendation straight away or consider the test features and the recommendation at the same time?
	11. If a doctor's recommendation is different from your judgments, what are you going to do?
	12. Why or why not do you follow the doctor's recommendation?
	13. Does a specialist or a GP matter? Does public or private hospital matter?
	14. What if a doctor recommend against a test?

Part 4: Role of majority choice

Duration	Discussion
10mins	**Topic: How is majority choice influencing your decision?**
	1. Is what other people do important for your testing decision?
	2. Suppose you receive some information that 70% of people in your situation choose to do the test, are you going to follow them straight away or balance the test features and other people's choice?
	3. If the majority choice is different from your judgments, what are you going to do?
	4. Why or why not do you follow the majority choice?
	5. What if the doctor recommendation is against the majority choice?
	6. Now, considering all test features, doctor recommendation and information on what other people do, what is the most important factor in your decision?

Part 5: Service delivery

Duration	Discussion

5mins	**Topic: How do you prefer genetic testing services to be offered in Singapore?**
	1. Do you prefer the above genetic test for allopurinol to be available in Singapore?
	2. Do you want it to be offered in hospitals or clinics?
	3. What's the most acceptable way of collecting your sample?
	4. Do you want the test results to be included in your medical record so that other doctors can see it in the future?
	5. What do you think of genotyping for many diseases related genes at the same time and include in your medical record?
	6. Are you concerned about genetic test?

Part 6: Interview experience and suggestions

Duration	Discussion
5mins	**Topic: How do you think of the interview experience, and how can it be improved?**
	1. Do you have problem understanding the information given at the beginning of the interview?
	2. Can you make sense of the probabilities given?
	3. Will the following make it easier to understand probability? Graphic representation Example of real life probabilities
	4. Do you have difficulties calculating the cost presented above?
	5. Do you have other comments or suggestions regarding the interview and the study?

Appendix B. Final experimental design

Block	Choice Set	Alternative	Risk	Test cost	Gout treatment cost	Doctor recommended	Most common choice
1	2	1	1 in 600	400	250	No	no info
1	2	2	1 in 5,000	20	1,500	Yes	no info
1	2	3	1 in 500	0	200	No	no info
1	3	1	1 in 1 million	20	4,000	No	No
1	3	2	1 in 600	400	1,500	No	Yes
1	3	3	1 in 500	0	200	Yes	No
1	4	1	1 in 5,000	20	400	no info	Yes
1	4	2	1 in 1,000	1,000	250	no info	No
1	4	3	1 in 500	0	200	no info	No
1	5	1	1 in 1,000	20	250	Yes	Yes
1	5	2	1 in 5,000	400	4,000	No	No
1	5	3	1 in 500	0	200	No	No
1	6	1	1 in 600	1,000	400	no info	No
1	6	2	1 in 1,000	200	4,000	no info	No
1	6	3	1 in 500	0	200	no info	Yes
1	7	1	1 in 1 million	200	250	No	no info
1	7	2	1 in 600	20	4,000	No	no info
1	7	3	1 in 500	0	200	Yes	no info
1	8	1	1 in 1,000	1,000	250	No	No
1	8	2	1 in 1 million	400	400	Yes	Yes
1	8	3	1 in 500	0	200	No	No
1	9	1	1 in 5,000	200	400	No	Yes
1	9	2	1 in 600	20	1,500	Yes	No
1	9	3	1 in 500	0	200	No	No
2	2	1	1 in 1,000	200	4,000	No	Yes
2	2	2	1 in 1 million	1,000	400	Yes	No
2	2	3	1 in 500	0	200	No	No
2	3	1	1 in 1,000	400	400	No	No
2	3	2	1 in 600	200	250	No	No
2	3	3	1 in 500	0	200	Yes	Yes
2	4	1	1 in 1 million	1,000	1,500	No	Yes
2	4	2	1 in 1,000	200	250	Yes	No
2	4	3	1 in 500	0	200	No	No
2	5	1	1 in 5,000	400	250	Yes	no info
2	5	2	1 in 1,000	20	400	No	no info
2	5	3	1 in 500	0	200	No	no info
2	6	1	1 in 1 million	1,000	250	no info	No
2	6	2	1 in 600	400	1,500	no info	Yes
2	6	3	1 in 500	0	200	no info	No
2	7	1	1 in 600	20	4,000	No	Yes

2	7	2	1 in 5,000	200	1,500	Yes	No
2	7	3	1 in 500	0	200	No	No
2	8	1	1 in 600	1,000	250	no info	Yes
2	8	2	1 in 5,000	20	400	no info	No
2	8	3	1 in 500	0	200	no info	No
2	9	1	1 in 5,000	20	400	No	No
2	9	2	1 in 1 million	400	4,000	Yes	Yes
2	9	3	1 in 500	0	200	No	No
3	2	1	1 in 1,000	1,000	4,000	Yes	Yes
3	2	2	1 in 600	20	1,500	No	No
3	2	3	1 in 500	0	200	No	No
3	3	1	1 in 1 million	400	250	No	Yes
3	3	2	1 in 600	200	1,500	Yes	No
3	3	3	1 in 500	0	200	No	No
3	4	1	1 in 600	20	4,000	Yes	No
3	4	2	1 in 5,000	400	1,500	No	No
3	4	3	1 in 500	0	200	No	Yes
3	5	1	1 in 5,000	1,000	4,000	No	Yes
3	5	2	1 in 600	400	1,500	Yes	No
3	5	3	1 in 500	0	200	No	No
3	6	1	1 in 1 million	200	400	No	no info
3	6	2	1 in 5,000	20	250	Yes	no info
3	6	3	1 in 500	0	200	No	no info
3	7	1	1 in 1 million	200	1,500	No	No
3	7	2	1 in 1,000	1,000	400	No	No
3	7	3	1 in 500	0	200	Yes	Yes
3	8	1	1 in 600	200	1,500	No	Yes
3	8	2	1 in 1 million	400	250	No	No
3	8	3	1 in 500	0	200	Yes	No
3	9	1	1 in 5,000	1,000	4,000	no info	No
3	9	2	1 in 1,000	200	400	no info	Yes
3	9	3	1 in 500	0	200	no info	No
4	2	1	1 in 600	20	250	No	no info
4	2	2	1 in 1,000	1,000	1,500	No	no info
4	2	3	1 in 500	0	200	Yes	no info
4	3	1	1 in 5,000	1,000	250	No	Yes
4	3	2	1 in 1,000	400	4,000	No	No
4	3	3	1 in 500	0	200	Yes	No
4	4	1	1 in 1 million	200	4,000	Yes	Yes
4	4	2	1 in 1,000	400	250	No	No
4	4	3	1 in 500	0	200	No	No
4	5	1	1 in 5,000	200	4,000	No	no info
4	5	2	1 in 1 million	1,000	400	Yes	no info
4	5	3	1 in 500	0	200	No	no info
4	6	1	1 in 1 million	20	1,500	No	No

4	6	2	1 in 600	1,000	400	Yes	No
4	6	3	1 in 500	0	200	No	Yes
4	7	1	1 in 5,000	400	4,000	Yes	Yes
4	7	2	1 in 1 million	1,000	1,500	No	No
4	7	3	1 in 500	0	200	No	No
4	8	1	1 in 1 million	200	4,000	No	No
4	8	2	1 in 5,000	20	250	Yes	Yes
4	8	3	1 in 500	0	200	No	No
4	9	1	1 in 1,000	20	1,500	no info	Yes
4	9	2	1 in 600	200	400	no info	No
4	9	3	1 in 500	0	200	no info	No

Appendix C. Survey instrument

A survey on the preferences for genetic testing to prevent severe side effects of medicines

For interviewer use only:

Fill in the spaces that you can at the beginning of the interview, and then enter Time Ended and Total Interview Time after completing the survey.

INTERVIEWER NAME: _____

CASE NO.: _____

DATE OF SURVEY (DD/MM/YYYY): _____/_____/2014

VENUE OF INTERVIEW:

Hospital: ☐ SGH ☐ NUH

Location: ☐ Waiting room ☐ Private room/office_____

TIME STARTED:_____

TIME ENDED:_____

TOTAL INTERVIEW TIME:_____MINUTES

VERSION: V19_block1

---------- **SURVEY STARTS HERE** ---------- ----------

INTRODUCTION

Hello！I am a Doctoral student from the National University of Singapore (NUS). We are conducting a survey to look at patient's preferences for taking a genetic test to minimize severe side effects of some commonly used medicines. I would appreciate if you could spare 15 to 20 minutes to help answer some questions. A $5 NTUC voucher will be given to you at the end as a token of appreciation.

Please feel free to call the study coordinator Di Dong, at Tel: 8298 5633 if you need any clarification on this survey.

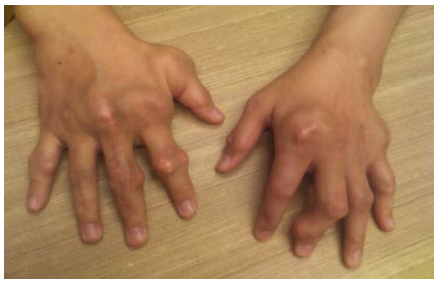
[Obtain respondent's signed consent before proceeding]

SECTION S: SCREENING QUESTIONS	
Question S1	Are you a Singaporean or PR? ☐ YES [PROCEED TO SURVEY] ☐ NO **[THANKS & TERMINATE]**
Question S2	Have you been diagnosed with diabetes? ☐ YES [PROCEED TO SURVEY] ☐ NO **[THANKS & TERMINATE]**

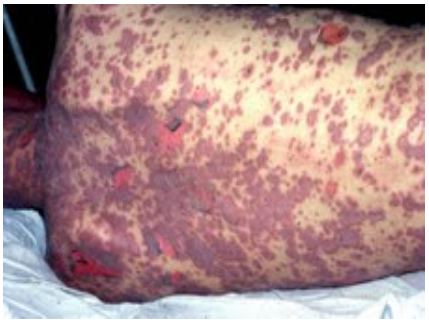
SECTION A: BACKGROUND INFORMATION ON GOUT TREATMENT

Life-threatening side effect of gout medicines

- Gout is a common form of arthritis that can cause severe pain and swelling in the joints (see picture below), often in the fingers and toes. Men and those with diabetes have higher risk of developing gout.



- For chronic gout patients, the standard treatment uses a medicine called allopurinol. It is very effective to treat chronic gout, but can in rare cases cause a severe side effect called Stevens-Johnson syndrome (shown in the picture below). This side effect can lead to death.



- 10% chance of death
- Extreme pain for 2 weeks
- High medical cost (S$5,000-S$20,000)
- May have long-term complications

Different gout treatments (with and without genetic testing)

- This life-threatening side effect is related to the genetics of individuals. A genetic test can be done via a blood test to identify whether or not an individual is more likely to have the side effect.
    - Test positive means you may have the severe side effect if you take the standard medicine.
    - Test negative means you will not have the severe side effect with the standard medicine.

- Now, we have different treatments for gout:
  -Genetic-testing guided treatment

  -Standard allopurinol treatment without testing

- Note that the test is not 100% accurate, so genetic test-guided treatment can not completely prevent the risk of severe side effect. It only reduces the risk.

SECTION B: FEATURES OF GOUT TREATMENTS

[Hypothetical scenario] Suppose you developed gout and had to decide which treatment to choose. Some factors one might consider in making this decision are listed below.

Factor 1: Whether genetic testing is involved

There are two possibilities:

- Test (genetic test-guided treatment)
- No test (no genetic test involved)

Factor 2: The chance of getting the severe side effect

These different treatments are equally effective in treating gout, but differ in the chance of getting the severe side effect. Remember that genetic-test guided treatments also have risk of the severe side effect.

The chance of getting the severe side effect may be:

- 1 out of 500 patients
- 1 out of 600 patients
- 1 out of 1,000 patients
- 1 out of 5,000 patients
- 1 out of one million patients

Question B1	**Among 500 patients who take a medicine, 1 will have a severe side effect.** **How would you feel about this risk?** I would feel at risk. I would not worry about it.
Question B2	**Comparing the two scenarios below, which indicates higher risk?** 1 out of 500 patients get the severe side effect 1 out of 1,000 patients get the severe side effect

Factor 3: Cost of the test

Assume the test must be paid out of pocket. You cannot use health insurance or Medisave. Please think carefully about how the cost of the genetic test would influence your budget (e.g., for food and clothing) before making a decision. The 4 possible test costs are:

- S$20
- S$200
- S$400
- S$1,000

Question B3	**If a genetic test costs S$400, would you consider taking the test to reduce the chance of the severe side effect?** Definitely would Probably would Probably would not Definitely would not

Factor 4: Cost of gout medicines (over 2 years)

As gout medicines should be taken daily for at least 2 years, costs over 2 years are shown here. Assume all treatment costs must be paid out of pocket. You cannot use Medisave or insurance.

If you choose standard allopurinol treatment without genetic testing, the cost is:

- S$200 over two years

If you choose genetic test-guided treatment, treatment cost will depend on your test results. Those who test positive (assumed to be 20% of individuals or 2 in every 10 who take the test) need to take a more expensive alternative medicine, whereas individuals who test negative (8 in 10) can take the standard medicine allopurinol at a cost of $200 over two years. . The four possible costs of the genetic test-guided treatment for those who test positive are:

- S$250 if test positive (2 in 10 chance)
- S$400 if test positive (2 in 10 chance)
- S$1,500 if test positive (2 in 10 chance)
- S$4,000 if test positive (2 in 10 chance)

Question B4	**If the alternative gout medicine costs S$2,000 over two years, and you need to take this medicine if you have test positive (2 in 10 chance) would you consider choosing the genetic test-guided treatment?** Definitely would Probably would Probably would not Definitely would not

Factor 5: Your doctor's recommendation

When you make a decision, you may or may not receive advice from your doctor.

Question B5	**Would your doctor's recommendation influence your decision?** Definitely would Probably would Probably would not Definitely would not

Factor 6: Most common choice

When given several options to choose from, some people are interested to know how other people choose in the same situation. The most common choice here is defined as the option chosen by 80% of people in the same situation.

Question B6	**Would knowing what the most common choice is influence your decision?** Definitely would Probably would Probably would not Definitely would not

SECTION C: TRADE-OFF QUESTIONS

- Suppose you developed gout and were asked to choose your preferred treatment option among several different scenarios.

- Please answer the 10 questions I am going to show you. They may look similar but all differ. In each question, you need to think about the pros and cons of each option.

- When making decisions we ask what you would prefer for yourself, not what you think would be best for your friends or other people.

- Remember, assume that costs must be paid out of pocket; you cannot use Medisave or health insurance.

<u>Example Question</u>: If you had to choose one of the treatment strategies below, which would you choose?

	Treatment A	Treatment B	Treatment C
Whether genetic testing is involved	Test	Test	No Test
The chance of getting the severe side effect	1 out of one million patients	1 out of 600 patients	1 out of 500 patients
Cost of the one-time genetic test	S$200	S$20	$0
Cost of gout medicines (over two years)	**S$1,500** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$400** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$200** over two years
Your doctor's recommendation		Doctor recommended	
Most common choice	No information	No information	No information

Question: If these were the only 3 options available, <u>which ONE would you choose?</u> (Please tick √) ☐ ☐ ☐

Explanation of the above scenario:

- If you chose <u>treatment A</u>, you take a genetic test, which costs S$200. Your chance of getting severe side effect is 1 in one million. If you test positive (2 in 10 chance), you pay S$1,500 over two years for the more expensive gout medicine. If you test negative, you pay $200 over two years for the standard gout medicine.
- If you chose <u>treatment B</u>, you take a genetic test, which costs S$20. Your chance of getting severe side effect is 1 in 600. If you test positive (2 in 10 chance), you will need to pay S$400 over two years for the more expensive gout medicine. If you test negative, you pay $200 over two years for the standard gout medicine.
- If you choose <u>treatment C</u>, you don't need to test, but your chance of getting severe side effect is 1 in 500. The gout medicine costs you S$200 over 2 years.
- In this scenario, treatment B is the doctor recommended option, and you have no information on what the most common choice is.

<u>Question 1</u>: If you had to choose one of the treatment strategies below, which would you choose?

	Treatment A	Treatment B	Treatment C
Whether genetic testing is involved	Test	Test	No Test
The chance of getting the severe side effect	1 out of one million patients	1 out of 600 patients	1 out of 500 patients
Cost of the one-time genetic test	S$200	S$200	$0
Cost of gout medicines (over two years)	**S$400** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$400** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$200** over two years
Your doctor's recommendation	No information	No information	No information
Most common choice	No information	No information	No information

Question: If these were the only 3 options available, <u>which ONE would you choose?</u> (Please tick √) ☐ ☐ ☐

Question 2: If you had to choose one of the treatment strategies below, which would you choose?

	Treatment A	Treatment B	Treatment C
Whether genetic testing is involved	Test	Test	No Test
The chance of getting the severe side effect	1 out of 600 patients	1 out of 5,000 patients	1 out of 500 patients
Cost of the one-time genetic test	S$400	S$20	$0
Cost of gout medicines (over two years)	**S$250** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$1,500** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$200** over two years
Your doctor's recommendation		Doctor recommended	
Most common choice	No information	No information	No information

Question: If these were the only 3 options available, which ONE would you choose? (Please tick √) ☐ ☐ ☐

<u>Question 3</u>: If you had to choose one of the treatment strategies below, which would you choose?

	Treatment A	Treatment B	Treatment C
Whether genetic testing is involved	Test	Test	No Test
The chance of getting the severe side effect	1 out of one million patients	1 out of 600 patients	1 out of 500 patients
Cost of the one-time genetic test	S$20	S$400	$0
Cost of gout medicines (over two years)	**S$4,000** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$1,500** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$200** over two years
Your doctor's recommendation			Doctor recommended
Most common choice		✔	

Question: If these were the only 3 options available, <u>which ONE would you choose?</u> (Please tick √)

☐ ☐ ☐

<u>Question 4</u>: If you had to choose one of the treatment strategies below, which would you choose?

	Treatment A	**Treatment B**	**Treatment C**
Whether genetic testing is involved	Test	Test	No Test
The chance of getting the severe side effect	1 out of 5,000 patients	1 out of 1,000 patients	1 out of 500 patients
Cost of the one-time genetic test	S$20	S$1000	$0
Cost of gout medicines (over two years)	**S$400** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$250** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$200** over two years
Your doctor's recommendation	No information	No information	No information
Most common choice	✔		

Question: If these were the only 3 options available, <u>which ONE would you choose?</u> (Please tick √)

☐ ☐ ☐

Question 5: If you had to choose one of the treatment strategies below, which would you choose?

	Treatment A	Treatment B	Treatment C
Whether genetic testing is involved	Test	Test	No Test
The chance of getting the severe side effect	1 out of 1,000 patients	1 out of 5,000 patients	1 out of 500 patients
Cost of the one-time genetic test	S$20	S$400	$0
Cost of gout medicines (over two years)	**S$250** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$4,000** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$200** over two years
Your doctor's recommendation	Doctor recommended		
Most common choice	✔		

Question: If these were the only 3 options available, which ONE would you choose? (Please tick √) ☐ ☐ ☐

<u>Question 6</u>: If you had to choose one of the treatment strategies below, which would you choose?

	Treatment A	**Treatment B**	**Treatment C**
Whether genetic testing is involved	Test	Test	No Test
The chance of getting the severe side effect	1 out of 600 patients	1 out of 1,000 patients	1 out of 500 patients
Cost of the one-time genetic test	S$1,000	S$200	$0
Cost of gout medicines (over two years)	**S$400** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$4,000** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$200** over two years
Your doctor's recommendation	No information	No information	No information
Most common choice			✔

Question: If these were the only 3 options available, <u>which ONE would you choose?</u> (Please tick √) ☐ ☐ ☐

Question 7: If you had to choose one of the treatment strategies below, which would you choose?

	Treatment A	Treatment B	Treatment C
Whether genetic testing is involved	Test	Test	No Test
The chance of getting the severe side effect	1 out of one million patients	1 out of 600 patients	1 out of 500 patients
Cost of the one-time genetic test	S$200	S$20	$0
Cost of gout medicines (over two years)	**S$250** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$4,000** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$200** over two years
Your doctor's recommendation			Doctor recommended
Most common choice	No information	No information	No information

Question: If these were the only 3 options available, which ONE would you choose? (Please tick √) ☐ ☐ ☐

Question 8: If you had to choose one of the treatment strategies below, which would you choose?

	Treatment A	Treatment B	Treatment C
Whether genetic testing is involved	Test	Test	No Test
The chance of getting the severe side effect	1 out of 1,000 patients	1 out of one million patients	1 out of 500 patients
Cost of the one-time genetic test	S$1,000	S$400	$0
Cost of gout medicines (over two years)	**S$250** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$400** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$200** over two years
Your doctor's recommendation		Doctor recommended	
Most common choice		✓	

Question: If these were the only 3 options available, which ONE would you choose? (Please tick √) ☐ ☐ ☐

Question 9: If you had to choose one of the treatment strategies below, which would you choose?

	Treatment A	Treatment B	Treatment C
Whether genetic testing is involved	Test	Test	No Test
The chance of getting the severe side effect	1 out of 5,000 patients	1 out of 600 patients	1 out of 500 patients
Cost of the one-time genetic test	S$200	S$20	$0
Cost of gout medicines (over two years)	**S$400** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$1,500** if test positive (2 in 10 chance); S$200 if test negative (8 in 10 chance)	**S$200** over two years
Your doctor's recommendation		Doctor recommended	
Most common choice	✔		

Question: If these were the only 3 options available, which ONE would you choose? (Please tick √) ☐ ☐ ☐

SECTION D: BACKGROUND QUESTIONS		
Question D1	**In which year were you born?**_____	
Question D2	**Have you ever been diagnosed with gout, or hyperuricemia (excess uric acid in the blood)?**	
	☐	Yes
	☐	No
	If YES, have you ever taken long-term gout treatment (ie. take medicine daily even when you don't feel the pain)?	
	☐	Yes
	☐	No
Question D3	**Have you ever been diagnosed with hypertension (high blood pressure)?**	
	☐	Yes
	☐	No
Question D4	**How would you rate your general health status?**	
	☐	Very good
	☐	Quite good
	☐	Neither good nor poor
	☐	Quite poor
	☐	Very poor
Question D5	**Have you ever had severe side effects from medicines (such as serious drug allergy)?**	
	☐	Yes
	☐	No
Question D6	**Record the gender**	
	☐	Male
	☐	Female
Question D7	**Which ethnic group do you belong to?**	
	☐	Chinese
	☐	Malay
	☐	Indian
	☐	Other (Please specify:_____)
Question D8	**What type of housing do you live in?**	
	☐	HDB flat (1-2 room)
	☐	HDB flat (3 room)
	☐	HDB flat (4 room)
	☐	HDB flat (5 room and above/HUDC/EC)
	☐	Condominium/Private flat
	☐	Bungalow/ semi-detached/ terrace house/shop house
Question D9	**How many people are there in your household?**_____	
Question D10	**What is the total monthly income of your household (from all sources includes drawing down from savings)?**	
	☐	S$0-1,500

	☐	S$1,500-3,000
	☐	S$3,000-5,000
	☐	S$5,000-8,000
	☐	S$8,000-10,000
	☐	Above 10,000
Question D11	**What is your highest educational level completed?**	
	☐	No formal education
	☐	Primary
	☐	Secondary
	☐	Junior college/ Polytechnic/ Diploma
	☐	University and above
Question D12	**What is your current employment status?**	
	☐	Full-time employment
	☐	Part-time employment
	☐	Self-employed
	☐	Homemaker
	☐	Retired
	☐	Unemployed

----------Thank you very much!----------