SickKids®

The Hospital for Sick Children

Technology Assessment at SickKids (TASK)

THIOPURINE S-METHYLTRANSFERASE TESTING FOR AVERTING DRUG TOXICITY IN PATIENTS RECEIVING THIOPURINES: A META-ANALYSIS OF DIAGNOSTIC TEST ACCURACY

Report No. 2015-03

October 7, 2015

REPORT HIGHLIGHTS

The Report Highlights consists of a summary of the full report with the same name and should be evaluated in conjunction with the full report and its appendices. Full documents are available for download at:

http://lab.research.sickkids.ca/task/reports-theses/

Authors

Richard M. Zur, PhD, Child Health Evaluative Sciences, The Hospital for Sick Children

Lilla M. Roy, RN, BScN, MSc, Child Health Evaluative Sciences, The Hospital for Sick Children

Wendy J. Ungar, MSc, PhD, Child Health Evaluative Sciences, The Hospital for Sick Children

Co-Investigators

Shinya Ito, MD, FRCPC, Clinical Pharmacology and Toxicology, The Hospital for Sick Children, Pharmacology & Pharmacy, Department of Paediatrics, University of Toronto

Elizabeth Uleryk, MLS, E.M The Hospital for Sick Children

Joseph Beyene, MSc, PhD, Department of Clinical Epidemiology & Biostatistics, McMaster University

Knowledge user partners:

Chris Carew, MBA, Centre for Genetic Medicine, The Hospital for Sick Children

James Whitlock, MD, Division Head/Chief Haematology/Oncology, The Hospital for Sick Children; Professor, Paediatrics, University of Toronto

Acknowledgements

This research was supported by a Canadian Institutes of Health Research Knowledge Synthesis Grant, grant #305352.

We thank Mr. Ian Schiller, M.Sc., Division of Clinical Epidemiology, McGill University Health Centre and Dr. Nandini Dendekuri, Ph.D., Director, Technology Assessment Unit of the McGill University Health Centre for their technical assistance with this report. We thank Ms. Christine Millan for administrative assistance.

The views expressed in the material are the views of the authors and do not necessarily reflect those of The Hospital for Sick Children or the province of Ontario.

Conflicts of interest

The authors have no conflicts of interest to disclose.

Introduction

Thiopurine S-methyltransferase (TPMT) is an enzyme that metabolizes thiopurine drugs used to treat childhood leukemia, as well as inflammatory bowel disease (IBD) and dermatological conditions. A deficiency in TPMT activity can significantly increase the risk of an adverse drug event (ADE). Unless thiopurine drugs are avoided or doses are reduced in these patients, they are at greater risk for life-threatening bone marrow toxicity and liver toxicity, which may lead to myelosuppression, anemia, bleeding, leukopenia, infection and death.[1] There are currently two methods of TPMT testing – a phenotype or a genotype test. Although many studies have assessed the accuracy of both methods, it is uncertain which test is superior.

Rationale

Meta-analysis of the accuracy of the two methods of TPMT testing documented in numerous studies would provide a valuable summary estimate of performance. Previous meta-analysis has been limited by the technical challenges of pooling diagnostic test accuracy (DTA) results and the lack of a gold reference standard.

Key Messages

- Traditional phenotype testing has been joined by genotype testing and a personalized medicine approach in clinical decision making for selection of drug treatment and doses.
- Testing to detect a deficiency in thiopurine s-methyltransferase (TPMT) activity can be performed using either a phenotype or genotype approach. It is unclear which test is superior.
- Meta-analyses have been limited due to technical challenges with pooling diagnostic test accuracy (DTA) results and lack of a gold reference standard.
- Advanced methods using hierarchical summary receiver operating characteristics (HSROC) and latent class analysis are recommended.
- Estimates of sensitivity and specificity of phenotyping for patients with deficient TPMT activity was 75.9% (95% CI, 58.3% to 87.0%) and 98.9% (95% CI, 96.3% to 100%) respectively.
- Sensitivity and specificity of genotyping to detect multiple polymorphisms in deficient individuals was 80.7% (95% CI, 41.7% to 99.4%) and 99.9% (95% CI, 99.7% to 100%), respectively.
- Pooled estimates of DTA can facilitate the selection of diagnostic test for optimal management of patients with TPMT deficiency.

Objectives

The objective was to meta-analyze the sensitivity and specificity of phenotype and genotype TPMT testing reported in the literature using two methods of evaluation and a DTA method that accounts for the imperfect reference standard provided by genotype testing.

Target Population

This study included evidence from all patient populations, age groups and clinical indications for which TPMT testing can be used.

Methods

Systematic review and quality appraisal

The first phase of the research elicited data on genotype test and phenotype test accuracy from studies identified through a comprehensive systematic review and quality appraisal [2]. Two different testing approaches were used: 1) tests identifying patients with deficient or absent TPMT enzyme activity (patients that are homozygous for TPMT mutations) versus the rest of the population, and 2) tests identifying patients that have either low or intermediate TPMT enzyme activity (patients that are homozygous or heterozygous for TPMT mutations) versus the rest of the population.

Meta-analysis

The second phase of the study used a meta-analysis conducted using a hierarchical summary receiver operating characteristic (HSROC) approach. A latent class meta-analysis method that allowed for heterogeneity in cut-point definition in phenotype TPMT testing while also allowing for an imperfect reference standard was used to meta-analyze the sensitivity and specificity data for the two testing approaches.[3]

Results

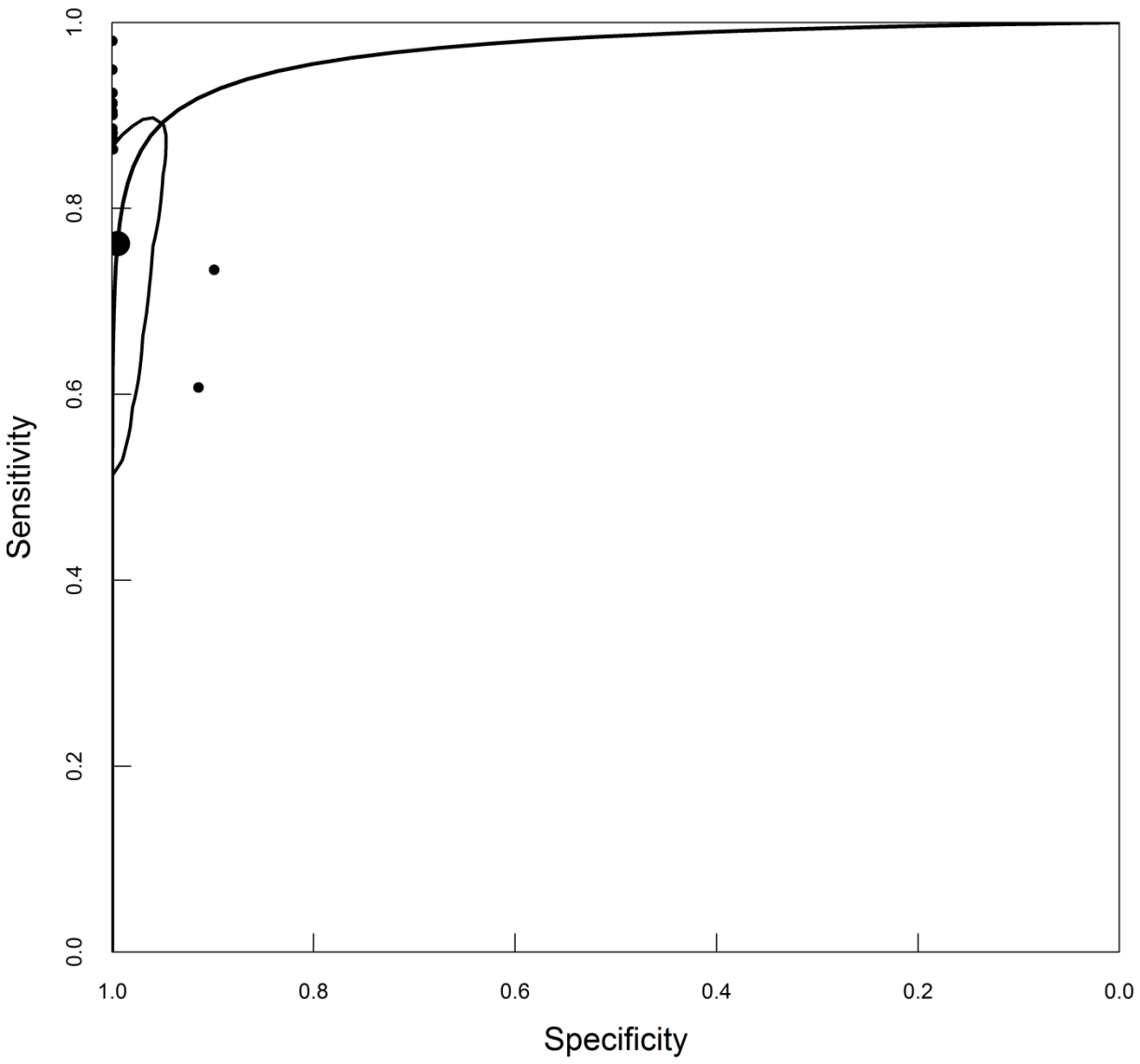
The latent class analysis of 13 studies resulted in a pooled sensitivity and specificity of phenotype testing for patients with deficient or absent TPMT enzyme activity (patients who are homozygous for TPMT mutations) of 75.9% (95% CI, 58.3% to 87.0%) and 98.9% (95% CI, 96.3% to 100%) respectively. For genotype testing evaluating only the most common TPMT*2 and TPMT*3 polymorphisms, the pooled sensitivity and specificity was 80.75 (95% CI, 79.1% to 99.4%) and 99.9% (95% CI, 99.7% to 100%), respectively. When evaluating more than TPMT*2 and TPMT*3, the pooled estimates were 80.7% (95% CI, 41.7% to 99.4%) and 99.9% (95% CI, 99.7% to 100%), respectively, for sensitivity and specificity.

For the second group which detected deficient or intermediate TPMT activity (homozygous or heterozygous TPMT mutations), the pooled sensitivity and specificity of phenotype testing from 27 studies was 91.3% (95% CI, 86.4% to 95.5%) and 92.6% (95% CI, 99.5% to 100%), respectively. For genotype testing evaluating TPMT*2 and TPMT*3 only, sensitivity and specificity estimates were 88.9% (95% CI, 81.6% to 97.5%) and 99.2% (95% CI, 998.4% to 99.9%), respectively. When more polymorphisms were considered, the pooled estimates were 93.5% (95% CI, 84.9% to 99.3%) and 99.9% (95% CI, 99.7% to 100%), respectively.

Discussion

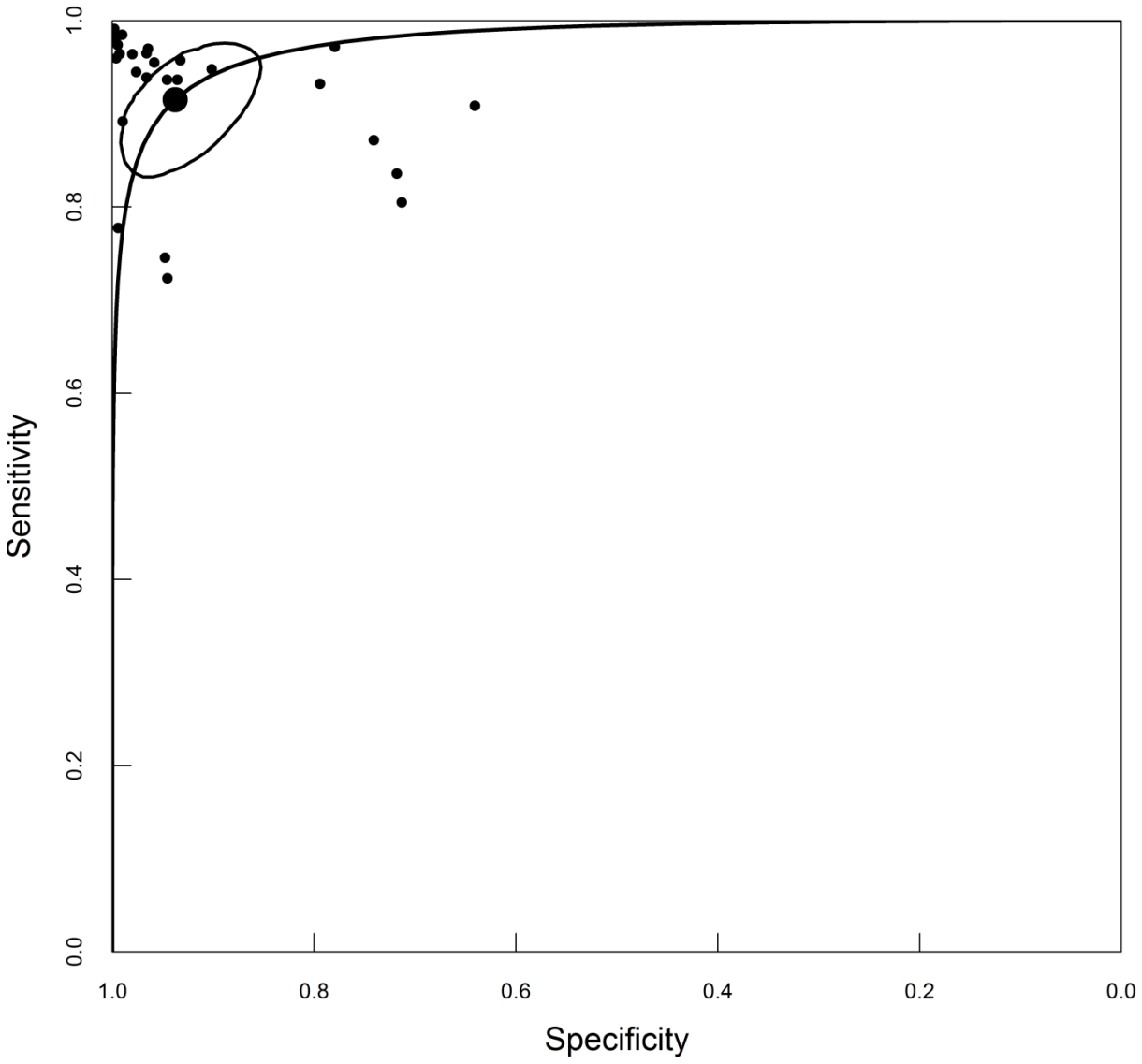
Higher values for specificity for both testing approaches indicate the value of testing for ruling in TPMT deficiency. The results of the meta-analysis indicate that genotype testing has higher sensitivity than phenotype testing, and sensitivity increases with more polymorphisms included in the test. This meta-analysis does not conclude that one test is superior to the other. Although there are more complex approaches, the methods used (latent class HSROC) is a straight-forward approach to provide a single pooled estimate for DTA meta-analysis.

Figure 1. Hierarchical summary ROC curve for the phenotype test discriminating deficient TPMT individuals versus others



The SROC curve was estimated from a latent class meta-analysis model of 13 studies assuming imperfect reference standards. Small dots represent the sensitivity and specificity of individual studies and the large dot represents the pooled sensitivity and specificity. The ellipse around the pooled sensitivity and specificity represented the 95% credible region for the pooled sensitivity and specificity.

Figure 2. Hierarchical summary ROC curve for the phenotype test discriminating individuals with T deficient or intermediate TPMT activity versus others



The SROC curve is estimated from a latent class meta-analysis model of 27 studies assuming imperfect reference standards. Small dots represent the sensitivity and specificity of individual studies and the large dot represents the pooled sensitivity and specificity. The ellipse around the pooled sensitivity and specificity represented the 95% credible region for the pooled sensitivity and specificity.

REFERENCES

1. Baker GR, Norton PG, Flintoft V, et al. The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. Canadian Medical Association Journal 2004;170:1678-86.
2. Roy LM, Ungar WJ, Zur RM. Thiopurine S-methyltransferase testing for averting drug toxicity in patients receiving thiopurines: A systematic review and quality appraisal; 2015 March 26, 2015. Report No.: 2015-02.
3. Dendukuri N, Schiller I, Joseph L, Pai M. Bayesian meta-analysis of the accuracy of a test for tuberculous pleuritis in the absence of a gold standard reference. Biometrics 2012;68:1285-93.