The Hospital for Sick Children
Technology Assessment at Sick Kids (TASK)

FULL REPORT

HEALTH TECHNOLOGY ASSESSMENT OF THIOPURINE METHYLTRANSFERASE TESTING FOR GUIDING 6-MERCAPTOPURINE DOSES IN PEDIATRIC PATIENTS WITH ACUTE LYMPHOBLASTIC LEUKEMIA

Authors:

Jennifer R. Donnan, MSc

Health Policy, Management & Evaluation, University of Toronto

Pharmacy Research Specialist, Newfoundland and Labrador Centre for Health Information

Wendy J. Ungar, MSc, PhD

Senior Scientist, Child Health Evaluative Sciences, The Hospital for Sick Children, Toronto; Associate Professor, Health Policy, Management & Evaluation, University of Toronto

Maria Mathews, PhD

Associate Professor, Health Policy/Health Care Delivery

Division of Community Health & Humanities, Health Sciences Centre, Memorial University of Newfoundland

Rebecca Hancock-Howard, MSc, PhD

Project Manager, Child Health Evaluative Sciences, The Hospital for Sick Children, Toronto

Collaborator:

Proton Rahman, MD, FRCPC

Professor of Medicine (Rheumatology), Memorial University of Newfoundland

Corresponding Author:

Wendy J. Ungar, MSc, PhD

EXTERNAL REVIEWERS

Mark Dobrow, MSc, PhD

Assistant Professor, Health Policy, Management & Evaluation, University of Toronto

Scientist and Lead, Cancer Services & Policy Research Unit, Cancer Care Ontario

Nicole Mittman, MSc, PhD

Executive Director, HOPE Research Centre

Scientist, Sunnybrook Health Sciences Centre

Assistant Professor, Department of Pharmacology, University of Toronto

ACKNOWLEDGEMENTS

CONFLICTS OF INTEREST

The authors declare that they do not have any conflicts of interest.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

6-MP	6-Mercaptopurine
ADE	Adverse drug event
ALL	Acute lymphoblastic leukemia
APEX	Arrayed primer extension;
ARMS	Amplification refractory mutation system
AS	Allele specific
CASP	Critical Appraisal Skills Program
CBC	Complete blood count
CEA	Cost-effectiveness analysis
CI	Confidence interval
CNS	Central nervous system
COG	Children's Oncology Group
DHPLC	Denaturing high performance liquid chromatography
DNA	Deoxyribonucleic acid
FDA	Food and Drug Administration
Hb	Hemoglobin
HPLC	High performance liquid chromatography
HTA	Health technology assessment
IBD	Irritable bowel disease
ITPA	Inosine triphosphate pyrophosphatase
LFT	Liver function test
LMS	Laboratory Medicine Service
MTX	Methotrexate
NPV	Negative predictive value
OCCI	Ontario Case Costing Initiative
OHIP	Ontario Health Insurance Program
OR	Odds ratio
PCR	Polymerase chain reaction
PPV	Positive predictive value
QALM	Quality adjusted life-month
QALY	Quality adjusted life-year
RC	Radiochemical

RFLP   Restriction fragment length polymorphism

ROC   Receiver operator characteristic

TPMT   Thiopurine methyltransferase

TMS   Tandem mass spectrometry

WBC   White blood cell

Executive Summary

Introduction

Leukemia is the most common form of cancer in the pediatric population, accounting for 25.3% of all childhood cancer diagnoses. Acute lymphoblastic leukemia (ALL) accounts for 75% of these leukemia diagnoses.

The treatment plan for childhood leukemia involves a multi-drug regimen over four phases, lasting two to three years. The goal is to first put the patient into clinical remission, then to target the cells that are clinically undetectable and finally to maintain the patient in remission. During the final maintenance phase of therapy, an immunosuppressive agent called 6-mercaptopurine (6-MP) is used. The risks of certain adverse drug events (ADE) as a result of 6-MP-treatment are influenced by genetic variations within the population in the enzyme responsible for metabolizing 6-MP, thiopurine methyltransferase (TPMT). The most serious dose dependant ADE over the short-term is myelosuppression, or more specifically, febrile neutropenia. Myelosuppression is bone marrow suppression characterized by a decrease in all the blood components, including red blood cells (anemia), white blood cells (leukopenia) and platelets (thrombocytopenia). If a patient presents with a fever (a sign of infection) and a low neutrophil count (febrile neutropenia), the patient requires hospitalization and immediate treatment with intravenous antimicrobials. Long-term dose dependant side effects include hepatotoxicity and secondary malignancy.

There are currently two methods of detecting TPMT enzyme deficiency: a phenotype test (enzymatic assay) that gives a metabolite activity reading and a genotype test that detects the presence of mutations in the genes responsible for producing the TPMT enzyme. Given the high cost of genetic testing and the importance of preventing serious ADEs, understanding the incremental cost-effectiveness of either form of testing compared to standard care (no testing) would be valuable to guide therapy.

Objectives

The primary objective was to review the literature systematically to determine the accuracy of the TPMT phenotype and genotype tests. The secondary objective was to determine the incremental cost of TPMT genotyping and phenotyping compared to standard weight-based dosing strategies per life-month saved.

Methods

<u>Systematic Review</u>

A systematic review of the literature was conducted to assess the accuracy of the TPMT technologies. Studies were included if they evaluated either a TPMT genotype or TPMT phenotype technology in comparison to a gold standard and showed results on the accuracy of the two tests, using either sensitivity and specificity or positive/negative predictive value. Studies were excluded if they were in a language other than English or evaluated any subject other than humans. The quality of the identified studies was assessed using a modified Critical Appraisal Skills Program (CASP) tool.

<u>Cost-Effectiveness Analysis</u>

A cost-effectiveness analysis (CEA) was carried out from the health care system perspective to compare three testing strategies for 6-MP dosing: genotype-based, phenotype-based and no testing (standard dosing based on weight and height). This analysis was performed on a hypothetical cohort of pediatric patients with ALL and receiving 6-MP for the maintenance phase of therapy. Costs included direct health care costs for testing, drugs, patient monitoring, physician services, and inpatient care for serious adverse events. The time horizon was set at three months to coincide with the period of identifying and treating myelosuppression at the start of 6-MP treatment. Myelosuppression was the only adverse drug effect evaluated. Given the short time horizon, the measure of effectiveness was life-months. To address uncertainty in some of the parameter estimates, univariate sensitivity analyses were conducted for variables of interest and a probabilistic sensitivity analysis (PSA) was conducted using Monte Carlo simulations. Mean costs and their 95% confidence intervals (CIs) were estimated from the PSA.

Results

<u>Systematic Review</u>

Seventeen studies were identified that met the inclusion criteria. Both TPMT phenotype and genotype technologies were considered accurate though there is no gold standard. Additionally, included studies were of low methodological quality according to the CASP tool. The sensitivity and specificity of the genotype test ranged from 55-100% and 94-100%, respectively. The sensitivity and specificity of the phenotype test ranged from 92-100% and 86-98%, respectively.

<u>Cost-Effectiveness Analysis</u>

Neither of the interventions showed a benefit in survival compared to standard dosing, as measured by life months. It is likely that no difference in effectiveness between the test strategies was detected because death following myelosuppression is an extremely rare occurrence and was the only outcome measure evaluated. Also, the homozygous TPMT mutation is so rare that approximately 300 children must be screened before one with a deficiency will be detected. Both testing strategies (genotyping and phenotyping) were more costly compared to standard weight-based dosing. In the base case analysis, the costs per child of the standard dosing, phenotyping and genotyping strategies were $654, $1,020, and $1,090, respectively. As there were no differences in effectiveness, incremental costs were calculated instead of incremental cost-effectiveness ratios. The incremental cost between the phenotyping and standard dosing strategies was $366; between the genotyping and standard dosing strategies was $436; and between the genotyping and phenotyping strategies was $70.

These conclusions were not altered in the PSA, which found that the mean costs per child of the standard, phenotyping and genotyping strategies were $669 (95% CI $547-791), $967 (95% CI $721-1,213), and $946 (95% CI $659-1,233), respectively. The PSA demonstrated that the cost differences between the phenotyping and genotyping tests are likely negligible. The univariate sensitivity analysis showed that the incremental costs between the strategies may be affected by changes in the price of the genotyping and phenotyping tests. If one of the tests was cheaper, it would become the more attractive strategy.

Discussion

This systematic review and cost-effectiveness analysis found that using TPMT phenotype or genotype tests prior to the first dose of 6-MP therapy did not prove to be cost-effective compared to standard weight-based dosing. This assessment highlights a number of important issues and gaps in the literature.

With respect to the TPMT tests, it was found that the phenotype tests identified more positive results compared to the genotype tests because they detected all deficiencies in the enzyme, not only those influenced by TPMT gene mutations. Genotype tests were accurate; however they were limited by the number of mutations the test was designed to detect. As a result, neither test could be considered the gold standard.

No difference in life-months was detected between the three strategies. Since there was no difference in effectiveness between the three arms of the decision tree, it was not possible to

calculate an ICER. The reduction in the occurrence of neutropenia is only one outcome measure that could be used to determine the benefits of TPMT testing. Future research should consider other ADEs such as liver toxicity, as well as efficacy outcomes such as long-term survival, rate of relapse and development of secondary malignancy. However, there is presently very little available evidence on the incidence and impact on survival for these outcomes. As a result, they could not be considered in this study.

The analysis showed that there would be an additional cost to offering either the phenotype test or genotype test prior to dosing 6-MP over the standard of care as described in the Children's Oncology Group protocols. Thus these alternatives were not cost effective to reduce the mortality and morbidity associated with 6-MP-induced neutropenia. The impact of dose reducing patients who received false positive test results was also not considered. It is possible that the false positives who are dose reduced will be under-dosed, potentially compromising their treatment.

Four previous economic evaluations have examined the assessment of TPMT activity prior to 6-MP dosing to prevent ADEs, however only one evaluated a pediatric ALL population. These evaluations have mainly concluded that the TPMT technologies were cost-effective, however many differences existed in the models used in those studies compared to the current study.

The study was limited by the data available through the systematic review. As studies in languages other than English were not included, it is possible that relevant studies were not identified. The assessment of quality-adjusted life-months or life-years was not possible due to a lack of data.

Conclusions

At this time there is insufficient evidence to recommend the use of phenotype or genotype testing prior to 6-MP therapy to guide initial doses in pediatric ALL patients. Institutions that follow the COG guidelines should not be affected by the results of this assessment. Institutions who have adopted the screening for TPMT status prior to the first dose of 6-MP should review their current practice. Currently the costs of these tests in the pediatric ALL population are funded by the health care system. The opportunity costs of using such tests outside clinical guidelines need to be taken into consideration. Policies should outline which clinical scenarios are eligible for publicly funded TPMT testing. Health care organizations will need to be prepared for a potential increase in public pressure for such tests as their availability becomes more widely known. Health technology assessment agencies can play a role in disseminating health economic evidence to inform decision making with respect to pediatric TPMT technologies.

1 INTRODUCTION

Leukemia is the most common form of cancer in the pediatric population, accounting for 25.3% of all childhood cancer diagnoses. Acute lymphoblastic leukemia (ALL) accounts for 75% of these leukemia diagnoses.[1] While leukemia still causes more deaths in children less than 20 years of age than any other childhood cancer, survival is improving as therapies are discovered. Today approximately 10% of children diagnosed with ALL will die as a result of their disease,[1] compared to 91% in the 1960s.[2]

The treatment plan for childhood leukemia involves a multi-drug regimen over four phases, taking between two and three years to complete.[3] The goal is first to put the patient into clinical remission, then to target the cells that are clinically undetectable and then finally to maintain the patient in remission. Central nervous system (CNS) prophylaxis consists of concurrent chemotherapy agents that are given throughout the entire three years and is intended to prevent relapse of leukemic meningitis. During the final phase of therapy (the maintenance phase), an immunosuppressive agent called 6-mercaptopurine (6-MP) is used. This agent is of particular interest because the risks for certain adverse drug events (ADE) are influenced by genetic variations within the population in the enzyme responsible for metabolizing 6-MP, thiopurine methyltransferase (TPMT). The most serious dose dependant ADE over the short-term is myelosuppression or more specifically neutropenia. Long-term dose dependant side effects include hepatotoxicity and secondary malignancy.

There are currently two methods of detecting a TPMT enzyme deficiency. The first is a phenotype test (enzymatic assay) that gives a metabolite activity reading. The second is a genotype test that detects the patient's individual genetic make-up for producing the TPMT enzyme. There is no available evidence to show which method is better or the most cost effective approach for preventing ADEs. Given the increasing interest in genotyping to guide medical

therapy, it is anticipated that the genotype test, though more costly, may be more accurate at identifying individuals with a deficiency. For children with ALL, the Children's Oncology Group[4] protocols (2008) recommend using one of the tests in specific clinical scenarios, and they state that the "genotyping test has a low false negative rate, and may be preferable to TPMT phenotype testing in cases where a history of red cell transfusions would potentially confound assessments of RBC [red blood cell] TPMT activity".[4] Practice varies across treatment centres. Some centres perform a genotype test before they start treatment with 6-MP to guide therapy, while other centres use a standard dosing strategy based on weight and height (hitherto referred to as 'weight-based') and only test (genotype or phenotype) when ADEs interrupt therapy.

The goal of this report is first to review the literature systematically to determine the accuracy of the TPMT phenotype and genotype tests and second to determine the incremental cost-effectiveness of the tests compared to standard care with respect to their ability to save life months and avert ADEs.

1.1 Hypothesis

Using the TPMT phenotype or genotype test to guide initial doses of 6-MP will prove to be a cost-effective screening tool for preventing ADEs and improving survival relative to weight-based dosing.

1.2 Research Questions

Primary question:

1. What is the accuracy of each of the various phenotype and genotype tests for detecting TPMT polymorphisms?

Secondary question:

2. From the health care system perspective, what is the cost-effectiveness of TPMT genotyping or phenotyping compared to no testing for individualizing 6-MP dosing and preventing ADEs and improving survival in pediatric patients with ALL?

1.3 Research Goals

1. The primary goal of this report was to systematically review the literature to determine the accuracy of phenotype and genotype tests used to detect TPMT polymorphisms. This was assessed by comparing sensitivity, specificity and positive and negative predictive values from studies evaluating two different forms of TPMT testing.

2. The secondary goal of this report was to determine the incremental cost of TPMT genotyping and phenotyping compared to weight-based dosing strategies per life-month saved.

1.4 Rationale

Patients taking 6-MP are at risk for a number of serious, life threatening events that can be dose-related. By knowing an individual's TPMT genotype or phenotype, clinicians can determine whether a patient requires dose reductions of 6-MP, thus reducing the chance that these events occur. This of course assumes that physicians know how to interpret the genotype and phenotype information and translate it into an appropriate dose for the patient.

There is increasing evidence that both of these methods of testing are effective predictors of ADEs with 6-MP. The Food and Drug Administration (FDA) in the United States has recommended that product information on 6-MP include warnings on the increased risk for adverse events in patients deficient in the TPMT enzyme.[5] They were unable to make the recommendation that either method of testing be performed in all patients who are to receive 6-MP because of the limited amount of evidence. This evaluation will contribute to the evidence to support such recommendations.

1.5   Ethical Considerations

This study was approved by the Research Ethics Board of the University of Toronto and was conducted in accordance with the Tri-Council Policy Statement.[6]

1.6   Literature Review

Pharmacogenetics is the study of human response to medications with respect to genetic variations. Numerous genetic variations have been studied in the medical literature, but not many have been examined to the same degree as TPMT mutations. TPMT has a very important role in the metabolism of 6-MP, an essential component of the treatment protocol for pediatric ALL.[4] Every medication has a concentration in which it is effective and minimally toxic, known as the therapeutic range. Chemotherapy agents usually have a very narrow therapeutic range. ADEs from chemotherapy treatments can not only decrease a patient's quality of life, but some ADEs can be life threatening. To understand how TPMT genetic testing can play a role in reducing ADEs in ALL patients, this literature review will present a background for ALL and its treatments, present a closer look at 6-MP and its metabolic pathway, discuss the difference between genotype and phenotype tests and look at the link between genetics and drug metabolism.

1.6.1 Acute Lymphoblastic Leukemia

Leukemia is a malignancy in the circulation system whereby the blood forming cells of the bone marrow undergo unregulated reproduction. There are four main types of leukemia: acute lymphoblastic leukemia (ALL), acute myeloid leukemia (AML), chronic lymphoblastic leukemia and chronic myeloid leukemia. The leukemias are classified based on cell origin, patient life expectancy, clinical presentation, rate of progression and response to therapy.   Unlike solid tumor cancers, which can be staged according to the degree of progression, the severity of leukemia is categorized by the acute or chronic diagnosis. Acute leukemia is a rapidly progressing disease in which immature cells replicate without the normal regulation of the body.

Both forms of acute leukemia (ALL and AML) originate from a single leukemic cell that replicates itself. In the case of ALL, it begins with a lymphoblast. These immature leukemia cells inhibit blood cells from maturing naturally. Normally, immature blood cells undergo a process termed hematopoiesis, which results in the production of the various types of blood cells. This process is interrupted by the immature non-functional blasts, resulting in leukemia.[3]

If left untreated, acute leukemias will progress rapidly and result in death in two to three months.[3] The presenting symptoms of patients with leukemia are many and varied, but commonly consist of a one to three month history of fatigue, malaise and pallor due to anemia. Other potential symptoms are infection and fever due to a drop in normal leukocyte levels (granulocytopenia), bruising, frank bleeding and small hemorrhagic spots on the skin (petechiae) due to decreased platelets (thrombocytopenia), enlarged spleen (splenomegaly) and enlarged liver (hepatomegaly) or sternal tenderness due to leukemic infiltration.[3]

1.6.2 Treatment for Acute Lymphoblastic Leukemia

The treatment for ALL is divided into three phases and CNS prophylaxis as shown in Table 1. The first is the remission induction phase, which is intended to put the patient into complete clinical and hematologic remission. The basic drugs used in this phase are vincristine and prednisone, which achieve complete remission in 85% of children and 50% of adults. The addition of a third induction agent, usually an anthracycline (doxorubicin) or asparaginase, increases the rate of complete remission even further to >95% in children and 83% in adults. Beyond these three agents, the length of remission may be prolonged with additional medications but at the risk of increased toxicity. Only high risk cases are exposed to more then three induction agents. Since all adult cases of ALL are considered higher risk, their normal regimes consist of four to seven drugs.[3]

Table 1: Phases of Chemotherapy in Patients with Acute Lymphoblastic Leukemia

Phase	Length of Treatment	Purpose
Phase 1: Remission Induction	4-8 weeks	Intended to put the patient into clinical remission. Ninety-five percent of children reach this endpoint.
Central Nervous System Prevention Therapy	Throughout entire therapy	Starts during the induction phase and continues throughout therapy; designed to prevent relapse with leukemic meningitis.
Phase 2: Consolidation Therapy	3-9 months	Designed to eliminate the small number of leukemic lymphoblasts that remain after clinical remission is achieved.
Phase 3: Maintenance Therapy	Continues until months 30-36	Less intensive continuation of the chemotherapy regimen, designed to keep patients in remission.

(Day and Henry, 2002)[3]

CNS prophylaxis is started shortly after the induction phase has begun and continues throughout the rest of treatment. The reason for this prophylaxis is two-fold. First, the induction chemotherapy agents do not readily cross the blood-brain-barrier. Second, even if CNS involvement is not evident at diagnosis, this is where many patients relapse. Agents used in this phase for pediatric patients generally consist of intrathecal (spinal injection) methotrexate (MTX), or triple intrathecal chemotherapy consisting of MTX, cytarabine and hydrocortisone. In adult patients intrathecal MTX may be combined with cranial irradiation.[3]

The second phase (consolidation therapy) is started after complete clinical remission is achieved. The intention of this phase is to eradicate the cells that are clinically undetectable. Agents used in this phase vary depending on which research group's protocols are being followed and may include 6-MP for low-medium risk patients. The regime may consist of drugs that are not cross-

resistant to those used in induction and CNS prophylaxis, or may be a more dose-intensive use of agents used before.[3]

The third and final phase is the maintenance phase, wherein long-term drug exposure can eliminate slowly dividing cells and allow the immune system to eradicate leukemia cells. The goal is to prolong remission and prevent relapse. MTX and 6-MP are the agents of choice for this phase. Pulse therapy of vincristine and prednisone are also used intermittently throughout the maintenance phase.[3]

1.6.3 6-Mercaptopurine

6-MP is a pro-drug that is converted inside human cells to an active form. It is an immunosuppressant that produces it's cytotoxic (cell destroying) effects by incorporating itself into deoxyribonucleic acid (DNA) during the cell development cycle. Like most cytotoxic agents, 6-MP has a narrow therapeutic range, therefore any alterations in the metabolism and elimination of these drugs could lead to accumulation and potentially life-threatening ADEs within individuals.

The most concerning ADE following treatment with 6-MP is myelosuppression. Myelosuppression is bone marrow suppression characterized by a decrease in all the blood components, including red blood cells (anemia), white blood cells (leukopenia) and platelets (thrombocytopenia). The neutrophil is the most abundant of the various white blood cell types and plays an important role in the immune system. When there is a drop in the neutrophil count (neutropenia) the patient is at an increased risk of infection, which can be fatal in the ALL population. If a patient presents with a fever, a sign of infection, and a low neutrophil count (febrile neutropenia) the patient is immediately treated with intravenous antimicrobials. Myelosuppression is a dose-related effect and substantial dose reductions may be required in patients unable to tolerate the drug. Other potential side effects are presented in Table 2.

Table 2: 6-Mercaptopurine Side Effect Profile

Organ Site	Side Effect (Prevalence)
Allergy / immunology	hypersensitivity (2-3%)
Blood / bone marrow / febrile neutropenia	anemia (>10%)*
	leukopenia (>10%)*
	thrombocytopenia (>10%)*
Constitutional symptoms	fever (1-10%)
Dermatology / skin	alopecia (<1%)
	hyperpigmentation (1-10%)
	rash (1-10%)
Gastrointestinal	emetogenic potential (rare)
	abdominal cramps (1-10%)
	anorexia (1-10%)
	diarrhea (1-10%)
	intestinal ulceration (<1%)
	nausea and vomiting (1-10%)
	stomatitis (1-10%)
Hepatobiliary / pancreas	hepatotoxicity (30%)
Infection	predisposed to bacterial/parasitic infections due to immunosuppression
Metabolic / laboratory	hyperuricemia (1-10%)
Renal / genitourinary	renal toxicity (1-10%)
Secondary malignancy	leukemia and myelodysplasia (<1%)
	cysts and polyps (2-6%)
Sexual / reproductive function	increased risk of abortion if taken in first trimester of pregnancy
	oligospermia, transient (<1%)
Syndromes	tumour lysis syndrome (rare; certain patients may be at increased risk)

Reproduced with permission from the BC Cancer Agency Professional Drug Index (BC Cancer Agency, 2007)[7]
*Onset 7-10 days, nadir 14-16 days, recovery 21-28 days

1.6.4 Thiopurine Methyltransferase

The metabolism of 6-MP involves three enzymes that convert 6-MP to both active and inactive thiopurine metabolites. Xanthine oxidase and TPMT produce the inactive metabolites and hypoxanthine guanine phosphoribosyl transferase initiates the pathway to active metabolites. TPMT also produces an active metabolite, however the literature suggests that the pathway to the inactive metabolite might be more dominant.[8] If an individual has high TPMT activity, then that patient will require a higher dose of 6-MP to produce enough active metabolites. Likewise, if an individual has low TMPT activity, the patient will require a much lower dose to avoid toxicity from too many active metabolites.[9] Since many of the ADEs from 6-MP are dose-related, patients with mutations in the TPMT allele are at significant risk, as these mutations result in decreased TPMT activity. There is also speculation regarding an association between excessive myelosuppression and secondary malignancies.[10]

Individuals are classified into three categories based on their genotype. They can be either wild type, which is the normal genotype, or they can have a heterozygous mutation or a homozygous mutation as described in Table 3.

Table 3: TPMT Genotype Classification and Implications to Therapy

Genotype	Activity Level	Frequency in the population	Dose Adjustments
Wild Type	High TPMT activity	89%*	Start at normal dose, may need to increase**
Heterozygous Mutation	Intermediate TPMT activity	11%*	Start at 60% of normal dose**
Homozygous Mutation	TPMT deficiency	0.33%*	Start at <10% of normal dose**

*Weinshilboum, 1980[11]
** COG, 2008[4]

This mutation in TPMT can be explained by 23 polymorphisms including TPMT*3A which is the most prevalent in the Caucasian population.[12] A dose reduction of approximately 90% is required for patients with the homozygous mutation and dose reductions of 40% may be required for patients with the heterozygous mutant.[4]

1.6.5 Ethnic Variations in Thiopurine Methyltransferase Mutations

As with all gene mutations, variations in TPMT alleles are subject to ethnic differences. A substantial amount of literature has been published that identifies the prevalence of specific mutations in different populations.[13-33] Of the ethnic populations studied to date, TPMT mutation *3A is the most common followed by *3C. Most studies only looked for TPMT *2 and *3 mutations. The prevalence of the less common mutations is still unknown. There was also limited attention paid to sample size calculations, and since most papers did not provide confidence intervals it is difficult to place much strength in these comparisons. However, by comparing the prevalence of the TPMT *3A mutations, which was tested in all studies, it seems that there is potential for considerable variability to exist. There are limitations to these studies however. Many studies had small sample sizes and very few tested for a wide range of known mutations. As a result a true picture of ethnic trends is not possible. However the potential for ethnic variation needs to be taken into consideration when interpreting study results and implementing policy decisions.

1.6.6 Phenotype and Genotype Determination of Thiopurine Methyltransferase Deficiency

There are currently two available methods to diagnose a deficiency in TPMT. The first is a phenotype test that measures the activity of the enzyme in the blood.[34] The second is a genotype test that will detect the polymorphism in the DNA.[35] The phenotype test is thought to have a number of limitations because there are factors that could influence the results. For example,

recent blood transfusions, medications (6-MP itself can induce TPMT activity), alcohol and food

can all have an effect on enzyme activity, and as a consequence alter the results of this test.

Therefore, this test may have to be repeated on a regular basis throughout treatment.[36] Also,

inter-laboratory results do not always agree, probably due to differences in assay conditions in

different laboratories.[35]

The genotype test is not influenced by exogenous factors as it looks directly at the DNA

responsible for the enzyme production and therefore it does not have to be repeated. The

sensitivity of the genotype test however depends on the number of polymorphisms it is designed

to detect. Since it is thought that over 85% of polymorphisms are the result of three specific

mutations, many tests are only designed to pick up the most common ones, leaving up to

approximately 15% of the mutations undetected with the genotype test. Even if the genotype test

was designed to detect all known mutations, there is still the possibility that undiscovered

mutations exist leading to potentially false negatives.

1.6.7 Variety of Thiopurine Methyltransferase Phenotype and Genotype Technologies

1.6.7.1 Phenotype Technologies

The most common method for determining the activity of the TPMT enzyme is to conduct high

performance liquid chromatography (HPLC) followed by an enzymatic assay. Other methods

include tandem mass spectrometry and radiochemical assay. HPLC is a process conducted to

separate the protein of interest from the rest of the proteins in the sample processed from drawn

blood. It uses a column that holds a chromatographic packing material (stationary phase) to

which a solvent is added (mobile phase). The test sample is then added to the column and reacts

with the chromatographic material and solvent, then leaves through the end of the column, with a

pump facilitating movement through the column. The amount of time that it takes to move through

the column is referred to as the retention time and is unique to the protein. The protein then is isolated, incubated at body temperature (37° C) and exposed to the thiopurine drug. After one hour the metabolite is measured to determine how much of the drug was broken down by the enzyme (enzymatic assay) (J. Doré, personal communication, February 11, 2009 (Appendix 1)).

Mass spectrometry, especially liquid chromatography coupled to tandem mass spectrometry, is another technology that can be used to determine enzyme activity. Mass spectrometers have the capacity to differentiate between substrates and products of enzyme reactions based on molecular weight. To measure molecules by mass spectrometry the molecules must first be charged, which is achieved by converting the sample to an electrospray, which adds a positive charge to all the molecules separated in the liquid chromatography system and entering into the mass spectrometer ion source. These charged molecules then pass through a mass analyzer that is set up specifically to detect the mass of the molecule of interest. There is a chance that there is more than one molecule that has the same molecular weight. To reduce the risk of isolating two different molecules with the same molecular weight, a tandem mass spectrometer can be used. A tandem mass spectrometer has two mass spectrometers separated by a collision cell. The collision cell breaks down the isolated molecules to its daughter ions, which each have different molecular weights. The daughter ions then pass through the second mass spectrometer, which is set up to isolate molecules of a weight that is specific to a daughter ion of the molecule of interest. Since the probability of two molecules with the same molecular weight separated by an appropriately set up liquid chromatography system also having the same daughter ions is unlikely, the use of a tandem mass spectrometer allows the user to have more confidence that the isolated substance can be accurately detected and quantified. In the case of TPMT the substrate would be a thiopurine drug. The amount of the product formed by action of TPMT in a certain period of time can be used to indicate the activity of TPMT by monitoring the appearance of product or

disappearance of substrate by liquid chromatography coupled with tandem mass spectrometry (E. Randell, personal communication, March 19th, 2009 (Appendix 1)).

The radiochemical test is an enzymatic assay that uses radioactive substrates for the enzyme. Activity is determined by using different amounts of radioactive product. It does not generally require purification of the enzyme first like HPLC. Its advantages are speed and relatively inexpensive equipment. The drawback is the potential for a contaminating enzyme giving false positive activity and the radioactive risks (J. Doré, personal communication, February 11, 2009 (Appendix 1)).

1.6.7.2 Genotype Technologies

TPMT genotypes are generally determined using a method of polymerase chain reaction (PCR). Some of the older PCR techniques include restriction fragment length polymorphism PCR, restriction digest PCR and amplification refractory mutation system PCR, which all use the same basic technique to genotype. A single strand of DNA (primer) is amplified and then a restriction enzyme is applied which will show if the mutation is present. Restriction enzymes are unique to the allele mutation of interest. This process can only detect one mutation at a time. When there is more than one mutation of interest, the sample must be run through the process multiple times using the appropriate restriction enzymes. This test can become time consuming and in the case of TPMT, where there are at least 23 known mutations, TPMT tests may be limited to only the most common mutations making them less sensitive. The more mutations detected, the more expensive the test becomes, as each restriction enzyme has to be purchased separately.

A newer technology, called the multiplex PCR, allows for a simultaneous amplification of multiple mutations in a single reaction using more than one primer. There are some drawbacks to using this test starting with the cost for capital equipment. If purchasing equipment is possible, it is

necessary to have experienced personnel to set up the initial conditions and test standards would have to be implemented (J. Doré, personal communication, February 11, 2009 (Appendix 1)).

Biochips are the newest of the genotype technologies. The chips contain standardized samples and are automated to allow for a high throughput. The test requires only a very small amount of the required reagents. Multiple mutations can be tested in one reaction without the expense of restriction enzymes. Capital cost and maintenance for this technology is expensive, though the materials to run each test are not. For this reason, biochip technology is only available in large scale laboratories that conduct a high volume of tests (J. Doré, personal communication, February 11, 2009 (Appendix 1)).

1.6.8 Validity of Screening Testing

All diagnostic and screening tests are subject to some measure of error. The error could be a result of the capability of the technology, the skills and knowledge of the tester, or the variability between patients and confounding factors present when the test takes place. No matter what the reason, false negatives or false positives are bound to occur. To visualize how these test results are categorized a contingency table (2x2) has been created using phenotype screening for TPMT mutation as an example (Table 4).

Table 4: Screening Contingency Table

	Mutation Present	Mutation Absent	Totals
Test Positive	True Positive (9)	False Positive (2)	Total persons with positive test (11)
Test Negative	False Negative (1)	True Negative (88)	Total persons with negative test (89)

Totals	Total persons with mutation (10)	Total persons without mutation (90)	Total tested (100)

1.6.8.1 Sensitivity and Specificity

Sensitivity and specificity are two ways of measuring the accuracy of diagnostic or screening tests. Using the example of TPMT phenotype screening, sensitivity refers to the number of positive results in the patients who truly have the mutation. Mathematically this equals the true positives (those with the disease who were detected by the test) divided by the sum of the true positives and false negatives (the total number of people with the disease) multiplied by 100.[37]

$$Sensitivity = \frac{TruePositives(9)}{TruePositives(9) + FalseNegatives(1)} \text{ x } 100\%$$

$$= 90\%$$

Specificity refers to the number of patients without the mutation that will be confirmed with a negative test result. Mathematically this equals the true negatives (those without the disease who were detected by the test) divided by the sum of the true negatives and the false positives (the total number of people without the disease) multiplied by 100.[37]

$$Specificity = \frac{TrueNegatives(88)}{TrueNegatives(88) + FalsePositives(2)} \text{ x } 100\%$$

$$= 97.8\%$$

These calculations are simply estimates of the accuracy of the test. Different values may result when another person performs the test or if a different set of patients are tested. To get a better sense of how accurate the test truly is, a confidence limit can be determined for point estimates of sensitivity and specificity by calculating the standard error. This is represented in the equation

below, where p is the sensitivity or specificity in the decimal fraction and n is the total number of patients tested.[37]

$$\text{Standard Error} = \sqrt{\frac{p(1-p)}{n}}$$

The standard error of the sensitivity in the example above would be:

$$\text{Standard Error} = \sqrt{\frac{0.9(1-0.9)}{100}}$$

$$= 0.03$$

To calculate the 95% confidence interval (CI) the standard error is multiplied by the constant known as the z-score (1.96).[37]

95% CI =   Sensitivity ± (Standard Error * z)

        =   0.90 ± (0.03*1.96)

        =   0.90 ± 0.059

        =   [0.841, 0.959]

1.6.8.2 Positive and Negative Predictive Value

Positive and negative predictive values are another way of calculating the accuracy of diagnostic screening tests. The positive predictive value (PPV) tells us the probability that the mutation is present if there is a positive test result. Mathematically this is equal to the number of true positives divided by the sum of the true and false positives multiplied by 100 percent. Using the data presented in Table 4, the PPV is calculated to be 81.8%.[37]

$$\text{Positive Predictive Value} = \frac{TruePositives(9)}{TruePositives(9) + FalsePositives(2)} \text{ x } 100\%$$

$$= 81.8\%$$

The negative predictive value (NPV) tells us the probability that the mutation is absent when there is a negative test result. Mathematically this is equal to the number of true negatives divided by the sum of the true and false negatives. Using the data presented in Table 4, the NPV is calculated to be 98.9%.[37]

$$\text{Negative Predictive Value} = \frac{TrueNegatives(88)}{TrueNegatives(88) + FalseNegatives(1)} \text{x } 100\%$$

$$= 98.9\%$$

1.6.8.3 Receiver Operator Characteristic Curves

Not all screening or diagnostic tests give binary results. In many circumstances a continuous value is given and a cut-off point is needed to determine the threshold for ascertaining a positive value. In the case of TPMT phenotype tests, the activity of the enzyme is determined and it is up to the diagnostician to determine if a mutation in the gene is likely to be present. In such cases a receiver operator characteristic (ROC) curve is used, which demonstrates the impact on sensitivity by varying the cut-off point. ROC curves allow the establishment of cut-off points that maximize sensitivity and specificity.[37]

1.6.8.4 Test Accuracy and Gold Standards

To test sensitivity, specificity, PPV or NPV, the investigational test must have a gold standard test for comparison. There must be another way to confirm if the disease, or in this case gene mutation, does in fact exist in order to calculate true and false negatives and positives. The ideal comparator would be a test that had as close to 100% sensitivity and specificity as possible.

Pharmacogenetics is still a very new field and in many cases true gold standards do not exist, limiting the true accuracy of these validity calculations.

1.6.9 Standard of Care for TPMT Testing

For hospitals in North America that follow the COG protocols,[4] there are clear guidelines indicating when TPMT testing should be completed. Currently, dosing is determined by the weight of the patient and is based on normal drug requirements for an individual and does not account for potential genetic mutations. TPMT testing is not routinely done prior to 6-MP therapy to individualize therapy, but instead is done after side effects delay therapy by two or more weeks or if the degree of myelosuppression is disproportionate to therapy. The guidelines do not make a clear recommendation for one test over the other but also state that "the genotype test has a low false negative rate, and may be preferable to TPMT phenotype testing in cases where a history of red cell transfusions would potentially confound assessments of RBC TPMT activity" (pp. 46).[4] They also suggest that if the two tests were combined the accuracy would be greater than either of the two alone.

The COG recommends dose modification in patients with homozygous mutations to 10-20 mg/m^2/day three days a week.[4] It is predicted that approximately 35% of TPMT heterozygous patients require dose adjustments. Therefore dose modifications are only recommended for patients with heterozygous mutations if they have experienced significant myelosuppression. In this case the dose should be reduced by 30-50%.[4]

1.6.10   Health Technology Assessment

Health technology assessment (HTA), like evidence-based medicine and clinical practice guidelines, is a part of a group of best practice activities done in the health care setting. HTAs are

conducted differently than evidence-based medicine and clinical practice guidelines, which are created to help guide clinical decisions. The purpose of HTA is to guide policy decisions.[38]

Authors of the EUR-ASSESS (1997) project define HTA as "a multidisciplinary activity that systematically examines the technical performance, safety, clinical efficacy and effectiveness, cost, cost-effectiveness, organizational implications, social consequences, and legal and ethical considerations of the application of a health technology" (pp. 222).[39]

The decision to implement a new health technology, whether it is a drug, a diagnostic procedure or a new device, needs to be based on more then just efficacy. It must also be based on all of the other factors that influence its use. Most basic HTAs will include a systematic review of efficacy and safety as well as some method of economic analysis that determines if the technology gives enough benefit for the money spent. Other aspects of HTA that should be considered, depending on the nature of the technology, are the psychological, social, ethical and organizational implications.[38]

An example of a psychological issue would be the anxiety experienced by a patient, and their parents, who is unable to have access to a TPMT genotype test. Such feelings of fear and anxiety in patients might make this technology desirable for physicians to routinely recommend and for service providers to cover. On the other hand, having the test available may cause undue anxiety over the knowledge of a positive test result.

Social issues that might arise include equity. The introduction of a new but expensive technology that is not covered by health plans can create a social divide among those people who can afford to pay for it out of pocket and those who cannot.

Some technologies may foster judgments which raise ethical dilemmas. For example, prenatal testing for disabilities may seem disrespectful to the population who live with those disabilities.

Organizational and professional issues that might have an impact on a policy maker's decision to introduce a new technology include how much training is required for staff, what additional staff might be needed, or if the technology will reduce hospital stays or emergency room visits.

The goal for this HTA was to critically review the accuracy of TPMT testing technologies and to determine if TPMT testing prior to initial dosing of 6-MP is a cost-effective intervention for preventing ADEs.

2 METHODS

2.1 Technology Review

The technology review of TPMT phenotyping and genotyping strategies was completed in accordance with standard methods for conducting and reporting systematic reviews.[40] A study protocol was developed *a priori* and followed during the review.

2.2 Literature Search Strategy

Three data sources were used to find relevant data for this systematic review: electronic databases, websites and references of retrieved studies. Three databases were searched for relevant trials: Cochrane Database of Systematic Reviews and Controlled Trials (CENTRAL), PubMed/MEDLINE and EMBASE, from the inception of the databases to January 2009. Websites included the FDA Centre for Drug Evaluation and Research. Only studies written in the English language were eligible for inclusion. Electronic database searches were completed with the assistance of a professional librarian. Search terms 'thiopurine methyltransferase', 'TPMT', 'accuracy', 'sensitivity', 'specificity' and 'positive and negative predictive value' were used to identify relevant literature (Table 5).

Table 5: Literature Search Strategy

Database / Source	Search Strategy
PubMed/MEDLINE	1. Thiopurine Methyltransferase OR TPMT 2. Accuracy OR Sensitivity OR Specificity OR Positive Predictive Value 3. #1 AND #2
Embase	1. Thiopurine Methyltransferase OR TPMT 2. Accuracy OR Sensitivity OR Specificity OR Positive Predictive Value 3. #1 AND #2
Cochrane Database	1. Thiopurine Methyltransferase OR TPMT
Food and Drug Administration	1. Thiopurine Methyltransferase OR TPMT

The website for the Centre for Drug Evaluation and Research was searched for additional unpublished grey literature.[41] References from all retrieved studies and selected review articles were then reviewed for additional trials for inclusion.

2.3 Selection Criteria and Method

2.3.1 Inclusion Criteria

1. Studies that evaluated either a TPMT genotype or TPMT phenotype technology in comparison to a gold standard.

2. Studies that showed results on the accuracy of the two tests, using either sensitivity and specificity or positive/negative predictive value.

2.3.2 Exclusion Criteria

1.   Studies not published in the English language.

2.   Studies evaluating any subjects other than humans.

2.4   Data Extraction Strategy

2.4.1 Study Characteristics

An electronic data extraction spreadsheet was created using Microsoft Excel™ to systematically collect relevant data from each study. Characteristics recorded included participant ethnicity (e.g. ethnic background or the country in which the study was conducted), study population (e.g. health status, inclusion and exclusion criteria and age distribution), and test characteristics. Data pertaining to test characteristics were also recorded, including test type (e.g. PCR, HPLC) and reference activity ranges used to classify phenotypes.

2.4.2 Relative Accuracy of Phenotype and Genotype Technologies

The measurements describing the accuracy of the tests (e.g. sensitivity and specificity, positive/negative predictive values) were extracted. Where possible, the raw findings, broken down by activity ranges or wild type and heterozygous/homozygous mutant alleles, were compiled in 2x2 tables (i.e. true or false positive or negative values). Where raw data were not available, 2x2 tables were calculated using available values such as sensitivity, specificity and total sample. Reference activity ranges used were those defined by the individual studies.

2.5   Strategy for Quality Assessment

Quality of the included studies was assessed using a modified version of the Critical Appraisal Skills Program (CASP) tools for evaluating diagnostic test studies.[42] The CASP tool consisted of three main categories of questions, the first being, "Are the results of the study valid?" which

questions the studies' methods to produce a reliable result. For example "Was there a clear study question?", "Was there an appropriate reference standard?", "Was the evaluation blinded?", "Was the population described" and "Was there sufficient detail on how the test was performed?" The second category asks "What are the results?" referring to how the results were presented and how likely they were to have occurred by chance. The final category of the tool asks "Will the results help me and my patients/population?" referring to the applicability of the tests, results and outcomes to specific patient populations. This final category was removed in the modified version of the tool because clinical outcomes were not assessed in this study.

The tool did not, however, ask about sample size. Though heterozygous TPMT mutations are fairly common and represent about 11% of the population, homozygous mutations are not common, occurring in only 0.33% of the population. To validate the accuracy of various phenotype and genotype technologies, a large sample size is necessary. For this reason, in the modified version of the tool used in this study, a question on sample size was added.

Two additional questions were added that were specific to the accuracy of the TPMT phenotype technologies and that related to the reporting of recent blood transfusions and the concurrent use of medications in study subjects. If these confounding factors are not made transparent in the publication, then the results which are reported need to be interpreted with caution.

Data from the quality assessment were compiled using a Microsoft Excel spreadsheet. Questions were answered with one of three responses: yes, no or cannot tell. Cannot tell was used when the article did not provide enough detail to answer the question definitively with a yes or no.

2.6 Economic Evaluation

2.6.1 Study Design

A cost-effectiveness analysis (CEA) was carried out from the health care system perspective to compare three testing strategies for 6-MP dosing: genotype-based, phenotype-based and no testing (standard weight-based dosing). This analysis was performed on a hypothetical cohort of pediatric patients with ALL and receiving 6-MP for the maintenance phase of therapy. The time horizon was set at three months, as any intolerance to 6-MP due to TPMT enzyme deficiency would be evident at or before three months (J. Hand, personal communication, September 9, 2008 (Appendix 1)). A decision analytic model was used to carry out the analysis (Figure 1).

Figure 1: Decision Tree
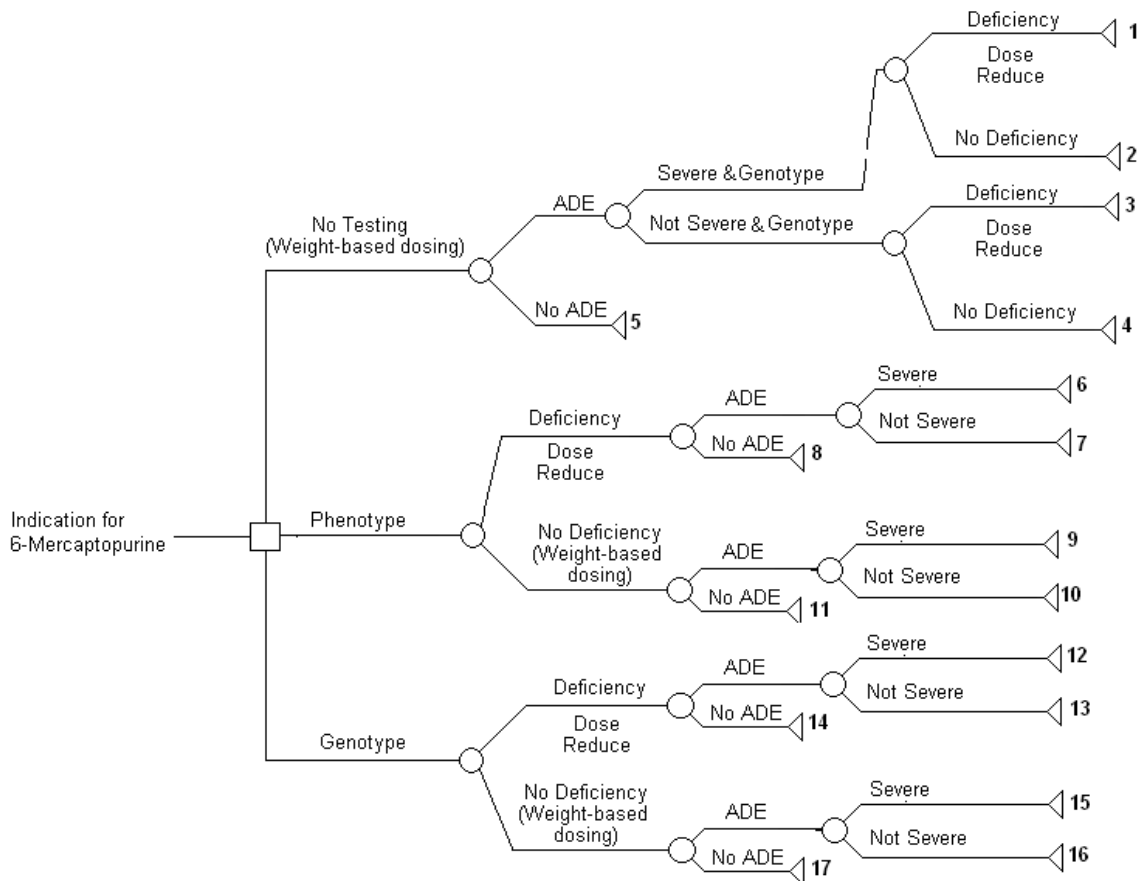
Note: Numbers 1-17 correspond to each terminal node

2.6.2 Decision Analysis

The decision tree consists of three arms. The first represents the standard of care where 6-MP is dosed based on patient weight in the absence of TPMT mutation or activity knowledge. In the event that an adverse drug event interrupted therapy by two weeks, a TPMT genotype test is conducted. When an adverse event causes a minor interruption in therapy (less than two weeks), the COG guidelines indicate that a genotype test is not necessary.[4] These cases were included in the 'No ADE' branch of the decision tree.

For the purpose of this analysis, myelosuppression was the only adverse drug event considered as it is the adverse event that causes the most concern.[43] There is also a lack of evidence around the parameters of other adverse events. The risk of secondary malignancy, though serious, would not factor into this analysis because of the short time horizon; they are also suspected to be quite rare, so the potential impact on this analysis would be minimal.

Myelosuppression puts the patient at an increased risk of contracting an infection, which is what makes this event serious. When myelosuppression is combined with fever (febrile neutropenia), infection is suspected. The patient would then be immediately admitted to hospital for observation and treatment,[4] which not only comes at a high cost but impacts the patients' quality of life.

The second and third arms of the tree represent the two forms of TPMT testing. Since dosage adjustments are not recommended for heterozygous patients until after they experience an adverse event, only patients with homozygous mutations will follow the 'deficiency' branch. Patients with heterozygous mutations will follow the 'no deficiency' branch. Other than 6-MP dosage adjustments for patients with homozygous mutations, the second and third arms follow the same clinical pathway as the first arm following the COG protocol guidelines, though they have different branch probabilities based on the sensitivity and specificity of the TPMT tests.

Each branch where a deficiency is detected through a phenotype or genotype test is representative of both true and false positives. This model does not account for the potential adverse drug events (e.g. disease relapse) of patients who are under-dosed due to a false positive test result.

Initially, a base case decision analysis with point estimates was performed. Subsequently one-way sensitivity analyses and a multi-way probabilistic sensitivity analysis (PSA) with Monte Carlo simulations were performed.

2.6.3 Base Case Scenario

A base case scenario was created for pediatric patients with ALL. This scenario included the starting dose of 6-MP, dosage adjustments, laboratory monitoring, incidence of myelosuppression, rate of hospital admission, risk of death and accuracy of the TPMT tests. Previously reported values from our systematic review and values from other published literature were used to assign the decision probabilities to each chance node. The assumption was made that there is 100% compliance by physicians with dose adjusting according to TPMT activity/genotype. The base case decision analysis was conducted using Excel^{TM}.

2.6.4 Probabilities

For the weight-based dosing arm, the general risk for myelosuppression is required. The incidence of serious cases of myelosuppression in children with ALL being treated with 6-MP is not well documented in the literature. Frequency of myelosuppression in other disease states treated with thiopurines has ranged from 1.4% to 9%.[44, 45] A European study evaluating the same question in the pediatric ALL population, as we have outlined here, assumed a frequency of 3% based on the above evidence.[43] For this evaluation we assumed the same frequency, 3%, for the overall risk of myelosuppression. This overall risk was adjusted in the phenotyping and

41

genotyping arms to reflect the reduction in risk by tailoring the 6-MP dosage. The value of 3% was used for the risk of myelosuppression in the weight-based dosing arm.

Not all cases of myelosuppression are due to a mutation in the gene coding for the TPMT enzyme. Therefore, screening for TPMT status with either the phenotype or genotype test will not prevent all myelosuppression cases. Marra, et al. (2002)[45] made the assumption that 50% of myelosuppression cases can be prevented with testing followed by dose reduction as appropriate. There is more literature to suggest that approximately one third (29-32%) of myelosuppression cases can be prevented by tailoring doses to enzyme activity.[44, 46-48] A conservative assumption that 30% of myelosuppression cases could be prevented if TPMT deficient patients were appropriately dose adjusted was used in the base case analysis.

For the phenotype and genotype arms, a breakdown of myelosuppression risk was necessary for the TPMT deficiency and no TPMT deficiency branches. For TPMT homozygous patients who followed the deficiency branch, the 6-MP dose would have been dose adjusted, thereby reducing their risk of myelosuppression. Since we assumed that 30% of myelosuppression cases were due to TPMT deficiency, the risk would be reduced by 30%, resulting in a risk of 2.10% (as described above, 3% was the risk assumed for weight-based dosing). For patients who followed the no deficiency branch (a combination of heterozygous and wild type patients), the risk of myelosuppression was calculated by solving for unknown variables. This was done using a hypothetical cohort of 300 patients, where one patient would have a homozygous TPMT mutation, thirty-three would have a heterozygous TPMT mutation and 266 would be TPMT wild type.

$$OverallRisk = \frac{1(X) + 33(Y) + 266(Z)}{300}$$

Where:

X = Risk of myelosuppression for homozygous patients

Y = Risk of myelosuppression for heterozygous patients

Z = Risk of myelosuppression for wild-type patients

It was assumed that 100% of patients with a TPMT homozygous mutation (i.e. X = 100%), who are not appropriately dose reduced, would experience myelosuppression.[49] Since it cannot easily be assumed what percentage of heterozygous patients would experience myelosuppression, the equation would be solved for 'Y+Z'.

$$3\% = \frac{1(100\%) + 299(Y + Z)}{300}$$

Y+Z = 2.68%

Patients who followed the 'no deficiency' arm were assigned a 2.68% risk of myelosuppression. All risks for myelosuppression were then weighted by incorporating the sensitivity and specificity of the phenotype and genotype tests. The weighted values for true positives and false positives were added together to give the risk of myelosuppression for patients with TPMT deficiency. Weighted values for the true negatives and false negatives were added together to give the risk of myelosuppression for the no TPMT deficiency branch. Values derived from the systematic review (Section 3.1.2.5) were used in these calculations (Table 6). Values for the base case analysis were chosen based on the midpoint of values found derived from the literature during the systematic review. As neither of the studies included in the systematic review were of high methodological quality (see Section 4.1.3), studies were not weighted by quality when selecting the values for the base case analysis. The genotype test was estimated to have a sensitivity of 77.28% and specificity of 97.15%. The phenotype test was found to have 95.83% sensitivity and 92.25% specificity.

Only cases of myelosuppression that were combined with fever (febrile neutropenia) were assumed to require admission to hospital. For the purpose of this analysis, febrile neutropenia is

considered 'severe' and myelosuppression without fever is considered 'not severe'. The

designation of 'severe' and 'not severe' is not a reflection of degree of myelosuppression, but

rather reflects the need for hospitalization. Approximately 10-20% of all cases of

myelosuppression are admitted to hospital (J. Hand, personal communication, September 9, 2008

(Appendix 1)).[43, 49] The midpoint of 15% was used in the base case analysis.

Table 6: Probabilities and Sources Used in the Base Case Analysis

Model Parameters	Base Case Value	References
Probabilities		
Incidence of TPMT wild type allele	89%	Weinshilboum, 1980
Incidence of heterozygous TPMT mutation	11%	
Incidence of homozygous TPMT Mutation	0.33%	
Population incidence of myelosuppression with mercaptopurine therapy	3.00%	van den Akker-van Marle, 2006
Incidence of myelosuppression for heterozygous and wild-type patients who are not dose reduced	2.68%	Calculated
Incidence of myelosuppression for homozygous patients who are not dose reduced	100%	Winter, 2004
Proportion of myelosuppression due to TPMT deficiency	30%	Sanderson, 2004; Colombel, 2000; Ansari, 2002; Schwab, 2002
Proportion of myelosuppression cases needing inpatient services	15%	Winter, 2004; van den Akker Marle, 2006; Personal Communication, Dr. J. Hand, Sept. 9th, 2008
Sensitivity of genotype test	77.28%	Refer to section 3.1.2.5
Specificity of genotype test	97.15%	
Sensitivity of phenotype test	95.83%	
Specificity of phenotype test	92.25%	

2.6.5 Branch Probability Calculations

Branch probabilities were calculated based on the following formulae. The probabilities of 'severe' versus 'not severe' were taken directly from Table 6 (proportion of myelosuppression cases needing inpatient services), and no further calculations were necessary.

2.6.5.1 Calculations for Weight-based Dosing Strategy

1.   *p[ADE] = iADEa*

        p[ADE] = Probability of experiencing an ADE in the general population

        iADEa = Incidence of ADE in general population

2.   *p[noADE] = 1 – p[iADEa]*

        p[ADE] = Probability of not experiencing an ADE in the general population

3.   *p[Deficiency | Severe | ADE] = ((Ho + He)*SeG)+((1-(Ho+He))*(1-SpG))*

        p[Deficiency | Severe | ADE] = Probability of TPMT deficiency in a patient that had an ADE requiring hospitalization

        Ho = Incidence of TPMT homozygous deficiency

        He = Incidence of TPMT heterozygous deficiency

        SeG = Sensitivity of genotype test

        SpG =Specificity of genotype test

4.   *p[no Deficiency | Severe | ADE] = 1 – p[Deficiency | Severe | ADE]*

p[no Deficiency | Severe | ADE] = Probability of no deficiency in a patient that had an ADE that required hospitalization

5. *p[Deficiency | not Severe | ADE] = ((Ho + He)*SeG)+((1-(Ho+He))*(1-SpG))*

p[Deficiency | not Severe | ADE] = Probability of deficiency in a patient with an ADE that did not require hospitalization. This calculation takes into account both the true positives and the false positives by incorporating sensitivity and specificity values.

Ho = Incidence of TPMT homozygous deficiency

He = Incidence of TPMT heterozygous deficiency

SeG = Sensitivity of genotype test

SpG =Specificity of genotype test

6. *p[no Deficiency | not Severe | ADE] = 1 - p[Deficiency | not Severe | ADE]*

p[no Deficiency | not Severe | ADE] = Probability of no deficiency in a patient with an ADE that did not require hospitalization

2.6.5.2 Calculations for Phenotype and Genotype Dosing Strategies

1. *p[Deficiency] = ((Ho*Se)+((1-Ho)*(1-Sp)))*

p[Deficiency] = Probability of TPMT homozygous deficiency. This calculation takes into account both the true positives and the false positives by incorporating sensitivity and specificity.

Ho = Incidence of TPMT homozygous deficiency

Se = Sensitivity of test (phenotype or genotype depending on tree branch)

Sp = Specificity of test (phenotype or genotype depending on tree branch)

2.   *p[no Deficiency] = 1 - p[Deficiency]*

p[no deficiency] = Probability of TPMT wild type or heterozygous mutation, which consists of all true negatives and false negatives.

3.   *p[ADE | Deficiency] =*

$$\left(iADEa * (1 - pTPMT)\right)\frac{Se * Ho}{(Se * Ho) + ((1 - Sp) * (1 - Ho))} + (iADEb)\frac{(1 - Sp) * (1 - Ho)}{((1 - Sp) * (1 - Ho)) + (Se * Ho)}$$

p[ADE | Deficiency] = Probability of an ADE in a TPMT deficient patient. This calculation is a weighted average of the probability of myelosuppression in the true positives and the false positives determined by incorporating sensitivity and specificity.

iADEa = Incidence of an ADE in general population

pTPMT = Proportion of ADEs attributable to TPMT deficiency

iADEb = Incidence of ADE in TPMT heterozygous and wild type patients

Ho = Incidence of homozygous TPMT mutation

Se = Sensitivity of test (phenotype or genotype depending on tree branch)

Sp = Specificity of test (phenotype or genotype depending on tree branch)

4.   *p[no ADE | Deficiency] = 1 - p[ADE | Deficiency]*

p[no ADE | Deficiency] = Probability of no ADE in a TPMT deficient patient.

5. *p[ADE | No Deficiency] =*

$$\left(iADEc\right)\frac{(1-Se)*Ho}{((1-Se)*Ho)+(Sp*(1-Ho))}+(iADEb)\frac{Sp*(1-Ho)}{(Sp*(1-Ho))+((1-Se)*Ho)}$$

p[ADE | No Deficiency] = Probability of ADE in patient who is TPMT heterozygous and wild type. This calculation is a weighted average of the probability of myelosuppression in the true negatives and the false negatives determined by incorporating sensitivity and specificity.

iADEc = Incidence of ADE in TPMT homozygous patients who are not dose reduced (false negatives).

iADEb = Incidence of ADE in TPMT heterozygous and wild type patients (true negatives).

Ho = Incidence of homozygous TPMT mutation

Se = Sensitivity of test (phenotype or genotype depending on tree branch)

Sp = Specificity of test (phenotype or genotype depending on tree branch)

6. *p[no ADE | No Deficiency] = 1 - p[ADE | No Deficiency]*

p[no ADE | No Deficiency] = Probability of no ADE in TPMT heterozygous and wild type patients

2.6.6 Outcomes

A systematic review of economic evidence of TPMT induced side effects concluded that the QALYs lost from experiencing neutropenia are small, and due to insufficient evidence it is perhaps premature to try and calculate a QALY difference between groups.[50] Accurate survival data are not available for pediatric patients with ALL treated with 6-MP. Assumptions made by others have

been that for every 1,000 patients screened there would be one death averted;[43, 49] however, these assumptions were not specific to the pediatric ALL population. Another study looked at longitudinal data to determine incidence of mortality of pediatric oncology patients admitted to hospital with febrile neutropenia over a seven-year period.[51] When broken down by cancer type, ALL patients were found to have a 2.9% risk of death during admission for febrile neutropenia, and this value was used in the base case analysis.[51] No published data were found that presented the risk of death in patients who do not experience febrile neutropenia. Based on a three-month time horizon, we assumed that the risk of death for those patients would be rare and used 0% for our base case analysis.

The survival outcome used in this analysis was life-months. Since it was not clear from the literature at what point after therapy a patient might die from febrile neutropenia, it was assumed any deaths would occur two weeks after starting therapy to allow the patient enough time after therapy was initiated to experience neutropenia. For all treatment branches where patients were admitted to hospital, a terminal node value of 2.9275 life-months was assigned. This value was determined by taking the risk of death (2.9%) and multiplying it by the number of months remaining in the three month time horizon (2.5 months) and then subtracting the result from three months. For all other treatment branches, where patients were not hospitalized, the risk of death was 0%, so three months was the value for survival assigned at the terminal node to reflect the full time horizon of the analysis.

2.6.7 Costs and Resource Utilization

Costs were calculated according to a health care system perspective and only included direct medical costs of the patients and health care system. Particular items of interest included acquisition cost for 6-MP, laboratory costs, regular complete blood cell (CBC) and liver function

tests (LFT), enzymatic assay and genotype test fees. Hospitalization costs for the treatment of febrile neutropenia were estimated. Costs were calculated in 2008 Canadian dollars (Table 8).

The starting dose for 6-MP is 75 mg/m^2 per day. This dose was assigned for all patients in the weight-based dosing arm. If patients experienced febrile neutropenia and they were found to be heterozygous for the TPMT allele, the dose would be reduced to 45 mg/m^2 daily (60% of starting dose) once therapy has been restarted. If patients were found to be homozygous, the dose would be reduced to 10 mg/m^2 per day, three days a week (<10% of starting dose) once therapy has been restarted.[4] Patients in the phenotype and genotype arm would receive 75 mg/m^2 daily if they are wild-type or have a heterozygous mutation, and 10 mg/m^2 per day, three days a week, if found to have a homozygous mutation. Patients who have a heterozygous mutation are not dose reduced initially because the COG guidelines suggest that only a third of these patients will not tolerate the full dose. If a patient with a heterozygous mutation experiences neutropenia, that delays therapy for two or more weeks, after which the dose would be adjusted to 45 mg/m^2 daily.[4]

The acquisition cost of 6-MP was based on the cheapest generic alternative. In the province of Newfoundland and Labrador it is a part of the pharmacists' standards of practice to always dispense the cheapest equivalent alternative of any medication. The price used in this analysis was that of the Novopharm product which was priced at $3.67 per 50 mg tablet (Novopharm, personal communication, February 18, 2008 (Appendix 1)). In the base case analysis, the cost per dose was calculated for a three-year old child who weighs 15 kg and is three feet tall (equaling a body surface area of 0.62 m^2). A three year old child was selected as this is the age in which ALL is most common.[52] A normal 6-MP starting dosage regimen of 75 mg/m^2 would equal 46.5 mg. For these patients, the cost per dose would be $3.67. For TPMT heterozygous patients who require dose reductions of 40%, the dose would be reduced to 27.9 mg, at a cost of $2.10 per dose. For a TPMT homozygous patient, whose dose is reduced to 10 mg/m^2, three days a week, the starting dose would be 6.2 mg, at a cost of $0.61 per dose (Table 7).

Table 7: Cost per Dose of 6-Mercaptopurine

Regimen	Patient Dose	Contents of suspension	Cost of suspension	# Doses in suspension	Cost per Dose
75 mg/m^2 daily	46.5 mg	50 mg x 14 tablets = 700 mg	$3.67 x 14 tablets = $51.38	15 (14 will be used)	$3.67
45 mg/m^2 daily	27.9 mg	50 mg x 8 tablets = 400 mg	$3.67 x 8 tablets = $29.36	14 (14 will be used)	$2.10
10 mg/m^2 three days per week	6.2 mg (three days per week)	50 mg x 1 tablets = 50 mg	$3.67 x 1 tablets = $3.67	8 (6 will be used)	$0.61

Note: Doses are calculated based on a three-year old child who is 15 kg and three feet tall (body surface area of 0.62 m^2); Each tablet of 6-MP contains 50 mg and costs $3.67.

A professional fee (dispensing fee) is generally added to all prescription prices to account for pharmacy overhead costs and professional services. This is a flat fee applied to every new and refill prescription and in Ontario most pharmacies charge $11.99.   A professional fee, referred to as the compounding fee, is applied to medications that require special preparation. While in some regions pharmacists may compound the drug for parents, in others parents are sent home with instructions on how to mix pills with water, and some children are able to consume the pills with no preparation . It was assumed in this analysis that parents would compound the drug themselves if necessary and therefore only dispensing fees and not compounding fees were included. A dispensing fee was applied for every two weeks of treatment, for those weeks that the patient receives 6-MP as an outpatient.

The cost of TPMT genotype tests varies depending on the institution. Most institutions have to send their samples to an outside lab. At the Janeway Children's Hospital in St. John's, NL, samples are sent to the Mayo Clinic in the United States at a cost of $380.50 US.[53] Some

institutions who have developed their own tests can offer this test on site for a fraction of the cost; however exact prices were not available. These tests are still in the development stages and competitive pricing is necessary to ensure that samples are referred to particular labs (P. Gordon, personal communication, January, 2008 (Appendix 1)). In the base case analysis the US Mayo Clinic price was used. Given the volatility in currency conversion rates, a conservative rate of 1.2098 (Central bank rate January 28th, 2009) was used resulting in a price of $459.63 CDN. TPMT phenotype tests are available at most hospital laboratories however commercial pricing is not available. The price of the TPMT phenotype test at the Mayo Clinic is $333.40 US ($402.912 CDN as of January 28th, 2009) and this value was used in the base case analysis.

Physician fees, for both inpatient and outpatient services, were obtained from the Ontario Health Insurance Program (OHIP) fee schedule.[54] Pediatric hematologists are reimbursed for outpatient medical specific re-assessments at a rate of $50.50 (A264). In the absence of treatment obstacles, patients would visit the outpatient clinic monthly (J. Hand, personal communication, September 9, 2008 (Appendix 1)). Inpatient consultations are reimbursed at $147.80 (C265) for the first visit on day one of admission, and $55.45 (C122, C123, C124) for daily subsequent visits by the most responsible physician.[54]

Laboratory monitoring of CBC and LFTs were assumed to be carried out monthly except in the presence of myelosuppression when the tests are carried out every seven days. Costing for laboratory tests was also retrieved from the OHIP fee schedule.[55] Laboratory services are reimbursed based on the number of Laboratory Medicine Service (LMS) units each test uses. Each LMS unit is reimbursed at a rate of $0.517. A CBC includes white blood cell (WBC) Differential (L372), Platelet Count (L396), RBC Count (L397), WBC Count (L399), Hematocrit (L417), and Hemoglobin (L418). When more than one of these counts is needed the code L393 is used which provides the complete profile at a cost of $8.27 (16 LMS Units). An LFT profile generally includes alanine aminotransferase (L223), aspartate aminotransferase (L222), alkaline

phosphatase (L191), gamma-glutamyl transpeptidase (L107), prothrombin time (L445) and total bilirubin (L030). These tests are costed separately and come to a total of $19.13 (37 LMS units). Laboratory costs for this analysis were calculated based on one CBC and LFT per month in patients not experiencing ADEs.

The cost for an inpatient stay was obtained from the Ontario Case Costing Initiative (OCCI),[56] and was calculated based on International Classification of Disease Codes (ICD10 Codes) D70 (Agranulocytosis, including neutropenia) for patients aged 0-17 years. It is possible that not all of these patients had neutropenia secondary to 6-MP therapy for ALL, but it is not expected that inpatient costs for neutropenia cases caused by other illnesses would be markedly different. The most recent year of data available, which also had the highest number of cases for any year in the database, was 2007-2008. The average case cost for that year was $9,641 in 2008 Canadian dollars, with an average length of stay of 5.9 days. The total cost includes both direct (nursing, diagnostic imaging, pharmacy and laboratory services) and indirect costs (overhead) but does not include physician services. This cost and length of stay was used in the base case analysis.

If the patient was to experience neutropenia it was assumed that they would be taken off therapy for two weeks. This estimate was based on a study which looked at the average number of weeks of missed 6-MP therapy based on TPMT status.[57] It was found that between 7-24% of the intended weeks of therapy were missed.[57] Applying this to a three-month time horizon gives a range of 0.8 to 2.9 weeks of missed therapy. Two weeks is approximately the middle of this range and was therefore used for the base case analysis.

Table 8: Cost and Resource Utilization Items and Sources

Item	Base-Case Value	References
Phenotype Test	$402.91	Mayo Clinic, 2009
Genotype Test	$459.63	
CBC Test	$8.272	OHIP* Fee Schedule, 2008 (L393)
LFT Test	$19.129	OHIP Fee Schedule, 2008 (L223, L222, L191, L030)
Physicians fee for service – Out-Patient	$50.50	OHIP Fee Schedule, 2008 (A264)
Physician fee for service – Inpatient (1st visit)	$147.80	OHIP Fee Schedule, 2008 (C265)
Physician fee for service – Inpatient (follow-up)	$55.45	OHIP Fee Schedule, 2008 (C122, C123, C124)
6-Mercaptopurine	$3.67 / 50 mg tablet	Novopharm, Personal communication, February 18, 2009
Dispensing fee	$11.99	
Average cost per inpatient stay	$9,641	OCCI**, 2008
Average length of hospital stay for febrile neutropenia	5.9 days	
Average weeks of missed therapy due to myelosuppression	2 weeks	Relling, 1999

*OHIP: Ontario Health Insurance Plan, ** OCCI: Ontario Case Costing Initiative

2.6.7.1 Cost Calculations

Please refer to Table 9 for a summary of all of the applicable cost items used in each branch. The number in each cell represents the frequency of the cost item incorporated into each tree branch. The total cost at the terminal node of the branch was determined by multiplying the price by the volume of use of each of the applicable cost items and summing all the cost items.

Table 9: Health Care Resource Use in Decision Analysis

Tree Code	Phenotype Test	Genotype Test	CBC & LFT Test	Inpatient Costs	Physician Fees			6-MP - full dose*	6-MP - 60% dose*	6-MP - ~10% dose*	Compounding Fee
					Outpatient	Inpatient 1st consult	Inpatient follow-up				
Weight-based Dosing Arm											
1	0	1	4	1	4	1	4.9	14	0.9709(56)	0.0291(24)	5
2	0	1	4	1	4	1	4.9	70	0	0	5
3	0	1	5	0	5	0	0	14	0.9709(56)	0.0291(24)	5
4	0	1	5	0	5	0	0	70	0	0	5
5	0	0	3	0	3	0	0	84	0	0	6
Phenotype Arm											
6	1	0	4	1	4	1	4.9	0	0	30	5
7	1	0	5	0	5	0	0	0	0	30	5
8	1	0	3	0	3	0	0	0	0	36	6
9	1	0	4	1	4	1	4.9	14+0.8896 *(56)	0.1104(56)	0	5
10	1	0	5	0	5	0	0	14+0.8896 *(56)	0.1104(56)	0	5
11	1	0	3	0	3	0	0	84	0	0	6
Genotype Arm											
12	0	1	4	1	4	1	4.9	0	0	30	5
13	0	1	5	0	5	0	0	0	0	30	5
14	0	1	3	0	3	0	0	0	0	36	6
15	0	1	4	1	4	1	4.9	14+0.8896 *(56)	0.1104(56)	0	5
16	0	1	5	0	5	0	0	14+0.8896 *(56)	0.1104(56)	0	5
17	0	1	3	0	3	0	0	84	0	0	6

*Refer to Table 10 for more detail.

The assumptions regarding health resource use listed below were followed. Phenotype and genotype tests were only administered once in each of the applicable branches. CBC and LTF tests were administered three, four or five times depending on the presence of an ADE and its severity. Patients who did not experience an ADE received the tests on three occasions, once for each month of therapy. Patients who experienced an ADE also had a set of tests for each month but had an additional set of tests for each of the two weeks of missed therapy. Patients who were not hospitalized had a total of five sets of tests. Patients who required hospitalization also had five sets of tests, however one test was assumed to be done during the hospital stay and therefore covered under the inpatient costs, therefore only the costs for four sets of tests were added for these patients.

Outpatient physician fees followed the same pattern as the CBC and LFT tests. Patients who never experienced an ADE had three physician visits (A264). Patients who experienced an ADE required an additional visit for each week of missed therapy. Patients who were not hospitalized had a total of five outpatient physician visits. Patients requiring hospitalization saw the physician during the hospital stay (reflected in a different cost category), and one additional visit as an outpatient, totaling four outpatient physician visits.

Inpatient costs were only applied to those patients requiring hospitalization. As previously described, patients were assumed to be hospitalized for an average of 5.9 days (SD 6.2). A physician fee for the first inpatient consultation (C265) was applied for the first day of stay. Physician fees for inpatient visits after the initial consultation were multiplied by (length of stay-1) to reflect a visit per day for the remainder of the stay (C122, C123, C124).

Table 10 outlines the process used to determine which of the three 6-MP doses (full dose, 60% dose or <10% dose) would be used and for how long in each of the tree branches. In the weight-based dosing arm, branches one and three represent patients who experience an ADE and who

were found by genotype testing to carry a TPMT deficiency, therefore requiring dosage adjustment. In a group of patients who all have TPMT deficiencies, 2.91% would be homozygous and 97.09% would be heterozygous. In the phenotype and genotype arms, branches 9, 10, 15 and 16 represent a combination of heterozygous and wild-type patients who experienced an ADE. The proportion of these patients who are heterozygous (11.04%) will require a dosage adjustment, while the wild-type patients will not (88.96%).

Table 10: Determination of 6-Mercaptopurine Dose for Each Tree Branch

Tree Code	6-Mercaptopurine Dose Description
Weight-Based Dosing Arm	
1	2 weeks full dose + 2 weeks no therapy + 8 weeks (2.91% at 10% dose, 97.09% at 60% dose)*
2	2 weeks full dose + 2 weeks no therapy + 8 weeks full dose
3	2 weeks full dose + 2 weeks no therapy + 8 weeks (2.91% at 10% dose, 97.09% at 60% dose)*
4	2 weeks full dose + 2 weeks no therapy + 8 weeks full dose
5	12 weeks full dose
Phenotype Arm	
6	2 weeks 10% dose + 2 weeks no therapy + 8 weeks 10% dose
7	2 weeks 10% dose + 2 weeks no therapy + 8 weeks of 10% dose
8	12 weeks 10% dose
9	2 weeks full dose + 2 weeks no therapy + 8 weeks (11.04% at 60% dose, 88.96% at full dose)**
10	2 weeks full dose + 2 weeks no therapy + 8 weeks (11.04% at 60% dose, 88.96% at full dose)**
11	12 weeks full dose
Genotype Arm	
12	2 weeks 10% dose + 2 weeks no therapy + 8 weeks 10% dose
13	2 weeks 10% dose + 2 weeks no therapy + 8 weeks 10% dose
14	12 weeks 10% dose
15	2 weeks full dose + 2 weeks no therapy + 8 weeks (11.04% at 60% dose, 88.96% at full dose)**
16	2 weeks full dose + 2 weeks no therapy + 8 weeks (11.04% at 60% dose, 88.96% at full dose)**
17	12 weeks full dose

* In a group of patients who all have TPMT deficiencies, 2.91% would be homozygous and 97.09% would be heterozygous.
** These branches represent a group of patients with either TPMT wild type (11.04%) or with a heterozygous mutation (88.96%).

Compounding fees were multiplied by either five or six depending on how many refills were expected for each of the tree branches. Patients who never experienced an ADE were assumed to not miss any weeks of therapy and were allotted six refills, reflecting one refill every two weeks for three months of therapy. Patients who experienced an ADE were assumed to miss two weeks of therapy and were therefore allotted five refills.

2.6.7.2 Univariate Sensitivity Analyses

After performing the base case analysis, one-way sensitivity analyses were performed altering the price of the genotyping and phenotyping tests according to their measure of uncertainty (Table 11), as these represent the technology of interest costs and also had a high degree of uncertainty. Additional one-way sensitivity analyses were conducted to examine the effects of varying the cost of the hospital stay, the proportion of myelosuppression cases requiring hospitalization, and phenotype and genotype sensitivity and specificity.

2.6.7.3 Probabilistic Sensitivity Analysis

A probabilistic sensitivity analysis was performed where all uncertain variables were varied simultaneously along specified ranges in a Monte Carlo simulation with 10,000 replications, using TreeAge[TM] software. The variable ranges and distributions used in the PSA are presented in Table 11.

The incidence of myelosuppression with 6-MP was varied between 1.4% and 9%, based on values reported in the literature.[44, 45] The percentage of myelosuppression that can be attributed to deficiencies in the TPMT enzyme was varied between 20% and 50% based on the assumptions made by Marra et al (2002).[45] These two variables were assigned a triangular distribution as few data were available to inform the choice of distribution. The percentage of hospital admissions for myelosuppression was varied between 10% and 20% based on expert

opinion (J. Hand, personal communication, September 9, 2008 (Appendix 1)).[43, 49] As there were no data on the distribution of this variable it was assigned a uniform distribution.

The accuracy of the TPMT genotype and phenotype tests was varied to reflect the results found in the systematic review. For the genotype test, values ranged from 54.55% to 100% for sensitivity[28, 58] and 94.3% to 100% for specificity.[28, 59] For the phenotype test, values ranged from 91.67% to 100% for sensitivity and 86.21% to 98.28% for specificity.[60, 61] As these studies were all considered to be of similar quality, these variables were assigned a uniform distribution.

Costs for the genotype and phenotype tests were varied based on possible costs incurred by different institutions. Based on expert opinion (P. Rahman, personal communication, 2008 (Appendix 1)), laboratories who have the ability to offer their own phenotype and genotype testing services could do so for approximately $100 CDN each. Internet searches identified a UK lab that provided the test for €50 (approximately CDN$74, based on a conversion rate of 1Euro=1.472CDN as of Feb. 9, 2010)[62] to hospitals in Europe and £29 (CDN$49 based on a conversion rate of 1GBP=1.675CDN as of Feb. 9, 2010)[62] for UK hospital laboratories.[63] The Mayo Clinic value was the highest price identified.[53] Values within this range were used in sensitivity analysis. A similar approach was used to determine costs and ranges for the phenotyping test. The distributions of these variables were set as triangular given the small amount of data available.

Rather than varying the average inpatient case cost along the range provided by OCCI, which included long-stay outliers, the OCCI average cost per day of inpatient stay was calculated by dividing the average case cost by the average length of stay, yielding $1,634 per day. This value was held constant and multiplied by the length of stay, which was allowed to vary according to the OCCI range (average length of stay 5.9 days, SD 6.2, distribution set as normal).

The estimated risk of death from myelosuppression has ranged from 2-9%;[51] this variable was

assigned a triangular distribution in the PSA with these values as the upper and lower bounds.

Table 11: Variables and Ranges Used in the Univariate and Probabilistic Sensitivity Analyses

Model Parameters	Base Case Value	Lower Range	Upper Range	Distribution	References
Probabilities					
Incidence of myelosuppression with 6-MP in the general population	3%	1.4%	9.0%	Triangle	Van den Akker-van Marle, 2006; Sanderson 2004
Incidence if myelosuppression with 6-MP in wild-type and heterozygous patients	2.68%	1.07%	8.70%	Triangle	Calculated
Proportion of myelosuppression due to TPMT deficiency	30%	20%	50%	Triangle	Van den Akker-van Marle, 2006; Marra, 2002
Proportion of myelosuppression cases needing inpatient services	15%	10%	20%	Uniform	Expert opinion Dr. J. Hand, 2008; Winter, 2004; Van den Akker-van Marle, 2006
Sensitivity of genotype test	77.28%	54.55%	100.00%	Uniform	Spire-Vayron de la Moureyre, 1998; Nasedkina, 2005
Specificity of genotype test	97.15%	94.30%	100.00%	Uniform	Spire-Vayron de la Moureyre, 1998; Nasedkina, 2005
Sensitivity of phenotype test	95.83%	91.67%	100.00%	Uniform	Alves, 2001; Kham, 2008
Specificity of phenotype test	92.25%	86.21%	98.28%	Uniform	Anglicheau, 2002; Kham, 2008
Costs					
Phenotype Test	$402.91	$100	$500	Triangle	Mayo Clinic, 2008
Genotype Test	$459.63	$100	$500	Triangle	Mayo Clinic, 2008
Per diem inpatient cost	$1634	-	-	Fixed	OCCI, 2008
Length of stay	5.9	-	SD 6.2	Normal	OCCI, 2008
Outcomes					
Risk of death from myelosuppression	2.9%	2	9	Triangle	Basu, 2005

All other costs (physician fees, dispensing fees, lab tests, 6-MP acquisition costs, dose per day) were held constant as these costs are stable and highly consistent across provinces.

2.6.7.4 Threshold Analysis

A threshold analysis was conducted to determine the costs of the phenotype and genotype tests that would result in zero incremental costs compared to the weight-based dosing arm.

2.6.8 Presentation of Findings

Findings of the systematic review of phenotype and genotype test accuracy were tabulated and summarized. The results of the cost effectiveness analysis were organized by treatment arm and broken down by the expected outcomes and expected sub-costs at each terminal node. Incremental costs were calculated comparing each of the testing arms to the weight-based dosing arm as well as comparing the incremental cost between each testing arm. The results of the one-way sensitivity analysis were summarized in tornado diagrams depicting the extremes of values found in the literature. The PSA results were tabulated and depicted in a scatter plot.

3 RESULTS

The results are presented in two main sections. The first section outlines the findings of the systematic review of the literature focusing on the accuracy of TPMT testing technologies. The second section describes the results of the CEA.

3.1 Literature Review

3.1.1 Quantity of Research Available

Though there is an abundance of literature on the pharmacogenetic applications of testing for TPMT deficiencies, few studies actually looked at the accuracy of the array of existing technologies. A thorough literature search using PubMed, Embase, the Cochrane Database the FDA Centre for Drug Evaluation and Research and a hand search of references from included studies, resulted in a total of 26 articles that potentially met the inclusion criteria.

Of these, one study was published in Chinese and therefore was excluded,[64] six did not contain enough detail on the assignments of phenotypes and/or genotypes,[65-70] one evaluated the accuracy of a test to detect adverse drug events[71] and one compared the tests in various animal species.[72] All of these were excluded. Nine studies were finally included in this review (Figure 2).

Figure 2: Literature Search Flow Diagram

	Pubmed	Embase	Cochrane Database	FDA	References
Returned Results :	71	102	20	0	10

Articles Meeting Inclusion Criteria: 26

Excluded Articles:
- 1 Written in Chinese
- 6 Insufficient detail
- 1 Not comparing accuracy of tests
- 1 Compared tests among species

Included Articles: 17
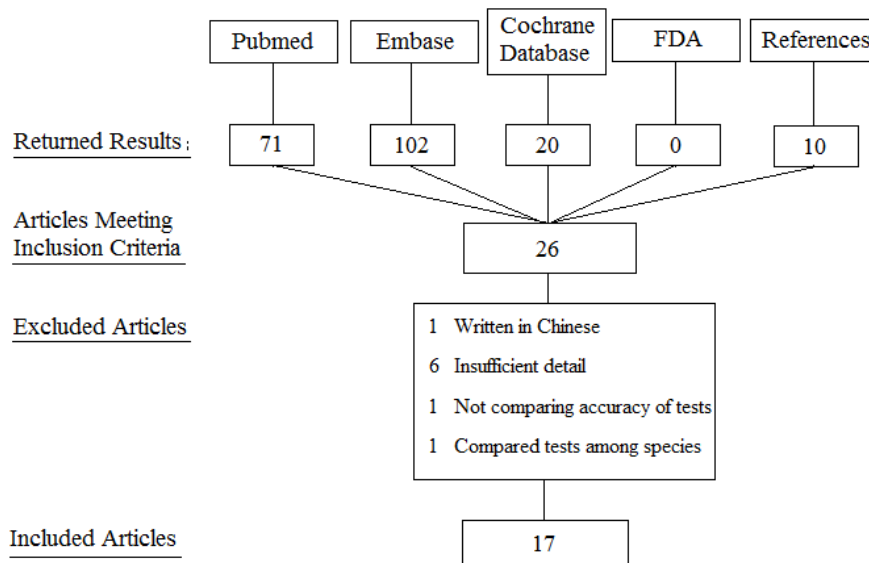
3.1.2 Trial Characteristics

Seventeen studies were identified that reported a comparative analysis between either the TPMT phenotype or genotype test and/or a gold standard.[19, 23, 27-29, 35, 58-61, 73-79]

3.1.2.1 Inclusion and Exclusion Criteria

Limited detail was provided in the papers with respect to inclusion and exclusion criteria of each of the studies (Table 12). Inclusion criteria were limited to the ethnic origin and health status of the study subjects. Only three studies considered confounding factors such as concurrent medication and blood transfusions in the exclusion criteria.[28, 73, 75]

Table 12: Inclusion and Exclusion Criteria for Included Studies

Study	Inclusion Criteria	Exclusion Criteria
Kham, 2008	Healthy blood donors; Asians (Chinese, Malays, Indians)	None noted
Yates, 1997	Caucasian patients	None noted
Winter, 2007	Patients with IBD who had received either azathiopurine or mercaptopurine	Prior knowledge of TPMT status
Schaeffeler, 2004	Healthy German adult blood donors	Regular drug use (oral contraceptives and vitamins excepted)
Loennechen, 2001	Caucasian and Saami adults admitted to specialist cardiology medical center	None noted
Spire-Vayron de la Moureyre, 1998	Healthy Volunteers and persons who require TPMT activity testing prior to receiving thiopurine therapy	None noted
Nasdekina, 2005	Control Samples of DNA from St. Jude's Children's Research Hospital, TN, USA	None noted
Lu, 2005	70 children with acute lymphoblastic leukemia; 55 healthy blood donors	None noted
Schaeffeler, 2001	Caucasian adult patients receiving thiopurines, and Caucasian adult patients who had stopped thiopurine therapy due to adverse events	None noted
Anglicheau, 2002	Autoimmune skin disease or IBD patients	None noted
Wusk, 2004	Healthy persons or patients with IBD; Caucasian volunteers	None noted
Rossi, 2001	Healthy blood donors	None noted
Alves, 2001	Healthy European Volunteers	Therapy with a diuretic
Hon, 1999	Healthy volunteers and children with ALL; Caucasian and African Americans	None noted
Reis, 2003	Volunteers recruited at the Brazilian National Cancer Institute Blood Bank	None noted
Larovere, 2003	Healthy Argentine blood donors	None noted
Indjova, 2003	Persons of Bulgarian origin 12-85 years of age	Therapy with thiopurines, blood transfusions and renal insufficiency

IBD: Irritable Bowel Disease; TPMT: thiopurine methyltransferase

3.1.2.2 Gold Standard and Investigational Tests

The seventeen trials included eight different methods of genotyping and three different methods of phenotyping (Table 13).

Table 13: Gold Standard and Investigational Tests Evaluated

Study	Investigational Test	Gold Standard
Yates, 1997	PCR Genotype	RC Assay – Phenotype
Winter, 2007	PCR- RFLP Genotype	TMS -Phenotype
Schaeffeler, 2004	PCR-RFLP Genotype	HPLC Phenotype
Loennechen, 2001	PCR Genotype	RC Assay – Phenotype
Spire-Vayron de la Moureyre, 1998	PCR Genotype	RC Assay – Phenotype
Hon, 1999	PCR Genotype	Phenotype method not described
Reis, 2003	PCR Genotype	RC Assay Phenotype
Larovere, 2003	AS-PCR Genotype & PCR-RFLP	HPLC Phenotype
Indjova, 2003	PCR Genotype	HPLC Phenotype
Nasedkina, 2005	Biochip PCR-Genotype	Multiplex PCR Genotype
Lu, 2005	APEX Genotype	ARMS-PCR or PCR-RFLP Genotype
Schaeffeler, 2001	DHPLC Genotype	PCR-RFLP & AS-PCR Genotype
Anglicheau, 2002	HPLC & RC Assay – Phenotype*	PCR Genotype
Wusk, 2004	HPLC Phenotype	PCR Genotype
Kham, 2008	RC Assay – Phenotype	PCR–RFLP Genotype
Rossi, 2001	RC Assay – Phenotype	PCR Genotype
Alves, 2001	HPLC Phenotype	PCR Genotype

APEX: arrayed primer extension; ARMS: amplification refractory mutation system; AS: allele specific; DHPLC: denaturing high performance liquid chromatography; HPLC: high performance liquid chromatography; PCR: polymerase chain reaction; RC: radiochemical; RFLP: restriction fragment length polymorphism; TMS: Tandem mass spectrometry
* Both methods were tested against the PCR genotype test

Though the phenotype test had been developed first, it was only used as the gold standard in nine of the studies. Three studies compared two different methods of genotyping. The most common method for genotyping was PCR, though other methods including restriction fragment length polymorphism PCR, arrayed primer extension, biochip technology, allele specific PCR, denaturing high performance liquid chromatography and amplification refractory mutation system PCR were also used. The most common form of phenotype conducted was the high performance liquid chromatography (HPLC) assay, however radiochemical assay and tandem mass spectrometry were also used.

3.1.2.3 Phenotype Test

TPMT activity is defined as the formation of 1 nmol of metabolite (e.g. 6-methylmercaptopurine or 6-methylthioguanine) per milliliter of packed red blood cells per hour incubation at body temperature (37° C). TPMT activity is measured by isolating the enzyme, incubating it at 37° C and exposing it to a thiopurine drug (e.g. 6-MP, thioguanine or azathioprine). The metabolite is then measured after one hour to determine how much of the drug was broken down. In studies in which the phenotype test was used as the gold standard[19, 23, 28, 29, 35, 58, 75, 76, 79] activity ranges were defined for comparison against the investigational test, which was a genotype test for these three scenarios (Table 14). In the study by Anglicheau, et al. (2002),[74] which used the phenotype test as the investigational test, ranges were set based on how well they matched with the genotype comparator. Two studies used a form of phenotype test and did not discuss the reference ranges used.[27, 61]

Table 14: TPMT Activity Ranges

Study	Measurement Units	TPMT Deficient	Low TPMT Activity	Normal TPMT Activity	High TPMT Activity	Activity Method
Yates, 1997	nmol/ml pRBC/hr	<5	5-10	>10	NR	RC
Winter, 2007	nmol/g Hb/hr	<10	10-25	26-50	>50	TMS
Schaeffeler, 2004	nmol/g Hb/hr	<5	5-10	>10	NR	HPLC
Loennechen, 2001	U/mL pRBC	<5	5-9.5	> 9.5	NR	RC
Spire-Vayron, 1998	U/mL pRBC	<5	5-13.7	>13.7	NR	RC
Hon, 1999	U/mL pRBC	<5	5-10.1	>10.1	NR	NR
Reis, 2003	mmol/mL pRBC /hr	NR	NR	>11.3	NR	NR
Larovere, 2003	pmol x 10^{-7} / pRBC/hr	<3	3-7	>7	NR	NR
Indjova, 2003	nmol/mL erythrocyes/hr	<2.8	2.8-10	>10	NR	NR
Anglicheau, 2002[†]	nmol/ml pRBC /hr	NR	4.3-7.7	5.9-28.7	NR	HPLC
		NR	7-10	6.2-28	NR	RC
Kham, 2008	nmol/ml pRBC /hr	<5	NR	NR	NR	RC
Alves, 2001	mmol/mL pRBC /hr	<9	9-18	>18	NR	HPLC

NR: Not reported; RC: Radiochemical Assay; TMS: Tandem mass spectrometry; HPLC: high performance liquid chromatography; pRBC: packed red blood cells; hr: hour; Hb: hemoglobin
† This study evaluated two methods of phenotypes

3.1.2.4 Genotype Test

To date there are at least 23 single nucleotide polymorphisms in the TPMT gene. TPMT *1 is considered the wild type or normal allele and TPMT *2 through *22, with some sub-categories (e.g. *3A, *3B, *3C, *3D), are associated with reduced or deficient TPMT activity. All of the included studies used a method of genotyping as the gold standard or investigational test, however the tests varied slightly in the specific polymorphisms in which they were designed to detect. Table 15 outlines the polymorphisms that each of the studies looked for in their evaluation.

Table 15: Polymorphisms Identified in Study Tests

Study	Population	TPMT Polymorphisms Studied
Yates, 1997	Caucasian	TPMT *2, *3
Winter, 2007	Study carried out in the UK, no ethnic group specified	TPMT *2, *3A, *3B, *3C
Schaeffeler, 2004	Caucasian (German)	TPMT *2, *3A, *3B, *3C, *3D
Loennechen, 2001	Caucasian or Saami	TPMT *2, *3A, *3B, *3C, *6
Spire-Vayron de la Moureyre, 1998	European	TPMT *2, *3A, *3C *3D, *4, *5, *6, *7
Hon, 1999	Caucasian and African American	TPMT *2, *3A, *3B, *3C
Reis, 2003	Brazilian	TPMT *2, *3A, *3C
Larovere, 2003	European, Amerindian and Middle Eastern	TPMT *2, *3A, *3B, *3C, *4, *5, *6, *7, *8
Indjova, 2003	Bulgarian	TPMT *2, *3A, *3B, *3C, *4, *6
Nasedkina, 2005	Study carried out in the USA, no ethnic group specified	TPMT *2, *3A, *3B, *3C, *3D, *7, *8
Lu, 2005	Study carried out Singapore, no ethnic group specified	Not Specified
Schaeffeler, 2001	Caucasian (German)	TPMT *2, *3A, *3C *3D
Anglicheau, 2002	Study carried out in France, no ethnic group specified	TPMT *2, *3A, *3B, *3C
Wusk, 2004	Study carried out in the Switzerland, no ethnic group specified	TPMT *2, *3A, *3B, *3C, *7
Kham, 2008	Asians (Chinese, Malays, Indians)	TPMT *3A, *3C and *6
Rossi, 2001	Italian	TPMT *2, *3A, *3B, *3C
Alves, 2001	European	TPMT *2, *3A, *3C, *8

The presence of polymorphisms TPMT *2 and TPMT*3A, *3B, *3C and *3D were frequently

tested, though only TPMT *3A and *3C were tested in all sixteen of the included studies, where

specific mutations were disclosed. Mutations *6, *7 and *8 were only sought in specific

populations.

3.1.2.5 Sensitivity and Specificity

Sixteen of the seventeen included studies reported sensitivity and specificity or provided enough

raw data to have these measures calculated. One study simply stated concordance to be 100%

but did not specify the calculation that was used to determine this.[77] Contingency tables were

created, where possible, for each of the studies (Appendix 2). Positive and negative predictive

values were determined from the data in the 2x2 tables (Table 16).

Table 16: Reported Agreement Values

Study	Investi-gational Test	Gold Standard Test	Sensitivity (%)	Specificity (%)	Positive Predictive Value (%)	Negative Predictive Value (%)	Concord-ance (%)
Yates, 1997	Genotype	Phenotype	96.30	100.00	100.00	95.45	NR
Winter, 2007	Genotype	Phenotype	64.71*	100.00*	100.00	94.96	NR
Schaeffeler, 2004	Genotype	Phenotype	86.82	99.45*	94.92	98.45	NR
Loennechen, 2001	Genotype	Phenotype	95.83	100.00	100.00	99.58	
Spire-Vayron de la Moureyre, 1998	Genotype	Phenotype	54.55	94.30	66.67	90.85	87.00
Hon, 1999	Genotype	Phenotype	78.26	100.00	100.00	82.14	
Reis, 2003	Genotype	Phenotype	87.10	100.00	100.00	91.49	
Larovere, 2003	Genotype	Phenotype	57.14	100.00	100.00	75.68	
Indjova, 2003	Genotype	Phenotype	62.50	100.00	100.00	85.25	
Nasedkina, 2005	Genotype	Genotype	100.00	100.00	100.00	100.00	NR
Lu, 2005	Genotype	Genotype	NR	NR	NR	NR	100.00
Schaeffeler, 2001	Genotype	Genotype	100.00	100.00	100.00	100.00	NR
Anglicheau, 2002 (RC Assay)	Phenotype	Genotype	100.00	86.21*	27.27	100.00	NR
Anglicheau, 2002 (HPLC)	Phenotype	Genotype	100.00	89.66*	33.33	100.00	NR
Wusk, 2004	Phenotype	Genotype	100.00*	88.74*	41.86	100.00	NR
Kham, 2008	Phenotype	Genotype	100.00	98.28	63.64	100.00	NR
Rossi, 2001	Phenotype	Genotype	100.00	96.74	78.57	100.00	
Alves, 2001	Phenotype	Genotype	91.67	97.71	78.57	99.22	

NR: Not Reported; RC: Radiochemical; HPLC: High performance liquid chromatography
*These values were reported in the study. All remaining values were calculated from study data.

These results demonstrate that phenotype technologies have more positive test results than genotype tests. Phenotype tests show a high sensitivity and low PPV when compared to a genotype test as a gold standard. Genotype tests show a low sensitivity and high PPV when compared to a phenotype test as a gold standard. When two different genotype tests are compared head-to-head, the accuracy measured is high, showing consistency between technologies. None of the studies ascertained accuracy by comparing different phenotype tests head-to-head, therefore it is unknown if accuracy is consistent between them.

3.1.3 Quality Appraisal

A quality appraisal of included trials was carried out using a modified version of the CASP tool for evaluating diagnostic tests.[42] One question was added to determine if a sample size calculation was considered. Two questions were added to the tool that were specific to the tests in question. The first asked about the use of concurrent medications and the second asked if the study accounted for blood transfusions. These results are presented in Appendix 3.

None of the seventeen studies included in this review were considered to be of high quality based on the CASP tool for diagnostic studies. Though all studies had appropriate reference test standards, only ten outlined well-defined study questions.[19, 23, 27, 29, 58, 61, 73, 75, 78, 79]

Considering how infrequently homozygous mutations tend to occur, studies should be set up to power the comparisons between the tests. None of the studies calculated a sample size nor did they discuss the issue of statistical power in their analysis.

The use of blinding was not commonly mentioned in study methodologies. However, due to the fact that subjectivity would not influence test results, blinding may not have been necessary to ensure unbiased interpretation of the results.

Generally the study population was clearly defined and consisted of a mix of healthy participants [19, 27, 28, 59, 61, 73] as well as participants already taking a thiopurine agent such as those with irritable bowel disease (IBD)[58, 60, 61, 74, 77] or auto-immune skin disorders.[74] As clearly outlined in the literature, phenotype tests can be influenced by recent blood transfusions. With the exception of one study,[75] it is unclear whether the results presented in these studies took this into consideration in the analysis. Three studies stated that healthy blood donors were used as the study sample,[23, 28, 59] so it can be assumed that blood transfusions did not influence the results in these analyses.

Methodologies for carrying out each of the tests were described in varying amounts of detail depending on the nature of the journal in which they were published. Those published in chemistry or biochemistry focused journals tended to provide more detail than those published in more clinically based journals.

Overall, the study conducted by Schaeffeler et al. (2004)[28] had the highest methodological quality, with the most positive responses to the quality appraisal tool. Schaeffeler included the greatest number of patients (1214) and therefore was most likely to get more realistic estimates. There was no study that stood out as having the lowest methodological quality.

3.2 *Economic Analysis*

The decision analysis was carried out by calculating expected costs and life months associated with each intervention arm (Figure 3).

Figure 3: Decision Tree with Terminal Values
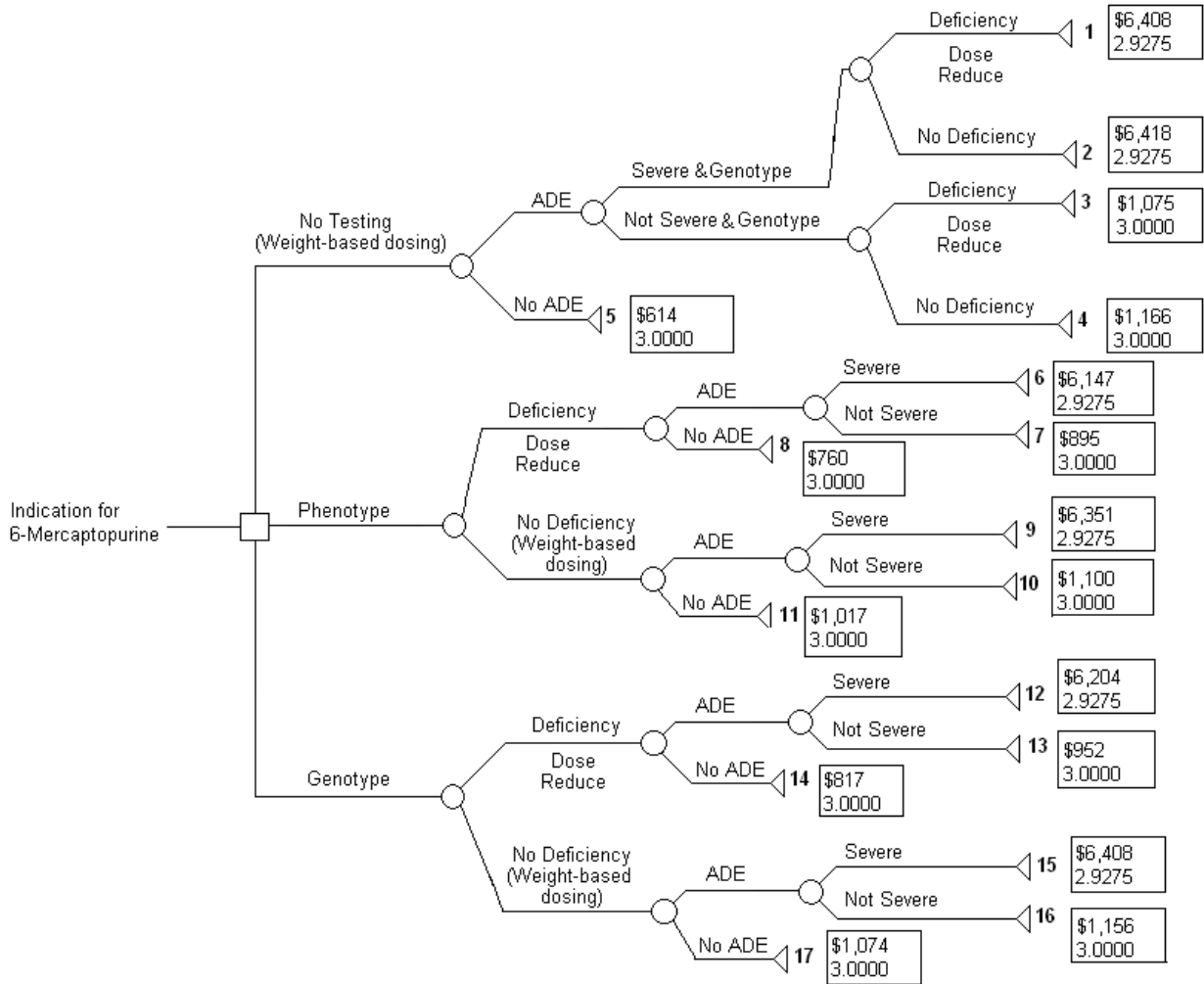
Note: The box at each terminal node contains the cost and life-months assigned to that branch

3.2.1 Base Case Analysis

Total costs at each terminal node were calculated by summing all the related costs for each branch (Table 9). Total life-months at each terminal node were determined by taking the risk of death (2.9% or 0% depending on the pathway) and multiplying it by the number of months remaining in the three month time horizon (2.5 months) and then subtracting the result from three months. Branch pathway probabilities were calculated by multiplying the probabilities at each chance node.

77

To determine the expected life-months for each of the three test options, life-months and probabilities were multiplied together. To determine the expected sub-costs at each terminal branch, the total cost for each pathway was weighted by its respective probability (Tables 17, 18 and 19).

Table 17: Expected Cost and Life-Months of the Weight-Based Dosing Arm

Terminal node #	Total Cost	Life Months	Probability	Expected Sub-Cost	Expected Life-Months
1	$6,408	2.9275	0.0005	$3.20	0.0015
2	$6,418	2.9275	0.004	$25.67	0.0117
3	$1,075	3.0000	0.0029	$3.12	0.0086
4	$1,166	3.0000	0.0226	$26.35	0.0679
5	$614	3.0000	0.97	$595.58	2.91
Total				$653.92	2.9997

The total expected cost of the weigh-based treatment arm, after summing all the treatment branches, was $653.92 and the expected life- months was 2.9997.

Table 18: Expected Cost and Life-Months of the Phenotype Arm

Terminal node #	Total Cost	Life-Months	Probability	Expected Sub-Cost	Expected Life-Months
6	$6,147	2.9275	0.0003	$1.84	0.0009
7	$895	3.0000	0.0018	$1.61	0.0054
8	$760	3.0000	0.0783	$59.51	0.2348
9	$6,351	2.9275	0.0037	$23.50	0.0109
10	$1,100	3.0000	0.0211	$23.21	0.0632
11	$1,017	3.0000	0.8948	$910.01	2.6844
Total				$1,019.68	2.9997

The total expected cost of the phenotype treatment arm, after summing all the treatment

branches, was $1,019.68 and the expected life-months was 2.9997.

Table 19: Expected Cost and Life-Months of the Genotype Arm

Terminal node #	Total Cost	Life-Months	Probability	Expected Sub-Cost	Expected Life-Months
12	$6,204	2.9275	0.0001	$0.62	0.0004
13	$952	3.0000	0.0007	$0.67	0.0021
14	$817	3.0000	0.0301	$24.59	0.0904
15	$6,408	2.9275	0.0040	$25.63	0.0117
16	$1,156	3.0000	0.0227	$26.24	0.0681
17	$1,074	3.0000	0.9423	$1,012.03	2.8270
Total				$1089.78	2.9997

The total expected cost of the genotype treatment arm, after summing all the treatment branches,

was $1,089.78 and the expected life-months was 2.9997.

Both TPMT testing arms were more costly than the weight-based dosing arm. The genotype arm

was slightly more costly than the phenotype arm, though with known uncertainty around some of

the included values, it is not possible to say that one is actually more expensive.

These results show that there is no clinically important difference in survival between each of the

testing strategies over the three-month time horizon. For this reason incremental cost

effectiveness ratios were not calculated. Results are presented as incremental costs instead

(Table 20).

Table 20: Incremental Costs

	Incremental Cost
Phenotype minus Weight-based	$365.76
Genotype minus Weight-based	$435.86
Genotype minus Phenotype	$70.10

Incremental costs show that using phenotype or genotype tests prior to starting a patient on a thiopurine agent increases costs with no increase in benefit.

## 3.2.2 Univariate Sensitivity Analysis

One-way sensitivity analyses were performed by varying the prices of the genotyping and phenotyping tests, which were highly variable across jurisdictions (Table 21).

Table 21: Incremental Costs in Univariate Sensitivity Analysis

	Phenotype minus Weight-based	Genotype minus Weight-based	Genotype minus Phenotype
Cost of phenotype test $100-500	$63 to $463	$436 to $436	-$27 to $373
Cost of genotype test $100-$500	$366 to $366	$87 to $475	-$290 to $110

As can be seen in Table 21, the only scenario in which the decision might be altered by the price inputs is in comparing the phenotyping and genotyping strategies. The results of the sensitivity analysis are presented in a tornado diagram (Figure 4). The vertical bar in the centre represents the cost determined from the base case analysis. Each of the horizontal bars depicts the impact of that associated variable on overall incremental cost.

Figure 4: Tornado Diagram of Genotype minus Phenotype
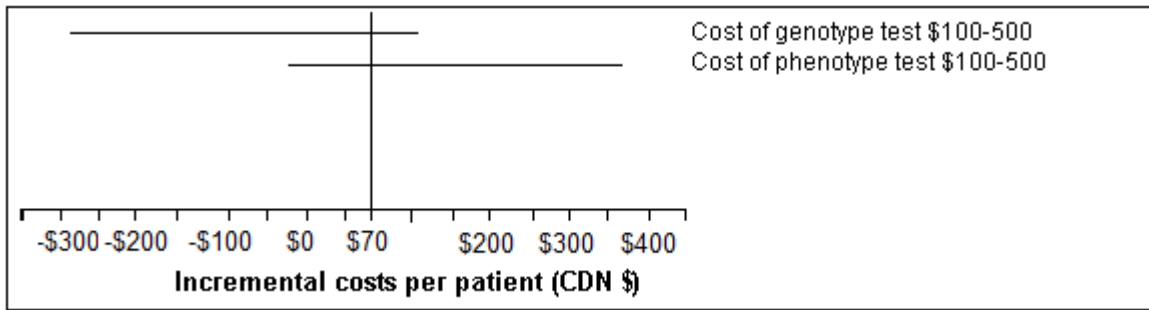
All costs in 2008 Canadian dollars

In the base case analysis, the incremental costs of the genotype and phenotype strategies were negligible. If one test was more expensive than the other, then the decision would favour the strategy that had the cheaper test. Reducing the cost of the phenotype or genotype tests made a considerable impact on the incremental costs of these treatments over weight-based dosing, though neither was dominant over weight-based dosing.

Other one-way sensitivity analyses were conducted to examine the effects of varying the cost of the hospital stay, the proportion of myelosuppression cases requiring hospitalization, and phenotype and genotype sensitivity and specificity (results not shown). In all of these sensitivity analyses, the incremental costs changed by less than $100, showing no significant impact.

3.2.3 Threshold Analysis

A threshold analysis was conducted to determine at what cost the phenotype and genotype tests would have to be offered for the incremental cost over weight-based dosing to be zero dollars. The threshold costs were $39 for the phenotype test and $12 for the genotype test. If the costs of these tests were to fall below these prices then the intervention would be deemed cost-effective, based on the single outcome of myelosuppression, compared to the standard of care, weight-based dosing.

3.2.4 Probabilistic Sensitivity Analysis

As described above, there were no differences in the effectiveness of the different strategies found in the base case analysis. The same was found in the PSA. The costs of each strategy (Table 22) and the incremental costs of the three strategies (Table 23) are presented below.

Table 22: Results of the Probabilistic Sensitivity Analysis

| Strategy	Mean Cost	Lower 95% CI	Upper 95% CI
Weight-based dosing	$669	$547	$791
Phenotype testing	$967	$721	$1,213
Genotype testing	$946	$659	$1,233

The mean costs of each strategy are approximately the same as those found in the base case analysis. The weight-based dosing strategy remained the least expensive. The PSA found that there is little meaningful difference between the phenotype and genotype dosing strategies as there is significant overlap between the 95% CIs.

Table 23: Incremental Costs in Probabilistic Sensitivity Analysis

	Phenotype minus Weight-based	Genotype minus Weight-based	Genotype minus Phenotype
Mean Costs	$298	$277	-$21
Lower 95% CI	$116	$73	-$343
Upper 95% CI	$480	$481	$301
Per cent of simulations where intervention is more costly (%)	99.79	99.48	56.08
Upper 95% CI contrasted with Lower 95% CI	$666	$686	$512
Lower 95% CI contrasted with Upper 95% CI	-$70	-$132	-$554

In virtually all simulations, weight-based dosing was less costly than genotyping or phenotyping. A scatterplot of the cost-effectiveness results of the PSA shows that in the majority of the simulations, the weight-based dosing strategy will be less costly than either the phenotyping or genotyping strategy, and that the costs of those two strategies are generally similar (Figure 5).

Figure 5: Cost-Effectiveness Scatterplot of Testing Strategies

CE Scatterplot
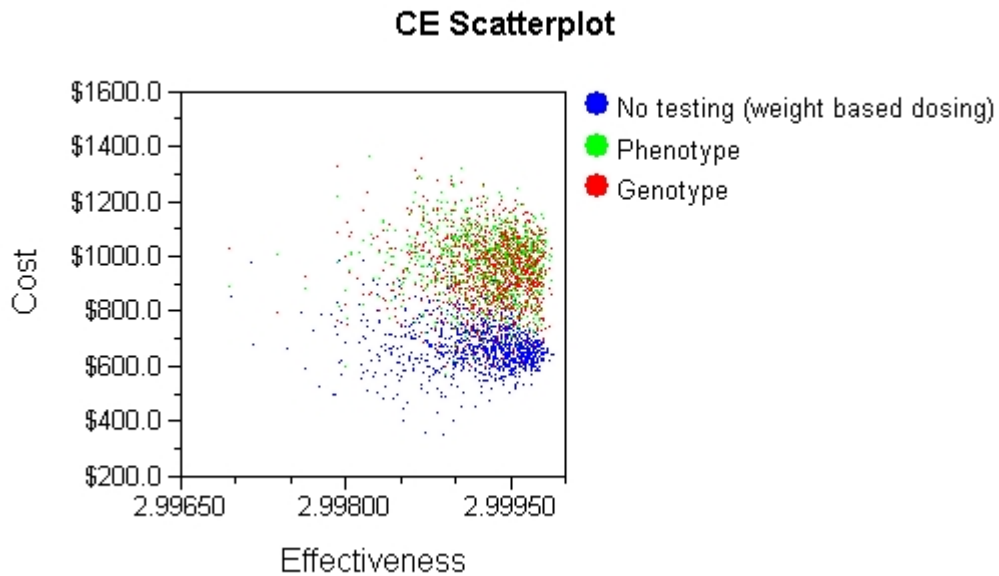
The sensitivity analyses indicated that pre-testing with phenotype or genotype technologies is not a cost-effective option. Altering uncertain variables in a PSA did not affect the conclusions. This is largely due to the fact that only TPMT homozygous patients were considered in the deficiency branches of the phenotype and genotype testing arms. The presence of a TPMT homozygous mutation is so rare (1/300), that the decision analysis model was not able to prevent sufficient cases of myelosuppression to demonstrate a distinguishable difference in life-months between testing arms. Variations in inputs for costs and probabilities, with the exception of phenotype and genotype test sensitivity, specificity and cost, affected each of the three arms equally. Varying sensitivity and specificity of the tests likely did not affect the results greatly because the ranges were narrow.

4   DISCUSSION

A systematic review and CEA were conducted to determine whether TPMT genotype or phenotype testing technologies are effective tools to assist in personalizing doses of 6-MP and thus prevent ADEs. Contrary to the hypothesis, using TPMT phenotype or genotype tests prior to the first dose of 6-MP therapy did not prove to be cost-effective compared to standard weight-based dosing. This assessment highlights a number of important issues and gaps in the literature. Additional research in these areas may allow for more certainty in the probabilities and costs identified which would strengthen future technology assessments of TPMT tests for the ALL population.

4.1   Systematic Review of the Literature

The first objective of this study was to review the literature describing the accuracy of the phenotype and genotype technologies by ascertaining sensitivity, specificity and positive and negative predictive values. As suspected, the phenotype tests identified more positive results compared to the genotype tests because they detected all deficiencies in the enzyme, not only those influenced by TPMT gene mutations. Genotype tests were accurate; however they were limited by the number of mutations the test was designed to detect. As a result, neither test could be considered the gold standard. Additionally, findings were drawn from studies of low quality; as sample sizes and confounding variables, such as blood transfusions and concurrent medications, were rarely considered in included studies.

4.1.1 Gold Standard

In any study designed to detect the accuracy of a diagnostic or screening test, it is essential that an appropriate gold standard be selected. The gold standard should be selected based on current clinical practice or an existing test known to have the highest accuracy. In the case of

pharmacogenetic testing a gold standard is not always available as the technology is still relatively new. There is no known existing test that is designed to detect 100% of known TPMT mutations.

When a phenotype test is presented as the gold standard for comparison against a genotype test, we would expect less than 100% sensitivity for the genotype test. This was the case for the three included studies that used phenotype technologies as the gold standard.[28, 35, 58] When a genotype test is presented as the gold standard and compared to the phenotype test, we would expect a very high sensitivity for the phenotype test, close to 100%,[59, 61, 74] however these results can be misleading. The phenotype test is likely to detect additional positive cases compared to the genotype test because it detects a decrease in activity which could be caused by any known or unknown mutations or external confounding factors (e.g. drug interactions, blood transfusions). This increased likelihood to detect positive cases leads to a lower PPV. Positive and negative predictive values are important in describing how accurately a test performs, although not as universally used as sensitivity and specificity due to the fact that they are dependant on prevalence of deficiency. It was found that very few studies presented these calculations.

A hypothetical ideal gold standard would be one in which the sensitivity and specificity are 100%. In a clinical setting, an ideal gold standard rarely exists due in part to cost and the time required to carry it out. In the case of TPMT, a genotype test designed to detect all 23 known mutations would be the best choice for the gold standard, but is prohibitive given the costs. For this reason it would not be ideal in a clinical setting to use such a genotype test, though it would be useful for research purposes to get a better sense of the true sensitivity, specificity and positive and negative predictive values of the existing phenotype and genotype technologies.

4.1.2 Sample Size

Of the nine studies included in this analysis, none reported a sample size calculation. Sample sizes ranged from 58 to 1,214 patients. With a prevalence of homozygous TPMT mutations occurring in one out of 300 persons, most of the studies included in the review were not powered to detect even one homozygous patient. Estimates of sensitivity and specificity can be imprecise if the sample used is not large enough. Small sample sizes will also lead to large confidence intervals, limiting the clinical usefulness of the point estimate. The lack of reporting of sample size found in this assessment appears to be a common fault in many diagnostic and screening studies. An assessment of the literature to determine how many studies which evaluated a test's accuracy performed a sample size calculation screened all issues of eight well known medical journals published in 2002 for studies on test accuracy.[80]   Of 43 eligible studies, only 5% (2) reported performing a sample size calculation.[80]

4.1.3 Consideration of Confounding Variables

There were minimal details regarding the inclusion and exclusion criteria used in the nine included studies. Though it is not expected that selection criteria would be extensive, they should account for potential confounding factors in evaluating test accuracy. TPMT phenotype technologies are known to be influenced by concurrent medications and blood transfusions. It can be assumed that two of the included studies excluded subjects with blood transfusions, due to the fact that blood donors rather than patients made up the study sample,[28, 59] though it was not clearly stated as an exclusion criterion. Use of concurrent medications was only considered in one of the studies.[28] TPMT genotypes are known to vary between ethnic groups, with certain mutations being more common in specific populations. Five of the studies included in this review provided details of the ethnic background of the study population.[28, 35, 59-61] The remaining four studies

alluded to the country in which the study took place but did not specify that inclusion was limited to a specific population.

4.2 Cost-Effectiveness Analysis

The second objective of the study was to determine the incremental cost of TPMT genotyping, phenotyping and weight-based dosing strategies for every life-month saved. The outcome of life-months was selected due to the short time horizon; another option could have been to use life-weeks. Since week-level information was not available, there was no additional benefit to using life-weeks. Life-months were therefore selected to facilitate interpretation of results. Since there was no difference in life-months between the three arms of the decision tree, it was not possible to calculate an ICER. The analysis showed that there would be an additional cost to offering either the phenotype test or genotype test prior to dosing 6-MP over the standard of care as described in the COG protocols, thus these alternatives were not cost effective to reduce the mortality and morbidity associated with 6-MP-induced neutropenia. The reduction in the occurrence of neutropenia is only one outcome measure that could be used to determine the benefits of TPMT testing. Future research should consider other ADEs such as liver toxicity, as well as efficacy outcomes such as long-term survival, rate of relapse and development of secondary malignancy. There is presently very little available evidence on the incidence and impact on survival for these outcomes. As a result, they could not be considered in this study.

The decision model did not account for the potential ADEs that may have occurred for those patients who were under-dosed based on a false positive test result. The impact on costs and on survival of dose reducing patients who received false positive test results was therefore not evaluated. However, contingency tables that demonstrate the rates of homozygous mutations using sensitivities and specificities of the phenotype and genotype tests as identified in the systematic review and used in the CEA, can shed some light on this issue (Table 24 & 25).

Table 24: Contingency Table for TPMT Homozygous Mutations using a Phenotype Test

	Homozygous Mutation Present	Homozygous Mutation Absent	Totals
Test Positive	True Positive (0.0033)	False Positive (0.1096)	Total with positive test (0.1129)
Test Negative	False Negative (0.0000)	True Negative (0.8871)	Total with negative test (0.8871)
Totals	Total persons with mutation (0.0033)	Total persons without mutation (0.9967)	Total tested (1.0000)

Phenotype Test: Sensitivity: 100%, Specificity: 89%

Table 25: Contingency Table for TPMT Homozygous Mutations using a Genotype Test

	Mutation Present	Mutation Absent	Totals
Test Positive	True Positive (0.0030)	False Positive (0.0199)	Total with positive test (0.0229)
Test Negative	False Negative (0.0003)	True Negative (0.9768)	Total with negative test (0.9771)
Totals	Total persons with mutation (0.0033)	Total persons without mutation (0.9967)	Total tested (1.0000)

Genotype Test: Sensitivity: 90%, Specificity: 98%

Both Tables 24 and 25 show the rate of false positives to be higher than that of true positives. Since TPMT homozygous patients receive dose reductions of greater than 90%, the risk of therapy failure in dose reduced false positive patients is expected to be high. The question is, *are there more patients put at risk for adverse outcomes than are saved from adverse outcomes by testing for TPMT prior to therapy?* This issue highlights the importance of accuracy of the testing technologies. To prevent under-dosing false positives, a test that is 100% specific is required.

To date there have been four other cost-effectiveness analyses that assessed TPMT technologies for their usefulness in preventing ADEs by guiding initial doses.[43, 45, 49, 81] However, only Van den Akker-van Marle (2006)[43] examined a pediatric ALL population. They used a cost-effectiveness model from a societal perspective and found that TPMT genotyping was a cost effective intervention at a cost of €4800 (approximately $7,879 CDN) per life year saved. They assumed a genotype specificity of 100% and therefore did not encounter the issue of under-dosing false positives. This value was not altered in the sensitivity analysis. Like the CEA presented in here, van der Akker van Marle only assessed the impact of TPMT testing on one ADE of myelosuppression, however the time horizon was extended over a life time which allowed the model to demonstrate a greater impact on survival.

By using a three-month time horizon, as done in this assessment, it was difficult to show a clinical difference in survival rendering the testing interventions more costly without added benefit.   The selection of a three-month time horizon was based on the assumption that any adverse drug events caused by the presence of a deficiency in the TPMT enzyme would occur in the first three months of therapy (J. Hand, personal communication, September 9, 2008 (Appendix 1)). Extending the time horizon beyond this point would not be a true reflection of the mortality due to neutropenia, but rather a reflection of the mortality associated with the underlying disease. Additionally, there was no evidence found to suggest how frequently neutropenia would occur during a treatment with 6-MP and therefore it is not clear how to model the ADE with a longer time

horizon. After the first occurrence of neutropenia, which is caused by TPMT deficiency, it is expected that dose adjustments would be made and thus reduce the risk of recurrence.

Van den Akker-van Marle also assumed a lower cost for the genotype test (€30 to €300, approximately $49 CDN to $493 CDN) and a longer hospital stay (seven days), both of which work in favor of making the testing option cost-effective.[43] Moreover, the study did not consider TPMT-based 6-MP dosing in the analysis. Additionally, there was no discussion or reference to the lack of conclusive evidence to suggest appropriate dose adjustments in pediatric ALL patients. Their recommendation to 'seriously consider' TPMT genotyping prior to commencing therapy, may therefore be a little premature.

The three remaining economic analyses also found TPMT testing to be cost-effective, however they were conducted in different patient populations and therefore are not generalizable to the ALL population.[45, 49, 81] In the study conducted by Oh (2004),[81] a rheumatoid arthritis and a systemic lupus erythematosus population were evaluated. Results showed TPMT genotyping with PCR to be cost-effective compared to conventional weight-based dosing, despite the short time horizon of one year. There are other differences between the current study and the one conducted by Oh et al. that may account for the opposing results. Firstly, the evaluation was conducted in Korea and costs may vary greatly in comparison to the North American health care system. The cost of the PCR test was estimated to be equivalent to $50 USD. A threshold analysis was conducted and the cost of the PCR test that would show equivalence between genotype and weight-based dosing arms was $278 USD. In the current study a cost of $380.50 USD was used for genotype tests based on actual prices from the Mayo Clinic. If this value had been used by Oh et al. the technology would not have been shown to be cost effective. Additionally it was assumed that patients who have a wild-type genotype would not experience myelosuppression and that all TPMT heterozygous and homozygous patients would experience mild or severe myelosuppression if not appropriately dose adjusted. Since not all

myelosuppression is caused by deficiencies in the TPMT enzyme, and since it is not clear if TPMT heterozygous patients will experience adverse drug events, these assumptions can overestimate the differences between the treatment arms.

Marra (2002)[45] conducted a study in a rheumatoid arthritis population treated with azathioprine over a life-long time horizon. They estimated that 50% of myelosuppression cases were due to TPMT deficiency and that 50% of cases would require hospital admission based on expert opinion, compared to the 30% and 15% used respectively for the current analysis. The cost of a PCR genotype test was assumed to be $100 CDN which was based on the cost of other PCR tests, because TPMT-specific costs were not available. As shown in the sensitivity analysis of the current study, the costs of the TPMT test exerts the greatest impact on incremental cost. The lower cost used by Marra, along with the longer time horizon and greater risk of myelosuppression, may account for the differences in results to this study.

Winter (2004)[49] studied an inflammatory bowel disease population, being treated with azathioprine, using a case model of a theoretical population. They determined that TPMT genotyping is good value for the money at a cost of £347 (approximately $630CDN) per life year saved. However, they assumed a genotype test to cost £30 (approximately $54CDN), which is considerably lower that the price used for this analysis. Patients who were admitted to hospital were also assumed to receive granulocyte colony-stimulating factor, which is a costly agent used to treat myelosuppression. This agent is not recommended in the pediatric ALL population and therefore this additional cost was not considered.

4.3 Limitations

Despite efforts to ensure high methodological quality, there were limitations encountered while conducting this health technology assessment. Data limitations would not allow for the determination of QALMs or QALYs. The accuracy and strength of a cost effectiveness analysis

largely depends upon the quantity and quality of previously published data. Uncertainties in the existing data led to limitations in this analysis which are discussed in further detail below.

4.3.1 Systematic Review

For this study, despite measures to ensure that all relevant material was included in the systematic review, it is possible that some studies were not found in the literature search. Only studies published in the English language were searched and evaluated for inclusion. Also, there may be unknown grey literature which was not found. The potential of not including all relevant material presents the risk of both publication and language bias. Studies that show poor accuracy of TPMT technologies are less likely to be published in peer reviewed journals and may only be available through grey literature.[82] These studies are also less likely to be published in the English language.[83] As a result the accuracy of TPMT technologies may be over-estimated.

4.3.2 Cost-Effectiveness Analysis

Though nine studies were found to meet the inclusion criteria of the systematic review, the type of technologies they evaluated varied substantially. Therefore it is not possible to conclude which test produces the best results. The various technologies also differ in price. Cost information was only available for one genotype test and one phenotype test, so this difference in costs for various technologies was not incorporated into the economic analysis.

Though pediatric oncology patients are very well studied, with most children being enrolled in a study protocol,[84] the rarity of the disease still limits the amount and type of evidence available. As a result many of the values used in this economic analysis, such as the proportion of patients with neutropenia that require inpatient care, were based on expert opinion rather than peer reviewed evidence. As more evidence becomes available, greater certainty can be placed on values included in future assessment.

Including a broader spectrum of outcomes would have strengthened the results. Other outcomes that may have benefited from genetic testing prior to the first dose of 6-MP includes rate of ADEs such as liver toxicity and secondary malignancy, as well as efficacy outcomes such as minimal residual disease and relapse rates. These outcomes are either quite rare (e.g. secondary malignancy) or not life threatening (e.g. liver toxicity) and therefore evidence is not available or limited to case reports and poorly controlled clinical studies. Additionally, long-term studies linking TPMT deficiency to these outcomes were not found. There is a need for additional research to explore the impact of TPMT testing technologies used, to guide initial dosing of thiopurine drugs on long-term survival and the incidence of ADEs. Another limitation was that the decision model did not account for the costs and impact on survival of potential ADEs that may have occurred for those patients who were under-dosed based on a false positive test result.

For the outcome measure of life-months, observed differences between each of the decision tree branches were not detected, and as a result incremental cost effectiveness ratios were not calculated. The use of quality-adjusted life months (QALM) as the outcome measure, rather than simply life-months, would have allowed for the patient's quality of life to be incorporated into the results and may have shown a difference between groups. However, an assessment of utility to reflect the experience of a patient with ALL or a patient with ALL and neutropenia was not found.

The lack of published utilities for health states in this population may be in part due to the difficulty in measuring utility in the pediatric population compared to an adult population. One reason for this may be the fact that it is difficult to differentiate between improvement in health and normal development, as children grow and develop at different rates. Additionally, most of the tools used to determine utility are actually designed for adults. As a result many attributes which are important to child health such as family relationships, autonomy, cognitive skills and body image are not considered. Finally, children, especially young children, do not have the ability to understand and interpret the measurement tasks. To overcome some of these obstacles

researchers often use proxies, such as the parent, to complete the utility assessment. However parents may let their own perceptions of health utility influence their perceptions of their child.[85]

There was uncertainty in the values used for the unit prices of the TPMT genotype and phenotype tests. Though laboratories in Canada were contacted and the test prices requested, they were reluctant to provide either internal costs for conducting the test or the price they would offer to outside institutions to conduct the test. Instead, prices from a U.S. laboratory, the Mayo Clinic, were used. Considering the for-profit nature of the health care system in the U.S., these prices may be inflated compared to what a Canadian laboratory could charge for this service. Additionally, testing provided internally at hospital laboratories, would be at a lower price (since they would be provided at cost and not require shipping or transport costs).

Costs for hospital stays, physician services and laboratory tests were based on Ontario provincial health insurance rates. Not only can prices for these services vary between provinces, standards of care may also vary, causing increased or decreased length of hospital stay. One study, which estimated the length of stay and mortality associated with pediatric oncology patients with febrile neutropenia in the U.S., found the average length of stay for a child with ALL suffering from febrile neutropenia to be nine days.[51] The OCCI data showed that the average length of stay was 5.9 days (SD 6.2 days) at an average cost of $9,641 (range $748 - $94,030) per hospital stay. Since the cost of each day in hospital is not equivalent, it would not be appropriate to extrapolate the cost from a nine day stay as suggested in the Basu paper (2005).[51]

4.4   Generalizability of Findings

There are other patient populations to whom parts of this assessment may be relevant, such as rheumatoid arthritis, IBD and kidney transplantation patients who use thiopurine agents frequently for clinical management and for whom accurate dosing is important. Additionally, users of HTA,

both within Canada and internationally can utilize this assessment to aid decision making and policy planning specifically in the pediatric ALL population.

The results of the systematic review of test accuracy is not specific to pediatric ALL patients and therefore can be generalized across patient groups. However as previously discussed, there are weaknesses in the results of this review that limit its strength, even though it does represent the best available evidence.

The findings of the CEA are generalizable to pediatric ALL patients from other jurisdictions. As shown in the sensitivity analysis, the costs of the phenotype and genotype tests were the only variables that made a noteworthy impact on incremental cost. Even if the prices of these tests vary between jurisdictions, the results presented in the threshold analysis could be used as a guide to assess cost-effectiveness in other places.

The decision analysis model used in this study was specific to the pediatric ALL population and therefore not generalizable to other patient populations. The risk of myelosuppression may vary in other patient groups if, for example, the dose intensity requirements are different or if there are concurrent interacting medications. Costs such as dispensing fees or compounding fees would be lower in adults as they could take tablets and fill their prescriptions for longer lengths of time, unlike children who use the liquid formulation which only has a two-week shelf life. However, this analysis highlighted the deficiencies in the literature that may be common for all populations who use thiopurine drugs.

4.5   Implications for Future Research

In addition to the need for research to establish the usefulness of TPMT phenotype or genotype tests in guiding initial doses of 6-MP in pediatric patients with ALL, there is also a need for research to guide appropriate initial 6-MP doses based on a patient's TPMT status, and evaluate the short-term and long-term outcomes related to TPMT-based dosing.

The manner in which standard dosing ranges are determined is changing. Traditionally, drug dosing ranges were determined by balancing clinical effectiveness with toxicity in the average patient. Some drugs have a narrow therapeutic range as is the case with 6-MP. This means that there is a fine balance in the dose that provides clinical effectiveness and avoids toxicity. Depending on the nature of the drug in question, various tests can be performed to calculate the bioavailability, half-life and volume of distribution as well as measure drug metabolites in a random sample of patients. These tests, in addition to pre-marketing clinical studies, help to determine an appropriate drug dose and frequency. This dosing range however is representative of what is safe and effective in the average patient, not accounting for individual variations.

It is known that not all individuals will respond to a drug in the same way due to variations in metabolizing enzymes or organ function; for example, kidney and liver disease are known to impact the clearance of specific drugs. Many studies have been done to determine appropriate dosing modifications in these circumstances. However it is more difficult to determine dosing modifications based on genetic mutations in metabolizing enzymes.

Thiopurine drugs are examples where mutations in metabolizing enzymes are known to alter drug metabolism. More than 10% of the population has a mutation in the allele responsible for producing the TPMT enzyme. This mutation decreases the amount of active enzyme available to metabolize 6-MP and the patient may require a dose reduction. When these patients are included in the random sample, the dose range determined to be suitable to the average patient may in fact be too low for the normal TPMT metabolizer. In one study, 601 children with ALL were treated with 75 mg/m^2 of 6-MP by physicians who were blinded to the patient's genotype, and were followed-up for an average of 12.4 years to measure the rates of disease relapse and secondary malignancy. The study found that there were significantly more patients who were TPMT wild type who relapsed compared to patients with TPMT deficiency (p=0.03).[86] There was no statistical difference in the rate of secondary malignancy between genotypes. Another study

looked at 814 children with ALL and measured minimal residual disease before and after therapy

with 6-MP. It found that patients with TPMT heterozygous mutations were significantly less likely

to have minimal residual disease compared to TPMT wild type patients (p=0.02).[87] Therefore it

may be necessary to increase 6-MP doses for patients who have the wild-type TPMT allele to

compensate for their increased metabolism. These studies suggest that other outcomes, in

addition to the rate of adverse drug events, need to be taken into account when demonstrating the

usefulness of TPMT testing technologies. With phenotype and genotype technologies now

available, traditional methods for determining drug doses may no longer be adequate. Dose

modifications recommended for patients with TPMT deficiency are not based on large clinical

trials, but on case reports of individuals who experienced ADEs and needed dosage reductions.[34,

88-90] It is still unclear how the results from phenotype and genotype tests should be interpreted

and what actions should be taken to maintain clinical efficacy and safety, though there is a

general movement toward genotyping for customized dosing across many diseases.

Studies have reported that only a third of the cases of neutropenia result from a deficiency in the

TPMT enzyme.[44, 46-48] Drug-drug or drug-disease interactions, blood transfusions and other

thiopurine metabolizing enzymes may also contribute to a patient's risk of ADEs. For example,

inosine triphosphate pyrophosphatase (ITPA) is another enzyme that metabolizes thiopurine

drugs. An examination of the impact of TPMT and ITPA nonfunctional variant alleles in patients

with ALL has been conducted.[91] In one group of patients, dosage adjustments were made based

on TPMT genotype. In this group, there were significantly more cases of febrile neutropenia in

patients with the ITPA variant allele compared to wild-type ITPA genotype (OR = 3.0; 95% CI =

1.2-7.3). In the second group, patients did not have doses adjusted based on TPMT genotype. In

this group, TPMT genotype had a greater influence than ITPA genotype on the incidence of febrile

neutropenia. These results pertained to only one additional enzyme; other unknown enzyme

variants may also influence the incidence of adverse drug events.

In summary, there are a number of areas where additional research is needed to support the use of TPMT phenotype or genotype tests to guide thiopurine doses. Moving away from traditional methods of drug dose determination towards incorporating patient specific characteristics, such as TPMT mutations, is the first step. Studying the impact of personalized medicine on efficacy and safety outcomes is then needed to inform decisions regarding adoption of dose-guiding tests. Finally, understanding that the efficacy and safety of drugs, such as 6-MP, is dependant on multiple factors and a personalized medicine regimen is only possible when all factors are taken into consideration.

4.6   Health Services Impact

Standard of patient care with respect to TPMT testing in pediatric ALL patients vary between Canadian provinces. This inconsistency may stem from the lack of conclusive evidence with respect to best practices. For example, the Janeway Children's Hospital in St. John's, NL, only tests for TPMT status when 6-MP therapy has been delayed for two or more weeks. Traditionally physicians at that hospital used the phenotype test. However, they have recently switched to genotype tests based on additions to the COG guidelines which support the use of the genotype test (J. Hand, personal communication, September 9, 2008 (Appendix 1)). Physicians at The Hospital for Sick Children in Ontario tend to conduct a TPMT genotype test prior to treatment (P. Gordon, personal communication, 2008 (Appendix 1)). The guidelines published by the COG, which recommends testing for TPMT only in the presence of myelosuppression (COG, 2008), continues to be the best strategy to use. Therefore, institutions that follow the COG guidelines should not be impacted by the results of this assessment. Institutions who have adopted the screening for TPMT status prior to the first dose of 6-MP should review their current practice.

As demonstrated in the sensitivity analysis of CEA, the costs of the TPMT phenotype and genotype tests were the only variables that made considerable impact on incremental cost. Since these tests come at high prices, without clinically evident benefits, policies and procedures should

be implemented at the institutional level to ensure appropriate utilization of these technologies. Currently the costs of these tests in the pediatric ALL population are funded by the health care system. The opportunity costs of using such tests outside clinical guidelines need to be taken into consideration. Policies should outline which clinical scenarios are eligible for publicly funded TPMT testing. These should include those described in the COG guidelines as well as within the context of clinical trials that are intended to fill the gap in the literature. Health care organizations will need to be prepared for a potential increase in public pressure for such tests as their availability becomes more widely known. Health technology assessment agencies can play a role in disseminating health economic evidence to inform decision making with respect to pediatric TPMT technologies.

5 CONCLUSIONS

The systematic review of TPMT test accuracy found that available technologies are accurate and consistent; however, there was a lack of good quality studies included in this review to place much confidence in these results. Seeming disregard for study sample size and confounding factors such as concurrent medications and blood transfusions were the main contributors to low quality. There were also inconsistencies in the selection of a gold standard which made it difficult to compare the results of the included studies against one another.

Testing for TPMT mutations prior to the administration of 6-MP in pediatric ALL patients is not a cost effective intervention. This analysis focused only on preventing the ADE of myelosuppression, as this was the only outcome which had reliable data to input into the decision tree. Other outcomes such as liver toxicity, secondary malignancy, minimal residual disease and disease relapse could all be affected by individualizing 6-MP dosing based on TPMT status. In future decision analysis, examining other additional outcomes may lead to more definitive evidence on the usefulness of these tests in this population.

Based on the findings of this HTA, three recommendations can be made. First there is a need for additional basic epidemiological research in the area of pediatric ALL. There is a need to identify the prevalence of myelosuppression and liver toxicity for patients taking 6-MP as well as overall survival. Research is needed to determine utilities for health states in children with ALL, and for children with ALL experiencing various ADEs, to allow for the consideration of quality of life in future analyses.

Secondly, there is need for research initiatives that focus on the impact of TPMT activity levels on outcomes such as myelosuppression, liver toxicity, minimal residual disease, secondary malignancy and survival. This research would not only add confidence to the values used in

future CEAs, it would also aid in identifying new dosing guidelines that are tailored to individual patients.

Finally, the current Children's Oncology Group guidelines provide the most appropriate guidance for the management of pediatric ALL patients on 6-MP, stating that initial doses should be based on weight alone. At this time there is insufficient evidence to recommend the use of phenotype or genotype testing prior to mercaptopurine therapy to guide initial doses in pediatric ALL patients.

REFERENCES

1.   Canadian Cancer Society & National Cancer Institute of Canada. Canadian Cancer Statistics 2007. 2007. (Accessed August 26, 2008, at http://www.cancer.ca/Canada-wide/About%20cancer/Cancer%20statistics/Canadian%20Cancer%20Statistics.aspx?sc_lang=en .)

2.   Rivera GK, Pinkel D, Simone JV, Hancock ML, Crist WM. Treatment of acute lymphoblastic leukemia. 30 years' experience at St. Jude Children's Research Hospital. N Engl J Med 1993;329(18):1289-95

3.   Day SD, Henry DW. Acute Leukemias. In: DiPiro JT, Talbert RL, Yee GC, Matzke GR, Wells BG, Posey LM, eds. Pharmacotherapy: A Pathophysiologic Approach. Fifth ed: McGraw Hill; 2002:2373-95.

4.   Children's Oncology Group. AALL0232: High Risk B-precursor Acute Lymphoblastic Leukemia (ALL): A Phase III Group-Wide Study., 2008. (Accessed September 16, 2008, at http://www.childrensoncologygroup.org/.)

5.   Maitland ML, Vasisht K, Ratain MJ. TPMT, UGT1A1 and DPYD: genotyping to ensure safer cancer therapy? Trends Pharmacol Sci 2006;27(8):432-7.

6.   Canadian Institutes of Health Research, National Sciences and Engineering Research Council of Canada, Social Sciences and Humanities Research Council of Canada. Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans. 1998 (with 2000, 2002, and 2005 amendments). Ottawa.

7.   BC cancer agency professional drug index. 2007. (Accessed January 19, 2008, at http://www.bccancer.bc.ca/HPI/DrugDatabase/DrugIndexPro/Mercaptopurine.htm.)

8.   Roberts RL, Gearry RB, Barclay ML, Kennedy MA. IMPDH1 promoter mutations in a patient exhibiting azathioprine resistance. Pharmacogenomics J 2007;7(5):312-7.

9.   McLeod HL, Krynetski EY, Relling MV, Evans WE. Genetic polymorphism of thiopurine methyltransferase and its clinical relevance for childhood acute lymphoblastic leukemia. Leukemia 2000;14(4):567-72.

10.   Karran P. Thiopurines, DNA damage, DNA repair and therapy-related cancer. Br Med Bull 2006;79-80:153-70.

11.   Weinshilboum RM, Sladek SL. Mercaptopurine pharmacogenetics: monogenic inheritance of erythrocyte thiopurine methyltransferase activity. Am J Hum Genet 1980;32(5):651-62.

12.   Szumlanski CL, Honchel R, Scott MC, Weinshilboum RM. Human liver thiopurine methyltransferase pharmacogenetics: biochemical properties, liver-erythrocyte correlation and presence of isozymes. Pharmacogenetics 1992;2(4):148-59.

13.   Ameyaw MM, Collie-Duguid ES, Powrie RH, Ofori-Adjei D, McLeod HL. Thiopurine methyltransferase alleles in British and Ghanaian populations. Hum Mol Genet 1999;8(2):367-70.

14.   Chang JG, Lee LS, Chen CM, et al. Molecular analysis of thiopurine S-methyltransferase alleles in South-east Asian populations. Pharmacogenetics 2002;12(3):191-5.

15.   Collie-Duguid ES, Pritchard SC, Powrie RH, et al. The frequency and distribution of thiopurine methyltransferase alleles in Caucasian and Asian populations. Pharmacogenetics 1999;9(1):37-42.

16.   Dokmanovic L, Urosevic J, Janic D, et al. Analysis of thiopurine S-methyltransferase polymorphism in the population of Serbia and Montenegro and mercaptopurine therapy tolerance in childhood acute lymphoblastic leukemia. Ther Drug Monit 2006;28(6):800-6.

17.   Haglund S, Lindqvist M, Almer S, Peterson C, Taipalensuu J. Pyrosequencing of TPMT alleles in a general Swedish population and in patients with inflammatory bowel disease. Clin Chem 2004;50(2):288-95.

18.   Hiratsuka M, Inoue T, Omori F, Agatsuma Y, Mizugaki M. Genetic analysis of thiopurine methyltransferase polymorphism in a Japanese population. Mutat Res 2000;448(1):91-5.

19.   Hon YY, Fessing MY, Pui CH, Relling MV, Krynetski EY, Evans WE. Polymorphism of the thiopurine S-methyltransferase gene in African-Americans. Hum Mol Genet 1999;8(2):371-6.

20.   Isaza C, Henao J, Lopez AM, Cacabelos R. Allelic variants of the thiopurine methyltransferase (TPMT) gene in the Colombian population. Methods Find Exp Clin Pharmacol 2003;25(6):423-9.

21.   Kham SK, Tan PL, Tay AH, Heng CK, Yeoh AE, Quah TC. Thiopurine methyltransferase polymorphisms in a multiracial asian population and children with acute lymphoblastic leukemia. J Pediatr Hematol Oncol 2002;24(5):353-9.

22.   Kubota T, Nishida A, Takeuchi K, et al. Frequency distribution of thiopurine S-methyltransferase activity in red blood cells of a healthy Japanese population. Ther Drug Monit 2004;26(3):319-21.

23.   Larovere LE, de Kremer RD, Lambooy LH, De Abreu RA. Genetic polymorphism of thiopurine S-methyltransferase in Argentina. Ann Clin Biochem 2003;40(Pt 4):388-93.

24.   Lu HF, Shih MC, Hsueh SC, Chen CM, Chang JY, Chang JG. Molecular analysis of the thiopurine S-methyltransferase alleles in Bolivians and Tibetans. J Clin Pharm Ther 2005;30(5):491-6.

25.   McLeod HL, Pritchard SC, Githang'a J, et al. Ethnic differences in thiopurine methyltransferase pharmacogenetics: evidence for allele specificity in Caucasian and Kenyan individuals. Pharmacogenetics 1999;9(6):773-6.

26.   Nishida A, Kubota T, Yamada Y, et al. Thiopurine S-methyltransferase activity in Japanese subjects: metabolic activity of 6-mercaptopurine 6-methylation in different TPMT genotypes. Clin Chim Acta 2002;323(1-2):147-50.

27.   Rossi AM, Bianchi M, Guarnieri C, Barale R, Pacifici GM. Genotype-phenotype correlation for thiopurine S-methyltransferase in healthy Italian subjects. Eur J Clin Pharmacol 2001;57(1):51-4.

28.   Schaeffeler E, Fischer C, Brockmeier D, et al. Comprehensive analysis of thiopurine S-methyltransferase phenotype-genotype correlation in a large population of German-Caucasians and identification of novel TPMT variants. Pharmacogenetics 2004;14(7):407-17.

29.   Spire-Vayron de la Moureyre C, Debuysere H, Sabbagh N, et al. Detection of known and new mutations in the thiopurine S-methyltransferase gene by single-strand conformation polymorphism analysis. Hum Mutat 1998;12(3):177-85.

30.   Taja-Chayeb L, Vidal-Millan S, Gutierrez O, Ostrosky-Wegman P, Duenas-Gonzalez A, Candelaria M. Thiopurine S-methyltransferase gene (TMPT) polymorphisms in a Mexican population of healthy individuals and leukemic patients. Med Oncol 2008;25(1):56-62.

31.   Tumer TB, Ulusoy G, Adali O, Sahin G, Gozdasoglu S, Arinc E. The low frequency of defective TPMT alleles in Turkish population: a study on pediatric patients with acute lymphoblastic leukemia. Am J Hematol 2007;82(10):906-10.

32.   Wei H, Zhou S, Li C, Zhang J, Wu J, Huang M. Phenotyping and genotyping studies of thiopurine S-methyltransferase in Kazaks. Pharm Res 2005;22(10):1762-6.

33.   Zhang JP, Guan YY, Wu JH, Xu AL, Zhou S, Huang M. Phenotyping and genotyping study of thiopurine S-methyltransferase in healthy Chinese children: a comparison of Han and Yao ethnic groups. Br J Clin Pharmacol 2004;58(2):163-8.

34.   Andersen JB, Szumlanski C, Weinshilboum RM, Schmiegelow K. Pharmacokinetics, dose adjustments, and 6-mercaptopurine/methotrexate drug interactions in two patients with thiopurine methyltransferase deficiency. Acta Paediatr 1998;87(1):108-11.

35.   Yates CR, Krynetski EY, Loennechen T, et al. Molecular diagnosis of thiopurine S-methyltransferase deficiency: genetic basis for azathioprine and mercaptopurine intolerance. Ann Intern Med 1997;126(8):608-14.

36.   Lennard L. Clinical implications of thiopurine methyltransferase--optimization of drug dosage and potential drug interactions. Ther Drug Monit 1998;20(5):527-31.

37.   Jenicek M. Identifying Cases of Disease: Clinimetrics and diagnosis. In: Jenicek M, ed. Epidemiology - The Logic of Modern Medicine. Montreal: Epimed International; 1995:81-114.

38.   Busse R, Orvain J, Velasco M, et al. Best practice in undertaking and reporting health technology assessments. Working group 4 report. Int J Technol Assess Health Care 2002;18(2):361-422.

39.   Granados A, Jonsson E, Banta HD, et al. EUR-ASSESS Project Subgroup Report on Dissemination and Impact. Int J Technol Assess Health Care 1997;13(2):220-86.

40.   Khan KS, Kunz R, Kleijnen J, Antes G, (Eds). Systematic Reviews to Support Evidence-Based Medicine: How to review and apply findings of healthcare research. London, UK: Royal Society of Medicine Press; 2007.

41.   Food and Drug Administration. Center for Drug Evaluation and Research. 2010. (Accessed February 2, 2010, at http://www.fda.gov/AboutFDA/CentersOffices/CDER/default.htm.)

42.   Public Health Resource Unit. Appraisal Tools: Diagnostic Test Studies. 2007. (Accessed February 23, 2008, at http://www.phru.nhs.uk/Pages/PHD/resources.htm.)

43.   van den Akker-van Marle ME, Gurwitz D, Detmar SB, et al. Cost-effectiveness of pharmacogenomics in clinical practice: a case study of thiopurine methyltransferase genotyping in acute lymphoblastic leukemia in Europe. Pharmacogenomics 2006;7(5):783-92.

44.   Sanderson J, Ansari A, Marinaki T, Duley J. Thiopurine methyltransferase: should it be measured before commencing thiopurine drug therapy? Ann Clin Biochem 2004;41(Pt 4):294-302.

45.   Marra CA, Esdaile JM, Anis AH. Practical pharmacogenetics: the cost effectiveness of screening for thiopurine s-methyltransferase polymorphisms in patients with rheumatological conditions treated with azathioprine. J Rheumatol 2002;29(12):2507-12.

46.   Colombel JF, Ferrari N, Debuysere H, et al. Genotypic analysis of thiopurine S-methyltransferase in patients with Crohn's disease and severe myelosuppression during azathioprine therapy. Gastroenterology 2000;118(6):1025-30.

47.   Ansari A, Hassan C, Duley J, et al. Thiopurine methyltransferase activity and the use of azathioprine in inflammatory bowel disease. Aliment Pharmacol Ther 2002;16(10):1743-50.

48.   Schwab M, Schaffeler E, Marx C, et al. Azathioprine therapy and adverse drug reactions in patients with inflammatory bowel disease: impact of thiopurine S-methyltransferase polymorphism. Pharmacogenetics 2002;12(6):429-36.

49.   Winter J, Walker A, Shapiro D, Gaffney D, Spooner RJ, Mills PR. Cost-effectiveness of thiopurine methyltransferase genotype screening in patients about to commence azathioprine therapy for treatment of inflammatory bowel disease. Aliment Pharmacol Ther 2004;20(6):593-9.

50.   Compagni A, Bartoli S, Buehrlen B, Fattore G, Ibarreta D, de Mesa EG. Avoiding adverse drug reactions by pharmacogenetic testing: a systematic review of the economic evidence in the case of TPMT and AZA-induced side effects. Int J Technol Assess Health Care 2008;24(3):294-302.

51.   Basu SK, Fernandez ID, Fisher SG, Asselin BL, Lyman GH. Length of stay and mortality associated with febrile neutropenia among children with cancer. J Clin Oncol 2005;23(31):7958-66.

52.   Hjalgrim LL, Rostgaard K, Schmiegelow K, et al. Age- and sex-specific incidence of childhood leukemia by immunophenotype in the Nordic countries. J Natl Cancer Inst 2003;95(20):1539-44.

53.   Mayo Clinic: Mayo Medical Laboratories. Reference Laboratory Services for Health Care Organizations. 2008. (Accessed October 18, 2008, at http://www.mayomedicallaboratories.com/.)

54.   Schedule of Benefits and Fees. 2008. (Accessed August 13, 2008, at http://www.health.gov.on.ca/english/providers/program/ohip/sob/sob_mn.html.)

55.   Ontario Ministry of Health and Long-Term Care. Schedule of Benefits. Laboratory Services. http://www.health.gov.on.ca/english/providers/program/ohip/sob/lab/lab_mn.html (Accessed Jan. 4, 2010) 2010.

56.   Ontario Case Costing Initative. Costing Analysis (CAT) Tool. 2010. (Accessed January 10, 2010, at http://www.occp.com/.)

57.   Relling MV, Hancock ML, Boyett JM, Pui CH, Evans WE. Prognostic importance of 6-mercaptopurine dose intensity in acute lymphoblastic leukemia. Blood 1999;93(9):2817-23.

58.   Winter JW, Gaffney D, Shapiro D, et al. Assessment of thiopurine methyltransferase enzyme activity is superior to genotype in predicting myelosuppression following azathioprine therapy in patients with inflammatory bowel disease. Aliment Pharmacol Ther 2007;25(9):1069-77.

59.   Kham SK, Soh CK, Liu TC, et al. Thiopurine S-methyltransferase activity in three major Asian populations: a population-based study in Singapore. Eur J Clin Pharmacol 2008;64(4):373-9.

60.   Schaeffeler E, Lang T, Zanger UM, Eichelbaum M, Schwab M. High-throughput genotyping of thiopurine S-methyltransferase by denaturing HPLC. Clin Chem 2001;47(3):548-55.

61.   Wusk B, Kullak-Ublick GA, Rammert C, von Eckardstein A, Fried M, Rentsch KM. Thiopurine S-methyltransferase polymorphisms: efficient screening method for patients considering taking thiopurine drugs. Eur J Clin Pharmacol 2004;60(1):5-10.

62.   www.xe.com - The World's Favorite Currency Site. 2010. (Accessed February 9, 2010, at http://www.xe.com/.)

63.   City Hospital Pathology Department, Birmingham. Thiopurine-S-Methyl Transferase (TPMT) Service. 2010. (Accessed February 3, 2010, at http://www.cityassays.org.uk/tpmt.html.)

64.   Wang LN, Zhang L, Nan F, et al. [HPLC determination and diversity analysis of thiopurine methyltransferase activity in human and pig]. Sichuan Da Xue Xue Bao Yi Xue Ban 2006;37(3):460-3.

65.   Coulthard SA, Rabello C, Robson J, et al. A comparison of molecular and enzyme-based assays for the detection of thiopurine methyltransferase mutations. Br J Haematol 2000;110(3):599-604.

66.   Dervieux T, Meyer G, Barham R, et al. Liquid chromatography-tandem mass spectrometry analysis of erythrocyte thiopurine nucleotides and effect of thiopurine methyltransferase gene variants on these metabolites in patients receiving azathioprine/6-mercaptopurine therapy. Clin Chem 2005;51(11):2074-84.

67.   Oselin K, Anier K, Tamm R, Kallassalu K, Maeorg U. Determination of thiopurine S-methyltransferase (TPMT) activity by comparing various normalization factors: reference values for Estonian population using HPLC-UV assay. J Chromatogr B Analyt Technol Biomed Life Sci 2006;834(1-2):77-83.

68.   Otterness D, Szumlanski C, Lennard L, et al. Human thiopurine methyltransferase pharmacogenetics: gene sequence polymorphisms. Clin Pharmacol Ther 1997;62(1):60-73.

69.   Relling MV, Hancock ML, Rivera GK, et al. Mercaptopurine therapy intolerance and heterozygosity at the thiopurine S-methyltransferase gene locus. J Natl Cancer Inst 1999;91(23):2001-8.

70.   Roberts RL, Barclay ML, Gearry RB, Kennedy MA. A multiplexed allele-specific polymerase chain reaction assay for the detection of common thiopurine S-methyltransferase (TPMT) mutations. Clin Chim Acta 2004;341(1-2):49-53.

71.   Heckmann JM, Lambson EM, Little F, Owen EP. Thiopurine methyltransferase (TPMT) heterozygosity and enzyme activity as predictive tests for the development of azathioprine-related adverse events. J Neurol Sci 2005;231(1-2):71-80.

72.   Salavaggione OE, Wang L, Wiepert M, Yee VC, Weinshilboum RM. Thiopurine S-methyltransferase pharmacogenetics: variant allele functional and comparative genomics. Pharmacogenet Genomics 2005;15(11):801-15.

73.   Alves S, Amorim A, Ferreira F, Prata MJ. Influence of the variable number of tandem repeats located in the promoter region of the thiopurine methyltransferase gene on enzymatic activity. Clin Pharmacol Ther 2001;70(2):165-74.

74.   Anglicheau D, Sanquer S, Loriot MA, Beaune P, Thervet E. Thiopurine methyltransferase activity: new conditions for reversed-phase high-performance liquid chromatographic assay without extraction and genotypic-phenotypic correlation. J Chromatogr B Analyt Technol Biomed Life Sci 2002;773(2):119-27.

75.   Indjova D, Atanasova S, Shipkova M, Armstrong VW, Oellerich M, Svinarov D. Phenotypic and genotypic analysis of thiopurine s-methyltransferase polymorphism in the bulgarian population. Ther Drug Monit 2003;25(5):631-6.

76.   Loennechen T, Utsi E, Hartz I, Lysaa R, Kildalsen H, Aarbakke J. Detection of one single mutation predicts thiopurine S-methyltransferase activity in a population of Saami in northern Norway. Clin Pharmacol Ther 2001;70(2):183-8.

77.   Lu Y, Kham SK, Tan PL, Quah TC, Heng CK, Yeoh AE. Arrayed primer extension: a robust and reliable genotyping platform for the diagnosis of single gene disorders: beta-thalassemia and thiopurine methyltransferase deficiency. Genet Test 2005;9(3):212-9.

78.   Nasedkina TV, Fedorova OE, Glotov AS, et al. Rapid genotyping of common deficient thiopurine S-methyltransferase alleles using the DNA-microchip technique. Eur J Hum Genet 2006;14(9):991-8.

79.   Reis M, Santoro A, Suarez-Kurtz G. Thiopurine methyltransferase phenotypes and genotypes in Brazilians. Pharmacogenetics 2003;13(6):371-3.

80.   Bachmann LM, Puhan MA, ter Riet G, Bossuyt PM. Sample sizes of studies on diagnostic accuracy: literature survey. British Medical Journal 2006;332(7550):1127-9.

81.   Oh KT, Anis AH, Bae SC. Pharmacoeconomic analysis of thiopurine methyltransferase polymorphism screening by polymerase chain reaction for treatment with azathioprine in Korea. Rheumatology (Oxford) 2004;43(2):156-63.

82.   Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. Publication bias in clinical trials due to statistical significance or direction of trial results. Cochrane Database Syst Rev 2009(1):MR000006.

83.   Tricco AC, Tetzlaff J, Sampson M, et al. Few systematic reviews exist documenting the extent of bias: a systematic review. J Clin Epidemiol 2008;61(5):422-34.

84.   National Cancer Institute. Educational Materials about Clinical Trials. 2009. (Accessed March 3, 2009, at http://www.cancer.net/patient/Learning+About+Cancer/Clinical+Trials/About+Clinical+Trials.)

85.   Griebsch I, Coast J, Brown J. Quality-adjusted life-years lack quality in pediatric care: a critical review of published cost-utility studies in child health. Pediatrics 2005;115(5):e600-14.

86.   Schmiegelow K, Forestier E, Kristinsson J, et al. Thiopurine methyltransferase activity is related to the risk of relapse of childhood acute lymphoblastic leukemia: results from the NOPHO ALL-92 study. Leukemia 2009;23(3):557-64.

87.   Stanulla M, Schaeffeler E, Flohr T, et al. Thiopurine methyltransferase (TPMT) genotype and early treatment response to mercaptopurine in childhood acute lymphoblastic leukemia. Jama 2005;293(12):1485-9.

88.   Evans WE, Horner M, Chu YQ, Kalwinsky D, Roberts WM. Altered mercaptopurine metabolism, toxic effects, and dosage requirement in a thiopurine methyltransferase-deficient child with acute lymphocytic leukemia. J Pediatr 1991;119(6):985-9.

89.   Lennard L, Gibson BE, Nicole T, Lilleyman JS. Congenital thiopurine methyltransferase deficiency and 6-mercaptopurine toxicity during treatment for acute lymphoblastic leukaemia. Arch Dis Child 1993;69(5):577-9.

90.   Lennard L, Lewis IJ, Michelagnoli M, Lilleyman JS. Thiopurine methyltransferase deficiency in childhood lymphoblastic leukaemia: 6-mercaptopurine dosage strategies. Med Pediatr Oncol 1997;29(4):252-5.

91.   Stocco G, Cheok MH, Crews KR, et al. Genetic polymorphism of inosine triphosphate pyrophosphatase is a determinant of mercaptopurine metabolism and toxicity during treatment for acute lymphoblastic leukemia. Clin Pharmacol Ther 2009;85(2):164-72.

Glossary

Clinical Remission: The eradication of leukemic cells to a clinically undetectable level.

Cytotoxic: A substance that causes cell destruction

Genotype: The genetic make-up of an individual

Granulocytopenia: Decreased leukocytes

Hematologic Remission: The eradication of all leukemic cells.

Hematopoiesis: The process that immature blood cells go through to produce the various types of blood cells

Hepatomegaly: Enlarged liver

Hepatotoxicity: Damage to the liver

Intrathecal: Injection onto the spine

Myelosuppression: Decrease bone marrow activity, resulting in less red blood cells white blood cells and platelets

Petechiae: Small hemorrhagic spots on the skin

Pharmacogenetics: The study of genetic variations that evoke differing responses to drug therapy

Phenotype: Physical or biochemical characteristics as determined by genetic makeup and environmental influences.

Splenomegaly: Enlarged spleen

Therapeutic Index: The ratio between the toxic dose and the therapeutic dose of a drug, used as a measure of the relative safety of the drug for a particular treatment.

Thrombocytopenia: Decreased platelets