Pharmacogenomics. Author manuscript; available in PMC 2016 July 04.

Published in final edited form as:
Pharmacogenomics. 2016 April ; 17(6): 633–656. doi:10.2217/pgs.16.12.

Thiopurine S-methyltransferase testing for averting drug toxicity in patients receiving thiopurines: a systematic review

Lilla M Roy[1], **Richard M Zur**[1], **Elizabeth Uleryk**[2], **Chris Carew**[3], **Shinya Ito**[4,5], and **Wendy J Ungar**[*,1,6]

[1]Child Health Evaluative Sciences, The Hospital for Sick Children Research Institute, Toronto, Canada

[2]Library Services, The Hospital for Sick Children, Toronto, Canada

[3]Centre for Genetic Medicine, The Hospital for Sick Children, Toronto, Canada

[4]Division of Clinical Pharmacology & Toxicology, The Hospital for Sick Children, Toronto, Canada

[5]Departments of Pharmacology & Pharmacy, Faculty of Medicine Department of Paediatrics, University of Toronto, Canada

[6]Institute for Health Policy, Management & Evaluation, University of Toronto, Canada

Abstract

Aim—Thiopurine S-methyltransferase (TPMT) testing is used in patients receiving thiopurines to identify enzyme deficiencies and risk for adverse drug reactions. It is uncertain whether genotyping is superior to phenotyping. The objectives were to conduct a systematic review of TPMT-test performance studies.

Materials & methods—Electronic and grey literature sources were searched for studies reporting test performance compared with a reference standard. Sixty-six eligible studies were appraised for quality.

Results—Thirty phenotype–genotype and six phenotype–phenotype comparisons were of high quality. The calculated sensitivity and specificity for genotyping to identify a homozygous mutation ranged from 0.0–100.0% and from 97.8–100.0%, respectively.

Conclusion—Clinical decision-makers require high-quality evidence of clinical validity and clinical utility of TPMT genotyping to ensure appropriate use in patients.

Keywords

enzyme deficiency; genetic test; phenotype; QUADAS-2; quality appraisal; sensitivity; specificity; systematic review; thiopurine s-methyltransferase

[*]Author for correspondence: Tel.:+14168137654, ext.303487, Fax: +14168135979, wendy.ungar@sickkids.ca.

Financial & competing interests disclosure
This study was supported by a Knowledge Synthesis grant from the Canadian Institutes of Health Research (grant #304352). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.
No writing assistance was utilized in the production of this manuscript.

With advances in the field of pharmacogenomics, it is increasingly common to use genetic or biomarker testing to predict an individual's drug responses [1]. This personalized medicine approach allows for more accurate selection of treatments as well as dosing of prescription medicines and the avoidance of potentially serious life-threatening adverse drug events (ADEs). The technologies that are used to test for drug metabolizing enzyme activity and for the presence of genetic variants that affect drug metabolism are rapidly evolving with regard to technical methods and scope [2]. This introduces uncertainty for clinical practitioners regarding which tests to use for their patients, and for health plan decision-makers regarding the value for money of new personalized medicine technologies.

One of the more common applications of personalized medicine is testing for deficiency in thiopurine s-methyltransferase (TPMT), an enzyme that metabolizes thiopurines [3]. The clinical consequences of deficient TPMT activity are significant. Unless thiopurine drug doses are reduced in these patients, they are at greater risk for life-threatening bone marrow toxicity, which may lead to myelosuppression, anemia, bleeding, leukopenia, infection and death and potentially life-threatening pancreatitis [4].

There are two approaches to testing for TPMT deficiency. Phenotype tests that measure levels of TPMT enzyme activity *in vitro* are common, but test results can be confounded by concomitant medications or blood transfusions [2,5–11]. Genotype tests are available that detect the presence of variants in the genes responsible for expressing the TPMT enzyme [12–16]. It remains uncertain whether an enzyme activity (phenotype) or genotype diagnostic test is the most appropriate strategy for clinical practice.

This uncertainty is especially true in the pediatric population, where thiopurine doses are calculated based on weight, and ADEs may result in significant morbidity [17]. A recent systematic review of clinical guidelines on TPMT testing revealed wide differences in testing recommendations as well as differences in thiopurine dosing recommendations for patients with identified TPMT deficiencies [18]. Improving information regarding the clinical validity and performance characteristics of alternative TPMT testing strategies will facilitate testing decision-making and treatment with thiopurines. The research objectives were to systematically review the literature on the performance characteristics of phenotype and genotype testing for TPMT deficiency and to appraise the quality of the TPMT testing literature.

Methods

Systematic review

Inclusion & exclusion criteria—Eligible studies were those conducted in humans that evaluated either a TPMT genotype or TPMT phenotype test compared with a reference standard, where the reference standard was as another phenotype or genotype test such that the comparison could be phenotype–phenotype, genotype–genotype or phenotype–genotype. Studies had to provide results or raw data permitting construction of contingency tables for calculation of sensitivity, specificity, negative predictive value (NPV), positive predictive

value (PPV) or concordance. Studies were not restricted based on age, disease group or language. Additional information is available in the full technical report [19].

Literature search—Electronic citation databases were searched, including Biosis, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Cochrane Database of Systematic Reviews (CDSR), Cochrane Central Register of Controlled Trials (CCTR), Database of Abstracts of Reviews of Effects (DARE), Health Technology Assessment (HTA), National Health Service Economic Evaluation Database (NHSEED), EMbase, International Pharmaceutical Abstracts (IPA), Medline and PubMed. Eligible grey literature was obtained directly from websites of government health agencies, health technology assessment agencies, health economic research groups, research institutes, academic organizations and from websites related to the diseases of interest that are treated with thiopurines. Search strategies were developed and terms were selected for each database in collaboration with a librarian experienced in systematic reviews and an experienced health technology assessment research team. The most comprehensive search strategy combined the search concepts in the following manner: TPMT (or related terms) and a thiopurine drug (common thiopurine drugs such as AZA, 6-mercaptopurine and thioguanine), and either a phenotype or genotype technology. This combination of terms maintained relatively high specificity for well-known studies, with 16/17 of previously identified studies detected in the results (Supplementary table 1).

Review for eligibility—Two reviewers (RM Zur and LM Roy) performed the screening and selection of studies. Discrepancies were resolved by establishing a set of decision rules, in consultation with the principal investigator (WJ Ungar) as needed. Agreement became consistent after comparing independent categorization of approximately 130 abstracts and titles between the two reviewers. Subsequently, one reviewer (LM Roy) screened the remaining titles and abstracts. A reference manager software program (EndNote X4, PA, USA) was used to maintain reference citations. Relevant full text articles were retrieved where possible through interlibrary loan or requested directly from the author, and reviewed for inclusion according to established decision rules.

Translation was required for papers published in Chinese, Dutch, French, German, Japanese, Korean, Polish, Serbian and Spanish. University student translators worked with a research team member to review eligible non-English publications.

Data extraction—A data extraction tool was created using Microsoft Access (version 2010) to ensure consistent abstraction of relevant data from each study, including study design, study sample and test characteristics. The particular ethnic groups studied were also recorded as the incidence of variants is correlated with ethnicity and this can affect calculations of positive and negative predictive value, and the stability of calculations of sensitivity and specificity. The alleles included in the genotype tests were recorded using standard nomenclature [20]. If test performance results including sensitivity, specificity, PPV, NPV and concordance were reported, then they were abstracted as reported by the authors. In addition, 2 × 2 or 3 × 3 contingency tables were populated for each included study using data reported in tables, text or inferred from graphs, to allow reviewers to calculate test performance characteristics independently. As no gold standard reference test

exists, for the purpose of calculation standardization, the phenotype TPMT test was designated as the reference test and the genotype as the index test for all statistical calculations of sensitivity and specificity, since the latter represents the innovative technology. For the purpose of this report, ‘absent’ and ‘deficient’ activity were considered equivalent to ‘low’-enzyme activity. The term ‘intermediate’ was used to describe intermediate-enzyme activity. The terms ‘high’ activity and ‘normal’ were both interpreted to represent the upper spectrum of enzyme activity, which was categorized as ‘high/normal’ (presumed wild-type genotype).

Reviewers first classified TPMT activity into 3 × 3 tables after considering the cutpoints reported by the study author, text descriptions and the distribution of the TPMT activity results (e.g., graphical results). ‘Low’ activity included reported ‘deficient’ or ‘absent’ activity, or enzymatic activity below approximately 5 U/ml packed red blood cells (pRBC). Reported activity above 5 U/ml pRBC and below approximately 10 U/ml pRBC was categorized as ‘intermediate’ activity. Enzymatic activity reported above 10 U/ml pRBC was classified as ‘high/normal’. These activity levels reflect a common classification initially reported by Weinshilboum *et al.* [21]. Where 3 × 3 tables were not possible, 2 × 2 tables were populated.

Quality appraisal

The Quality Assessment tool for Diagnostic Accuracy Studies (QUADAS) version 2 was used to evaluate the quality of included studies [22]. The QUADAS-2 contains four domains pertaining to risk of bias and applicability related to patient selection; the index test; the reference standard; and flow and timing of the study. A fifth domain (described below) was created for the purpose of this study to assess the risk of bias pertaining specifically to genomic tests. A study-specific QUA-DAS-2 appraisal tool was created in Microsoft Access (version 2010) by tailoring items to the study objectives to ensure consistent and reliable assessment between reviewers.

An overall determination of high versus low quality of included studies was made based on a pre-established algorithm created by the reviewers and reviewed for consistency until consensus was reached. Studies were considered to be of high quality if all five QUADAS domains demonstrated low bias and had low concern for applicability. If only one domain demonstrated high risk of bias, then the study was considered to be of high quality overall. If the study had two or more domains that were of high or uncertain bias, then the study was deemed as low quality overall.

Results

Systematic review

The search results are displayed in Figure 1. The search yielded 4071 publications from the database and grey literature sources. After the removal of duplicates, screening of the titles and abstracts of 2088 records resulted in 374 full text papers, which were screened for eligibility, including 37 requiring translations from Korean, German, Polish, French, Japanese, Chinese, Dutch, Spanish and Serbian. One hundred and twenty-one papers

appeared to meet inclusion criteria and were assessed for relevant data. Of these, 55 had insufficient data to populate contingency tables resulting in 66 papers included in the review.

Of those papers with sufficient data, 55 reported a phenotype–genotype comparison [5–11,13,14,16,23–67]. The remaining 11 papers reported a laboratory method comparison (either phenotype–phenotype or genotype–genotype) [6,7,15,67–74]. Studies comparing phenotype and genotype testing were published between 1996 and 2014. Studies comparing phenotype–phenotype or genotype–genotype were published between 1994 and 2013. Among the 66 eligible studies, sample sizes ranged from 15 [57] to 7195 [30]. Sixteen studies were conducted in adults, 11 in children, 13 in a mix of adult and pediatric populations, and the remaining 26 did not specify the sample age. Fourteen studies were conducted in healthy populations while 51 studies sampled patients, including 14 studies in acute lymphoblastic leukemia (ALL), 15 in inflammatory bowel disease (IBD), six that were not specified, 13 with ‘other’ patients, one with dermatological conditions and two with organ transplant patients. Only one study did not specify the disease population [68]. The prevalence of variants is known to vary by ethnic group. Many studies identified a particular ethnicity, race or nationality. Caucasian (n = 11) was the most commonly identified group, in whom *TPMT*2* and *TPMT*3* are the most common variants. This was followed by Chinese (n = 4), European (n = 5) and German (n = 1). Authors did not commonly identify whether participants were related to one another; only 18 studies reported that participants were unrelated.

Quality appraisal

Phenotype–genotype comparisons—Of the 55 papers with sufficient data to calculate sensitivity and specificity, 30 studies were of high quality (Table 1). Seven studies demonstrated ‘high’ or ‘unclear’ concern regarding applicability for at least one of the five domains. Fifteen of 30 high-quality studies showed ‘high’ or ‘unclear’ risk of bias for at least one of the five domains. Thirteen of the studies consistently demonstrated low scores (low risk of bias, low concern for applicability). Low-quality studies generally had more ‘unclear’ ratings than high-quality studies, as opposed to definitive high risk of bias ratings. Only nine low-quality studies were deemed of low quality due to two or more high risk of bias or concern for applicability. The remaining 16 studies had at least one element that was considered ‘unclear’ in addition to one or more elements of high risk of bias or ‘unclear’ risk or concern for applicability.

Among the 25 low-quality studies, the highest risk of bias was observed for Domain 4 (genomics), with 12 studies appraised as having high risk of bias. A high risk of bias was next most frequent in Domain 3 (reference test), with seven studies so categorized. For the domains reporting ‘unclear’ risk of bias or applicability, the most problematic domain was Domain 3 (Reference test) with 12 studies having insufficient information to determine whether bias was high or low. Concern for applicability was highest in Domain 3.

Genotype–genotype & phenotype–phenotype comparisons—Among studies comparing genotype–genotype or phenotype–phenotype tests, six were found to be of high quality (Table 1) and all were genotype–genotype test comparisons. Only one study

demonstrated low risk of bias and low concern for applicability in all domains [74]. The five low-quality studies did not have clear patterns of bias risk or concern for applicability. Two studies were phenotype–phenotype studies, and therefore, Domain 5 (genomics) did not apply.

Design characteristics of high-quality studies

Study objectives & eligibility criteria—High-quality phenotype–genotype studies were published between 1997 and 2013. Eleven studies stated their primary objective was to investigate the relationship (e.g., concordance) between phenotype and genotype testing for TPMT activity determination [11,14,25,27–30,32,37,39,62]. Two studies explicitly stated that investigating this relationship was a secondary objective [24,34] (Table 2).

Inclusion criteria for individual studies often specified that participants should meet specific disease criteria: healthy [33–35], ALL [16,25,32], IBD [11,29,64,65], transplant [14,36] and renal failure [37]. One study specified pediatric patients as an inclusion criterion [25]. Common exclusion criteria were a history of blood transfusions [26,31,32], concurrent medications such as methotrexate [14,26,58], insufficient functioning of major organs [26] and concurrent or history of a variety of acute, chronic or genetic diseases [14,35,37,58]. One study specified that blood samples >8 days old were excluded [27]. Most studies did not specify exclusion criteria [10,11,13,16,23–25,28–30,33,34,36,38,39,60,62,64–67]. Recruitment of patients and conduct of studies ranged from a 4-week period to over 7 years. Most studies did not specify the time period during which patients were recruited.

Sample characteristics—Sample sizes for high-quality phenotype–genotype studies ranged from 35 to 7195 individuals (Table 2). Four studies reported samples as pediatric [16,25,29,32] and 11 studies did not specify the age of the samples [11,13,24,27,28,30,33,36,60,67,77]. The remaining studies reported either adult or a mix of adult and pediatric samples. Race was not always specified [10,14,27–30,36,37,59], but several high-quality studies identified their sample as, for example, Caucasian, Scandinavian or from the United Kingdom [16,23,31,32,58,60,62,64–66].

Sample sizes for high-quality genotype–genotype studies ranged from 80–630 (Table 2). Two studies included a mix of children and adults [70,76], while one included adults only [15]. The remaining three did not specify the age group [71,74,75]. None of the studies specified the mean age of their subjects. One study was composed of IBD patients [71] and one did not specify a disease group [74]. The remaining studies had a variety of subjects including ALL, otherwise healthy blood donors and unspecified patients who were undergoing thiopurine treatment or who had TPMT testing requested. One study had Chinese subjects [70], two had Caucasian subjects [15,74] and the other three did not specify a race or ethnicity [71,75,76]. Only two studies specified that subjects were unrelated [70,74].

Laboratory test methods

Genotyping—Genotyping studies employed similar DNA amplification methods, with 80% (24/30) using PCR, 26% (8/30) allele-specific PCR (AS-PCR) and 6% (2/30) PCR-

single strand conformation polymorphism (PCR-SSCP). Methods such as denaturing HPLC (DHPLC) [62,64], multiplex amplification refractory mutation (ARMS) [27,28], pyrosequencing [16,30] and TaqMan SNP genotyping [14,33] were reported. Direct sequencing (n = 3) [11,24,67] and RFLP (n = 17) [10,13,23,26,29,31–39,59,60,66] were also reported. Only one study did not specify a genotyping method [65].

There were nine different methods of genotype testing reported (Tables 3 & 4). These included pyrosequencing (2/30), RFLP (includes restriction mapping, restriction analysis or restriction digestion) (17/30), DHPLC (2/30), AS-PCR (8/30), direct sequencing (3/30), PCR-SSCP (2/30), ARMS (2/30), PCR (24/30) and TaqMan methods (2/30). Twenty-six studies reported more than one method of genotyping, and one study did not report any method [65].

For the six high-quality genotype–genotype studies, test methods varied and included RFLP [70,71,76], arrayed primer extension technology (APEX) [76], ARMS-PCR [76], AS-PCR [70,71,75], DHPLC [70,74], LightSNiP [15], MALDI-TOF-mass spectrometry [74], PCR [15,70,75], SNaPshot™ (Thermo Fisher Scientific, MA, USA) sequencing [70] and TaqMan SNP genotyping [71]. Microchip RFLP and AS-PCR technologies were investigated in one study [71], and two studies referred to ‘sequencing’ as the genotyping method [15,70] (Table 4).

All but two quality phenotype–genotype comparisons tested for (at least) *TPMT*2* and *TPMT*3* (Table 3). The outlying studies did not test for *TPMT*2*, only *TPMT*3* [32,59]. Although this increased bias, the rest of the study characteristics were considered of high quality according to the QUADAS-2. All but one genotype–genotype study investigated at least *TPMT*2* and *TPMT*3*, the most common polymorphisms [15,70,71,74,75]. One study investigated nearly all of the known TPMT polymorphisms, ten in total [74].

Phenotyping—Phenotype test methods included radiochemical method (11/30), HPLC (13/30), competitive microwell immunoassay (1/30) and mass spectrometry (1/30), with four studies unclear about the method used (Table 3). There were also assay variations such as MS/MS and modifications to the traditional radiochemical assay. Measurement units for reporting enzyme activity varied across studies. Enzyme activity was most commonly measured per milliliter of pRBCs (U/ml pRBCs). Variation in units made direct comparison of enzyme activity cutpoints across studies difficult.

In general, TPMT activity was classified by authors as low, intermediate or high. However, terminology and classification of activity levels were inconsistent, with some studies using ‘deficient’ while others used ‘low’. Some studies adding a category of ‘very high’, and some studies used ‘normal’ in place of ‘high’. Tables 3 & 5 describe the phenotype test characteristics.

The choice of cutpoint to distinguish between activity levels was generally cited from previous research, although some authors calculated their own cutpoints after sample collection and analysis. Typically this was in the form of an receiver operating characteristic analysis [26,27,33,34,37]. The conventional classification system developed by

Weinshilboum *et al.* [21] classifies phenotype activity as deficient (<5 U/ml red blood cell [RBC]), intermediate (5–10 U/ml pRBC) and normal (>10 U/ml pRBC). This classification was used in three studies [13,23,66]. It was not clear whether cut-points varied by any particular study characteristic or population. For example, the cutpoint between intermediate and high-enzyme activity for ALL patients varied from 9 to 12 U/ml pRBCs, while the cutpoint between intermediate and low varied between 2.5 and 6 U/ml pRBCs. For patients with IBD, the cutpoint between intermediate and high enzyme activity varied between 8 and 45.5 nmol 6-MTG/gHb/h, or 4.75 and 15.5 U/ml RBC and the cutpoint between low and intermediate varied between 2.5 and 5.6 U/ml RBC. In contrast to these values, one study reported a cut-point of 25 between intermediate and high enzyme activity and a cutpoint of 10 between low and intermediate; however, the unit of this test was specified as picomoles [10]. Further, some studies did not specify the unit of measure. Cutpoints used for each study are presented in Supplementary Table 2.

Diagnostic test performance characteristics

Diagnostic test performance characteristics such as sensitivity, specificity, NPV and PPV were infrequently reported explicitly in studies comparing two tests, although a concordance rate was commonly reported (Supplemental Table 3). Using data from the high-quality phenotype–genotype publications, the sensitivity, specificity, NPV, PPV and concordance were calculated with genotyping as the index test and phenotyping as the reference standard. Table 5 presents test performance characteristics for genotyping when deficient was defined as the absence of TPMT activity (suggesting the presence of a homozygous mutation).

Fifteen studies provided data sufficient to calculate sensitivity for detection of a homozygous mutation [10,13,24,25,27,29–31,35,58,59,62,64,66,67]. Due to the absence of homozygous deficient patients (cell count of zero), it was not possible to calculate sensitivity and specificity in all studies. Calculated sensitivity of genotyping from these 15 studies ranged from 0.0 to 100.0% and with data that were available from 26 studies, specificity ranged from 97.8 to 100.0%.

Ten of the 15 studies with sufficient data had 100.0% for both values [13,24–25,27,35,58,62,64,66–67]. The other five studies only investigated *TPMT*2* and *TPMT*3*, although half of those studies with 100.0% calculated values also were limited to these polymorphisms [13,25,27,35,58]. The two studies with a sensitivity of 0.0% were conducted in samples of 130 (persons with positive test for low enzyme activity = 1; persons with negative test for low enzyme activity = 129) [10] and 53 (persons with positive test for low enzyme activity = 1; persons with negative test for low TPMT activity = 52) [59]. Most studies with calculated sensitivity and specificity of 100.0% generally had large sample sizes (n = 88–1214) with the number of persons with positive tests for low enzyme activity ranging from 1–7, and the number of persons with negative tests for low enzyme activity ranging from 34 to 1207. The largest study [30] had a sensitivity of 86.0% and tested for all polymorphisms.

Four of five studies with imperfect sensitivity and specificity did not specify the race of the population studied [10,29–30,59]. Among the six studies where polymorphisms beyond the

common *TPMT*2* and *TPMT*3* were examined [16,24,30,62,64,67], five were conducted in European populations; the sixth did not specify ethnicity or race [30].

Table 6 presents test performance characteristics for genotyping when deficient was defined as absent to intermediate TPMT activity (suggesting the presence of a homozygous, heterozygous or compound heterozygous mutation). This table was easier to populate compared with the previous table (Table 5) as the ability to detect mutations increased with the inclusion of heterozygous status, which is more commonly found. Twenty-five studies provided sufficient data to calculate both sensitivity and specificity. Calculated sensitivity ranged from 13.4–100.0% and specificity ranged from 90.9–100.0%. Of the 25 studies, only one had perfect sensitivity and specificity of 100.0%, the only study conducted in a Tunisian population [13].

There was no clear trend indicating whether additional SNPs increased the sensitivity. Six of nine (67%) studies with >75% sensitivity tested only *TPMT*2* and *TPMT*3*, whereas 12/16 (75%) of studies with <75% sensitivity tested only *TPMT*2* and *TPMT*3*. One study with >75% sensitivity had a sample size of 35 (number of persons with low + intermediate enzyme activity = 18, persons with high enzyme activity = 17) [67], while the remaining eight had sample sizes that ranged from 88–1214 (number of persons with low + intermediate enzyme activity ranged from 5–954, persons with high enzyme activity ranged from 17–6241) [13,23,27,38,58,62,64,66].

Only four studies in the genotype–genotype group reported test performance characteristics. Roman *et al.* [15] reported sensitivity, specificity, PPV and NPV. Schaeffeler *et al.* [62], Lu *et al.* [76] and Anglicheau *et al.* [6] reported concordance (Table 6).

Discussion

This review revealed a diverse and large body of literature assessing both phenotype and genotype technologies for TPMT testing across several diseases. Published studies compare phenotype and genotype technologies, as well as different laboratory methodologies within each technology (genotype–genotype testing and phenotype–phenotype testing) with increasing focus on genotype methods in recent years. It is clear that there are limitations to both genotype testing and phenotype testing, with neither accepted as a gold standard for identifying TPMT deficiency.

The quality appraisal revealed that the quality of the studies was varied. Inadequate reporting of information regarding index tests, reference tests, recruitment methods and study populations was the primary reason for low quality. There was a paucity of reporting by authors of test performance results, indicating a need for guidance on reporting for diagnostic technologies. This review found 30 high-quality studies comparing phenotype and genotype technologies and an additional six high-quality genotype-genotype studies.

When performance characteristics were reported, it was rare for 95% CIs to be included. The low prevalence of deficient TPMT activity (homozygous mutations) in the population made it challenging for many study authors to acquire a sufficient sample size to calculate test accuracy. A number of studies conducted a genotype test only for those subjects who had

demonstrated low TPMT enzyme activity on a phenotype. While this choice may reflect clinical practice or may be related to the comparatively high cost of genotype testing, a serial testing design inflates genotype test sensitivity and should be reported separately from estimates from general or heterogeneous patient populations. The highest quality studies included genotyping for *TPMT*1* (*1S* & *1A*), *TPMT*2*, *TPMT*3* (*3A*, *3B*, *3C* & *3D*), *TPMT*4*, *TPMT*5*, *TPMT*6*, *TPMT*7* and *TPMT*8*. As the number of polymorphisms tested increased, the sensitivity of the test was expected to increase, and with the exception of one study, this trend was weakly shown.

With regard to measurement of enzymatic activity for the phenotype test, limited consistency in cutpoints between low, intermediate and high activity categories was observed. Authors frequently used a receiver operating characteristic analysis to determine the cutpoint for their study population. In addition, measurement units for enzyme activity were variable, making the comparability of cutpoints difficult. Using a cutpoint that defined deficient as the absence of enzyme activity/presence of a homozygous mutation, 10 of 15 studies for which both sensitivity and specificity could be calculated demonstrated perfect (100%) sensitivity and specificity. The inference of perfect values may be misleading, however. Due to the low prevalence of homozygous mutations (0.3%), it is possible that the sample sizes of the studies were too small for a stable rate of detection of this rare mutation. Using a cut-point that defined deficient as low to intermediate enzyme activity/presence of heterozygous/compound heterozygous or homozygous mutation, only 1 of 25 studies for which both sensitivity and specificity could be calculated displayed perfect (100%) sensitivity and specificity. Raising the cutpoint for the definition of deficient activity to include the intermediate activity (heterozygous/compound heterozygous mutation) enabled the detection of more positive cases, resulting in more stable determinations of sensitivity and specificity from the data provided. However, such a cutpoint is not useful for isolating patients at highest risk of a severe ADE.

The clinical utility of TPMT testing lies in its ability to distinguish patients with homozygous mutations (deficient TPMT activity) from other patients to know in whom thiopurines should be avoided, as well as to identify who requires a reduced dose (heterozygous patients with intermediate TPMT activity). Only 15 studies included sufficient data to estimate sensitivity and specificity of genotyping for this purpose. It was evident that distinguishing between these different patient groups was not the priority in many studies.

The variation in sensitivity and specificity observed in the present review may also be related to the disease context. In more severe and life-threatening diseases such as ALL, a higher risk of drug-related adverse events such as myelosuppression may be tolerated to maximize the chemotherapeutic dose of the thiopurine. This would result in a preference for a higher threshold resulting in more false negatives (lower sensitivity) and fewer false positives (higher specificity). In contrast, a different set of thresholds, and consequently values for sensitivity and specificity, may be preferred for chronic diseases such as IBD and dermatological conditions.

Consideration of preanalytical components is important for the success of any diagnostic test, as the risk of error in the laboratory is highest during this phase [78]. Both phenotype and genotype tests contain laboratory and operator steps, which could introduce error. Genotyping offers a solution to the variability of TPMT phenotype activity measurement and potential misclassification due to confounding effects such as recent blood transfusions and certain medications [79]. Graham [80] suggests that selectively genotyping patients whose phenotype tests indicate low enzyme activity and who may be at highest risk of an ADE may be the best approach to avoid the confounding issues of phenotype testing. Again, the issue of choice of polymorphisms in the genotype must be considered.

A previous review identified 17 studies of the performance characteristics of phenotype or genotype testing [2]; however, not all of those studies were found to be of high quality when appraised using the QUADAS-2 tool in this review [5,6,8,9,49,51,53,67,81]. In the previous review, the genotype test performance characteristics, expressed in terms of sensitivity and specificity, ranged from 55 to 100% and from 94 to 100%, respectively. The sensitivity and specificity of the phenotype test ranged from 92 to 100% and 86 to 98%, respectively.

Poor reporting practices were a significant contributor to the exclusion of studies from the present review and were also found in the previous review that used a modified Critical Appraisal Skills Program tool [82]. In another review of papers studying the relationship between genotype and drug-related myelosuppression, a quality appraisal of 67 studies that used a pharmacogenetic assessment tool [83] did not detect any low-quality studies [84]. In a review comparing phenotype and genotype diagnostic accuracy that used the QUADAS-2, 37% of the studies were deemed low quality [78]. The range of quality appraisal tools, reporting practices and judgments regarding high and low quality underscore the importance of building a consensus on reporting of quality in evaluations of diagnostic tests. This issue will become more salient with the increasing use of pharmacogenomics in healthcare.

A meta-analysis of 16 studies of TPMT test performance was performed by the US Agency for Health-care Research and Quality (AHRQ) in 2010. In that review, pooled sensitivity for detecting homozygosity and heterozygosity was 70.7% (95% CI: 37.9–90.5) and pooled specificity was 99.9% (95% CI: 97.4–99.6). Sensitivity and specificity estimates from individual studies were statistically transformed to make them more normally distributed before independent mean estimates were calculated [78]. However, that review did not address the correlation between sensitivity and specificity in performing the meta-analysis. A further limitation of the AHRQ analysis was that it only considered TPMT testing for IBD patients and omitted adults or children with ALL. In addition, the AHRQ analysis assumed that all cutpoints for labeling results as positive or negative were the same across studies. The variation in cutpoints observed in the present review suggests that an assumption of cutpoint equivalence may have introduced bias into the AHRQ pooled estimates [78].

With regard to strengths and limitations, the search strategy was comprehensive by thoroughly searching all relevant citation databases, grey literature sources and by including foreign language articles. It is possible, however, that some relevant articles were missed. As the bulk of screening, reviewing and appraising was performed by a single reviewer with

consultation from two others, the present review might have been enhanced had two independent reviewers been available for all filtering, review and appraisal tasks.

Choosing the QUADAS-2 allowed the assessment to be tailored to the research objective and this tool is recommended by the Cochrane Diagnostic Test Accuracy Working Group, the world leader in systematic review and quality appraisal methods [85]. One disadvantage of the QUADAS is that it is a summary tool that was not designed to distinguish between low and high quality, requiring reviewers to develop judgment-based criteria regarding what constitutes low quality. The addition of a genomics domain to the QUADAS-2 significantly improved the ability to use this tool to assess bias pertaining to genomic testing and can be useful for future quality appraisals of assessments of genomic diagnostic tests. The calculations of sensitivity and specificity were hampered by the absence of cell count data in many studies and low cell counts may have contributed to unstable estimates.

In the absence of a gold standard, the present review set the reference test as the phenotype test. This is the older test and test results are subject to confounding from blood transfusions as well as drug interactions [79] with known imperfect sensitivity and specificity [82]. The range of polymorphisms included in the genotype test would also affect its sensitivity and specificity, thus both approaches have limitations.

Conclusion

The types of pharmacogenomic tests available for selection of drug treatment and dose to avert serious ADEs have been growing. This systematic review comparing phenotype testing and genotype testing for TPMT status demonstrates a broad but diverse base of evidence for these tests. The quality of the studies for assessing diagnostic test accuracy was mixed. The literature displayed a profound lack of patients with low TPMT activity or homozygous TPMT mutations, making estimates of sensitivity of the tests uncertain. Clinical and laboratory decision-makers require high-quality evidence of clinical validity and clinical utility of TPMT genotyping technologies to ensure appropriate and consistent use in patient populations who would benefit from this testing. In selecting a testing approach, clinical decision-makers must consider the patient population, the ethnicity of the patient and the variants that should be included in the test if a genotype test is preferred. Laboratory directors must also consider the availability and cost of tests that permit testing for a wide range of variants, the ability to automate testing, training required and other operator characteristics as well as the technology’s shelf-life.

Future perpsective

There is a growing use of personalized medicine applications such as pharmacogenomics in clinical diagnostics and selection of drug treatment and dose. The automation of laboratory processes including DNA extraction and PCR has made genotyping more rapid and less expensive. Although current tests may become less costly in the future, there may also be variants that have not yet been identified with current methods. Next generation sequencing including whole exome and whole genome sequencing is expected to provide greater yields of variants related to disease as well as drug metabolizing activity [86], but use of these

technologies may not be cost–effective for all applications and requires further evaluation. Consideration also needs to be given to the applicability of pharmacogenetic discoveries to ethnically diverse populations and to vulnerable populations such as children and the elderly.

There is a need for consistent guidelines for reporting diagnostic test accuracy findings. This will be increasingly important as new technologies evolve. Likewise, it is important that future studies adequately sample subjects with homozygous mutations and deficient TPMT activity to better estimate sensitivity of diagnostic tests.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

Papers of special note have been highlighted as:

• of interest;

•• of considerable intere

1. Tantisira, K., Weiss, ST. Overview of pharmacogenomics. In: Basow, DS., editor. UpToDate. MA, USA: 2013.
2. Donnan JR, Ungar WJ, Mathews M, Rahman P. Systematic review of thiopurine methyltransferase genotype and enzymatic testing strategies. Ther Drug Monit. 2011; 33(2):192–199. [PubMed: 21240057]
3. Macdermott, RP. 6-mercaptopurine (6-MP) metabolite monitoring and TPMT testing in the treatment of inflammatory bowel disease with 6-MP or azathioprine. In: Basow, DS., editor. UpToDate. MA, USA: 2013.
4. Baker GR, Norton PG, Flintoft V, et al. The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. CMAJ. 2004; 170(11):1678–1686. [PubMed: 15159366]
5. 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):165–174. [PubMed: 11503011]
6. 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):119–127.
7. 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):631–636. [PubMed: 14508387]
8. Kham SKY, 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):373–379. [PubMed: 18193212]
9. Larovere LE, De Kremer RD, Lambooy LHJ, De Abreu RA. Genetic polymorphism of thiopurine S-methyltransferase in Argentina. Ann Clin Biochem. 2003; 40(4):388–393. [PubMed: 12880540]
10. 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):1069–1077. [PubMed: 17439508]

11. 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. [PubMed: 14985890]
12. 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):1485–1489. [PubMed: 15784872]
13. Ben Salah L, Belkhiria El Haj Amor M, Chbili C, et al. Analysis of thiopurine S-methyltransferase phenotype–genotype in a Tunisian population with Crohn's disease. Eur J Drug Metab Pharmacokinet. 2013; 38(4):241–244. [PubMed: 23553048]
14. Liang JJ, Geske JR, Boilson BA, et al. TPMT genetic variants are associated with increased rejection with azathioprine use in heart transplantation. Pharmacogenet Genomics. 2013; 23(12): 658–665. [PubMed: 24121523]
15. Roman M, Cabaleiro T, Ochoa D, et al. Validation of a genotyping method for analysis of TPMT polymorphisms. Clin Ther. 2012; 34(4):77980878–884.
16. Wennerstrand P, Martensson LG, Soderhall S, Zimdahl A, Appell ML. Methotrexate binds to recombinant thiopurine S-methyltransferase and inhibits enzyme activity after high-dose infusions in childhood leukaemia. Eur J Clin Pharmacol. 2013; 69(9):1641–1649. [PubMed: 23660772]
17. Kongkaew C, Noyce PR, Ashcroft DM. Hospital admissions associated with adverse drug reactions: a systematic review of prospective observational studies. Ann Pharmacother. 2008; 42(7):1017–1025. [PubMed: 18594048]
18. Burnett HF, Tanoshima R, Chandranipapongse W, Madadi P, Ito S, Ungar WJ. Testing for thiopurine methyltransferase status for safe and effective thiopurine administration: a systematic review of clinical guidance documents. Pharmacogenomics J. 2014; 14(6):493–502. [PubMed: 25156214]
19. Roy, LM., Ungar, WJ., Zur, RM. Thiopurine S-methyltransferase testing for averting drug toxicity in patients receiving thiopurines: a systematic review and quality appraisal. TASK Report No. 2015–2002, The Hospital for Sick Children. 2015. www.sickkids.ca
20. Appell ML, Berg J, Duley J, et al. Nomenclature for alleles of the thiopurine methyltransferase gene. Pharmacogenet Genomics. 2013; 23(4):242–248. [PubMed: 23407052]
21•. Weinshilboum RM, Sladek SL. Mercaptopurine pharmacogenetics: monogenic inheritance of erythrocyte thiopurine methyltransferase activity. Am J Hum Genet. 1980; 32(5):651–662. Original investigation describing the mutation responsible for thiopurine S-methyltransferase (TPMT) deficiency. Proposes a classification system to categorize phenotype activity as deficient (<5 U/ml packed red blood cells [RBC]), intermediate (5–10 U/ml packed RBCs) and normal (>10 U/ml pRBC). [PubMed: 7191632]
22•. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011; 155(8):529–536. Description of the Quality Assessment tool for Diagnostic Accuracy Studies-2, the preferred quality appraisal checklist for diagnostic test literature. [PubMed: 22007046]
23. 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):183–188. [PubMed: 11503013]
24. Spire-Vayron De La Moureyre C, Debuysere H, Mastain B, et al. Genotypic and phenotypic analysis of the polymorphic thiopurine S-methyltransferase gene (*TPMT*) in a European population. Br J Pharmacol. 1998; 125(4):879–887. [PubMed: 9831928]
25. Fakhoury M, Andreu-Gallien J, Mahr A, et al. Should *TPMT* genotype and activity be used to monitor 6-mercaptopurine treatment in children with acute lymphoblastic leukaemia? J Clin Pharm Ther. 2007; 32(6):633–639. [PubMed: 18021342]
26. Fangbin Z, Xiang G, Minhu C, et al. Should thiopurine methyltransferase genotypes and phenotypes be measured before thiopurine therapy in patients with inflammatory bowel disease? Ther Drug Monit. 2012; 34(6):695–701. [PubMed: 23149442]
27. Ford L, Graham V, Berg J. Whole-blood thiopurine S-methyltransferase activity with genotype concordance: a new, simplified phenotyping assay. Ann Clin Biochem. 2006; 43(Pt 5):354–360. [PubMed: 17036413]

28. Ford L, Kampanis P, Berg J. Thiopurine S-methyltransferase genotype–phenotype concordance: used as a quality assurance tool to help control the phenotype assay. Ann Clin Biochem. 2009; 46(Pt 2):152–154. [PubMed: 19164342]

29. Gazouli M, Pachoula I, Panayotou I, et al. Thiopurine methyltransferase genotype and thiopurine S-methyltransferase activity in Greek children with inflammatory bowel disease. Ann Gastroenterol. 2012; 25(3):249–253. [PubMed: 24714152]

30. Hindorf U, Appell ML. Genotyping should be considered the primary choice for pre-treatment evaluation of thiopurine methyltransferase function. J Crohns Colitis. 2012; 6(6):655–659. [PubMed: 22398041]

31. Larussa T, Suraci E, Lentini M, et al. High prevalence of polymorphism and low activity of thiopurine methyltransferase in patients with inflammatory bowel disease. Eur J Intern Med. 2012; 23(3):273–277. [PubMed: 22385887]

32•. Lennard L, Cartwright CS, Wade R, Richards SM, Vora A. Thiopurine methyltransferase genotype-phenotype discordance and thiopurine active metabolite formation in childhood acute lymphoblastic leukaemia. Br J Clin Pharmacol. 2013; 76(1):125–136. High-quality publication comparing phenotype to genotype diagnostic tests in 1117 children with acute lymphoblastic leukemia. [PubMed: 23252716]

33. Milek M, Murn J, Jaksic Z, Lukac Bajalo J, Jazbec J, Mlinaric Rascan I. Thiopurine S-methyltransferase pharmacogenetics: genotype to phenotype correlation in the Slovenian population. Pharmacology. 2006; 77(3):105–114. [PubMed: 16691038]

34. 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.

35. Serpe L, Calvo PL, Muntoni E, et al. Thiopurine S-methyltransferase pharmacogenetics in a large-scale healthy Italian–Caucasian population: differences in enzyme activity. Pharmacogenomics. 2009; 10(11):1753–1765. [PubMed: 19891552]

36. Xin HW, Xiong H, Wu XC, Li Q, Xiong L, Yu AR. Relationships between thiopurine S-methyltransferase polymorphism and azathioprine-related adverse drug reactions in Chinese renal transplant recipients. Eur J Clin Pharmacol. 2009; 65(3):249–255. [PubMed: 19048245]

37. Zhang LR, Song DK, Zhang W, Zhao J, Jia LJ, Xing DL. Efficient screening method of the thiopurine methyltransferase polymorphisms for patients considering taking thiopurine drugs in a Chinese Han population in Henan Province (central China). Clin Chim Acta. 2007; 376(1–2):45–51. [PubMed: 16952345]

38. Jorquera A, Solari S, Vollrath V, et al. Phenotype and genotype of thiopurine methyltransferase in Chilean individuals (Spanish). Rev Med Chil. 2012; 140(7):889–895. [PubMed: 23282701]

39. Ma XL, Wu MY, Hu YM, Zu P, Li ZG. Relationships between thiopurine methyltransferase gene polymorphisms and its enzymatic activity (Chinese). Zhonghua Zhong Liu Za Zhi. 2006; 28(6): 456–459. [PubMed: 17152495]

40. Ansari A, Hassan C, Duley J, et al. Thiopurine methyltransferase activity and the use of azathioprine in inflammatory bowel disease. Aliment Pharmacol Ther. 2002; 6(10):1743–1750.

41. Arenas M, Simpson G, Lewis CM, et al. Genetic variation in the *MTHFR* gene influences thiopurine methyltransferase activity. Clin Chem. 2005; 51(12):2371–2374. [PubMed: 16306100]

42. Barlow NL, Graham V, Berg JD. Expressing thiopurine S-methyltransferase activity as units per litre of whole-blood overcomes misleading high results in patients with anaemia. Ann Clin Biochem. 2010; 47(Pt 5):408–414. [PubMed: 20595409]

43. Ebbesen MS, Nersting J, Jacobsen JH, et al. Incorporation of 6-thioguanine nucleotides into DNA during maintenance therapy of childhood acute lymphoblastic leukemia-the influence of thiopurine methyltransferase genotypes. J Clin Pharmacol. 2013; 53(6):670–674. [PubMed: 23585236]

44. Evans WE, Hon YY, Bomgaars L, et al. Preponderance of thiopurine S-methyltransferase deficiency and heterozygosity among patients intolerant to mercaptopurine or azathioprine. J Clin Oncol. 2001; 19(8):2293–2301. [PubMed: 11304783]

45. Ferucci ED, Hurlburt KJ, Mayo MJ, et al. Azathioprine metabolite measurements are not useful in following treatment of autoimmune hepatitis in Alaska Native and other non-Caucasian people. Can J Gastroenterol. 2011; 25(1):21–27. [PubMed: 21258664]
46. Gu LJ, Ye QD, Liang AB, et al. Significance of TPMT activity and TGNs level detection for individualizing 6-mercaptopurine chemotherapy (Chinese). Chung Hua Hsueh Yeh Hsueh Tsa Chi. 2003; 24(1):18–21.
47. 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):288–295. Erratum appears in *Clin. Chem.* 50(4), 788 (2004). [PubMed: 14656901]
48. Heckmann JM, Lambson EMT, 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. [PubMed: 15792824]
49. 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):371–376. [PubMed: 9931346]
50. Kasirer Y, Mevorach R, Renbaum P, et al. Thiopurine S-methyltransferase (*TPMT*) activity is better determined by biochemical assay versus genotyping in the Jewish population. Dig Dis Sci. 2014; 59(6):1207–1212. [PubMed: 24390675]
51. Reis M, Santoro A, Suarez-Kurtz G. Thiopurine methyltransferase phenotypes and genotypes in Brazilians. Pharmacogenetics. 2003; 13(6):371–373. [PubMed: 12777968]
52. Relling MV, Hancock ML, Boyett JM, Pui C-H, Evans WE. Prognostic importance of 6-mercaptopurine dose intensity in acute lymphoblastic leukemia. Blood. 1999; 93(9):2817–2823. [PubMed: 10216075]
53. 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):51–54. [PubMed: 11372592]
54. 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):557–564. [PubMed: 18987654]
55. Sies C, Florkowski C, George P, et al. Measurement of thiopurine methyl transferase activity guides dose-initiation and prevents toxicity from azathioprine. N Z Med J. 2005; 118(1210):U1234.
56. Tamm R, Oselin K, Kallassalu K, et al. Thiopurine S-methyltransferase (*TPMT*) pharmacogenetics: three new mutations and haplotype analysis in the Estonian population. Clin Chem Lab Med. 2008; 46(7):974–979. [PubMed: 18605963]
57. El-Azhary RA, Farmer SA, Drage LA, et al. Thioguanine nucleotides and thiopurine methyltransferase in immunobullous diseases: optimal levels as adjunctive tools for azathioprine monitoring. Arch Dermatol. 2009; 145(6):644–652. [PubMed: 19528417]
58. Ganiere-Monteil C, Medard Y, Lejus C, et al. Phenotype and genotype for thiopurine methyltransferase activity in the French Caucasian population: impact of age. Eur J Clin Pharmacol. 2004; 60(2):89–96. [PubMed: 15022030]
59. Langley PG, Underhill J, Tredger JM, Norris S, Mcfarlane IG. Thiopurine methyltransferase phenotype and genotype in relation to azathioprine therapy in autoimmune hepatitis. J Hepatol. 2002; 37(4):441–447. [PubMed: 12217596]
60. Marinaki AM, Arenas M, Khan ZH, et al. Genetic determinants of the thiopurine methyltransferase intermediate activity phenotype in British Asians and Caucasians. Pharmacogenetics. 2003; 13(2): 97–105. [PubMed: 12563179]
61. 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): 2001–2008. [PubMed: 10580024]
62. 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):407–417. [PubMed: 15226673]

63. Schmiegelow K, Al-Modhwahi I, Andersen MK, et al. Methotrexate/6-mercaptopurine maintenance therapy influences the risk of a second malignant neoplasm after childhood acute lymphoblastic leukemia: results from the NOPHO ALL-92 study. Blood. 2009; 113(24):6077–6084. [PubMed: 19224761]

64. 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):429–436. [PubMed: 12172211]

65. von Ahsen N, Armstrong VW, Behrens C, et al. Association of inosine triphosphatase 94C>A and thiopurine S-methyltransferase deficiency with adverse events and study drop-outs under azathioprine therapy in a prospective Crohn disease study. Clin Chem. 2005; 51(12):2282–2288. Erratum appears in *Clin. Chem.* 52(8), 1628 (2006) Aug. Note: Schutz, E (added). [PubMed: 16214825]

66. 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):608–614. [PubMed: 9103127]

67. 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):177–185. [PubMed: 9711875]

68. Lennard L, Singleton HJ. High-performance liquid chromatographic assay of human red blood cell thiopurine methyltransferase activity. J Chromatogr B Biomed Appl. 1994; 661(1):25–33. [PubMed: 7866549]

69. Lu Y, Kow-Yin Kham S, Foo TC, Hany A, Quah TC, Eng-Juh Yeoh A. Genotyping of eight polymorphic genes encoding drug-metabolizing enzymes and transporters using a customized oligonucleotide array. Anal Biochem. 2007; 360(1):105–113. [PubMed: 17113562]

70. Ma X-L, Zu P, Hu Y-M, Wu M-Y, Li Z-G, Bu D-F. Exploration of methodology for assay of single nucleotide polymorphism in thiopurine methyltransferase gene (Chinese). Zhongguo Shi Yan Xue Ye Xue Za Zhi. 2003; 11(5):458–463. [PubMed: 14575536]

71. Chowdhury J, Kagiala GV, Pushpakom S, et al. Microfluidic platform for single nucleotide polymorphism genotyping of the thiopurine S-methyltransferase gene to evaluate risk for adverse drug events. J Mol Diagn. 2007; 9(4):521–529. [PubMed: 17690215]

72. Kim JH, Cheon JH, Hong SS, et al. Influences of thiopurine methyltransferase genotype and activity on thiopurine-induced leukopenia in Korean patients with inflammatory bowel disease: a retrospective cohort study. J Clin Gastroenterol. 2010; 44(10):e242–e248. [PubMed: 20308917]

73. Osaki R, Imaeda H, Ban H, et al. Accuracy of genotyping using the TaqMan PCR assay for single nucleotide polymorphisms responsible for thiopurine sensitivity in Japanese patients with inflammatory bowel disease. Exp Ther Med. 2011; 2(5):783–786. [PubMed: 22977575]

74••. Schaeffeler E, Zanger UM, Eichelbaum M, Asante-Poku S, Shin JG, Schwab M. Highly multiplexed genotyping of thiopurine S-methyltransferase variants using MALD-TOF mass spectrometry: reliable genotyping in different ethnic groups. Clin Chem. 2008; 54(10):1637–1647. High-quality publication comparing detection of ten TPMT variants by alternative genotype tests in 586 health volunteers. [PubMed: 18687736]

75. Kim S, Lee HW, Lee W, Chun S, Min W-K. Validation of new allele-specific real-time PCR system for thiopurine methyltransferase genotyping in Korean population. Biomed Res Int. 2013; 2013:1–4. Article 305704.

76. Lu HF, Shih MC, Hsueh SC, Chen CM, Chang JY, Chang JG. Molecular anlaysis of the thiopurine S-methyltransferase alleles in Bolivians and Tibetans. J Clin Pharm Ther. 2005; 30(5):491–496. [PubMed: 16164497]

77. 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):593–599. [PubMed: 15352906]

78•. Booth RA, Ansari MT, Tricco AC, et al. Assessment of thiopurine methyltransferase activity in patients prescribed azathioprine or other thiopurine-based drugs. Evid Rep Technol Assess (Full Rep). 2010; 2010(196):1–282. Comprehensive meta-analysis of TPMT performance characteristics of TPMT testing in adults.

79. 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. [PubMed: 15298741]

80. Graham, V. Master of Philosophy dissertation. The University of Birmingham College of Medical and Dental Sciences; Birmingham, US: 2009. Thiopurine methyltransferase phenotyping and genotyping in clinical practice; p. 1-167.

81. Indjova D, Shipkova M, Atanasova S, et al. Determination of thiopurine methyltransferase phenotype in isolated human erythrocytes using a new simple nonradioactive HPLC method. Ther Drug Monit. 2003; 25(5):637–644. [PubMed: 14508388]

82. Donnan, JR., Ungar, WJ., Mathews, M., Hancock-Howard, RL. Health technology assessment of thiopurine methyltransferase testing for guiding 6-mercaptopurine doses in pediatric patients with acute lymphoblastic leukemia. TASK Report No. 2010–2002, The Hospital for Sick Children. 2010. www.sickkids.ca

83. Jorgensen ALW, Paula R. Methodological quality of pharmacogenetic studies: issus of concern. Stat Med. 2008; 27:6547–6569. [PubMed: 18837075]

84. Higgs JE, Payne K, Roberts C, Newman WG. Are patients with intermediate TPMT activity at increased risk of myelosuppression when taking thiopurine medications? Pharmacogenomics. 2010; 11(2):177–188. [PubMed: 20136357]

85•. Reitsma, J., Rutjes, AWS., Whiting, P., Vlassov, VV., Leeflang, MMG., Deeks, JJ. Chapter 9: Assessing methodological quality. In: Deeks, J.Bossuyt, PM., Gatsonis, C., editors. Cochrane Handbook for Systematic Reviews of Diagnostic Accuracy (Version 1.0.0). Vol. 28. 2009. Recommended guidelines on conducting systematic reviews of the diagnostic test literature

86. Wright, C., Burton, H., Hall, A., et al. Next steps in the sequence: the implications of whole genome sequencing for health in the UK. PHG Foundation; Cambridge, UK: 2011.

Executive summary

Background

- The absence or a deficiency of thiopurine S-methyltransferase (TPMT) can significantly increase the risk of adverse drug events in patients receiving thiopurines.
- There has long been phenotype blood testing to measure TPMT enzyme activity, and more recently a genotype test is used to identify individuals with genetic variants to assess TPMT status. Guidelines disagree on which test to recommend and uncertainty persists.
- The objectives were to systematically review the literature on the performance characteristics of thiopurine testing for TPMT deficiency, to appraise the quality of the literature and to identify the characteristics of high-quality studies.

Literature retrieval & quality appraisal

- The search identified 4071 publications for review. Full text review was performed on 373 papers and 66 met eligibility criteria and underwent quality appraisal with the Quality Assessment tool for Diagnostic Accuracy Studies.
- In total, 30/55 phenotype–genotype and 6/11 phenotype–phenotype comparisons were deemed of high quality.
- Low-quality studies demonstrated high levels of bias and concerns for applicability,
- High-quality studies were published between 1997 and 2013 and examined a range of genotype and phenotype test methods.

Test performance

- Based on data from 15 studies, the calculated sensitivity for genotyping to identify a homozygous mutation ranged from 0.0–100.0%.
- Based on data from 26 studies, the calculated specificity for genotyping to identify a homozygous mutation ranged from 97.8–100.0%.
- Based on data from 25 studies, the calculated sensitivity for genotyping to detect a homozygous or heterozygous mutation ranged from 13.4–100.0% and specificity ranged from 90.9–100.0%.

Genotyping

- Genotyping studies employed similar DNA amplification methods, with 80% (24/30) using a method of PCR, 57% (17/30) using PCR with RFLP, 26% (8/30) Allele-specific-PCR and 6% (2/30) PCR-single strand conformational polymorphism.

- Methods such as denaturing HPLC, multiplex amplification refractory mutation, pyrosequencing and TaqMan SNP genotyping were reported and direct sequencing was used in three studies.

Conclusion

- There are limitations to both genotype and phenotype testing, and neither test can be referred to as the ‘gold standard’ for identifying TPMT deficiency.
- Lack of reporting of diagnostic test accuracy indicates a need for guidance on reporting of test performance characteristics.
- The number of polymorphisms included in genotype tests ranged from two to nine, with most studies including *TPMT*2* and *TPMT*3*, the most common genetic variants in persons with deficient TPMT activity.
- The variation in sensitivity and specificity observed in the present review may be related to the disease context and low prevalence of a homozygous TPMT mutation.
- The tolerance for the risk of serious adverse drug events, and consequently values for sensitivity and specificity, may be different for chronic disease such as inflammatory bowel disease compared with life-threatening diseases such as acute lymphoblastic leukemia.
- Clinical decision-makers require high-quality evidence of clinical validity and clinical utility of TPMT genotyping technologies to ensure appropriate use in patient populations who would benefit from this testing.

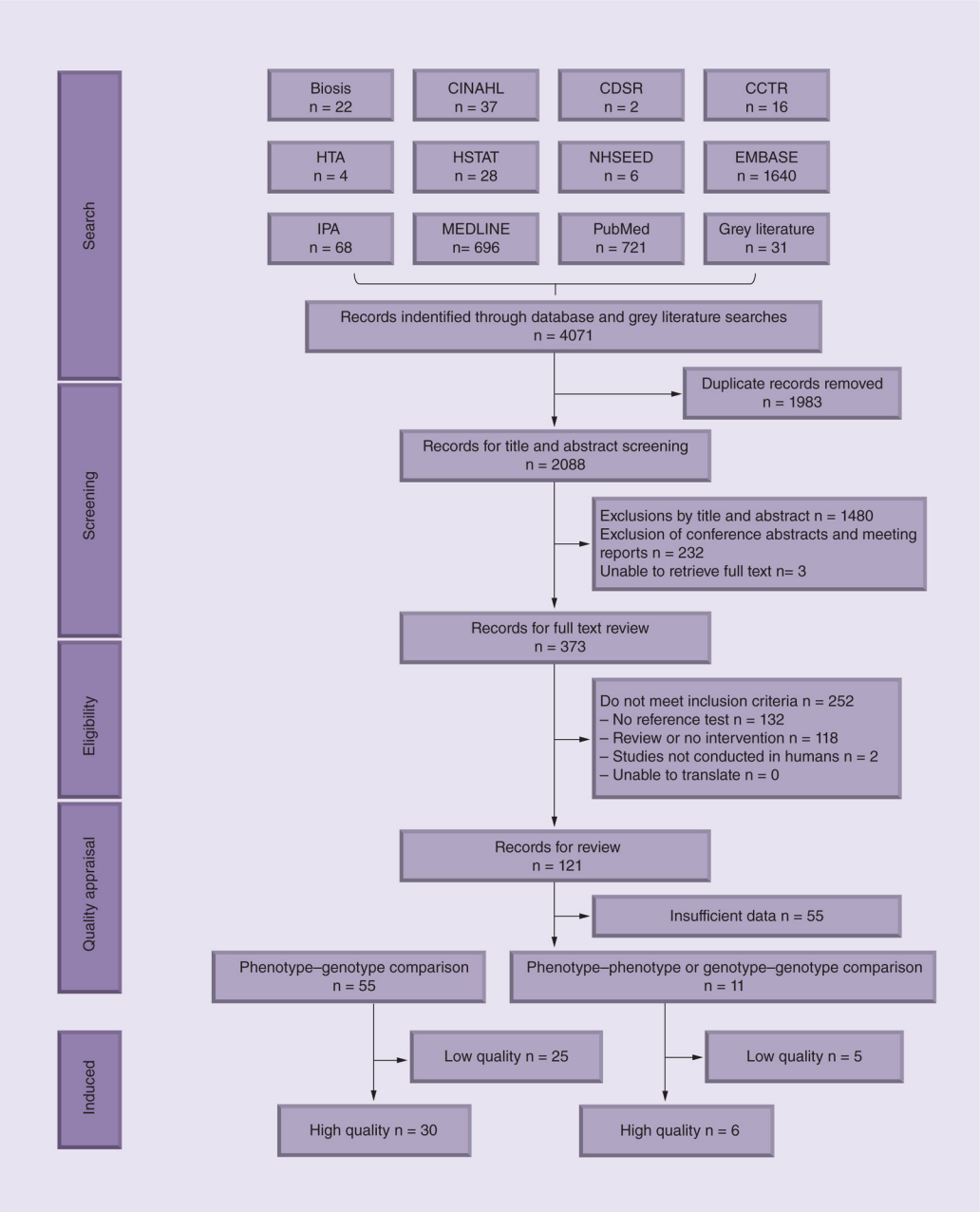

Figure 1.
PRISMA flowchart.

Pharmacogenomics. Author manuscript; available in PMC 2016 July 04.

Table 1

Quality Assessment tool for Diagnostic Accuracy Studies-2 results for high-quality studies.

Study	Year	Domain 1 – patient selection		Domain 2 – index test		Domain 3 – reference test		Domain 4 – flow and timing	Domain 5 – genomics	Ref.
		Risk of bias	Concern for applicability	Risk of bias	Concern for applicability	Risk of bias	Concern for applicability	Risk of bias	Risk of bias	
Phenotype–genotype studies (n = 30)										
Ben Salah	2013	Low	Low	Low	Low	Unclear	Low	Low	Low	[13]
Fakhoury	2007	Low	Low	Low	Low	Low	Low	Low	Low	[25]
Fangbin	2012	Low	Low	Low	Low	High	Low	Low	Low	[26]
Ford	2006	Low	Low	Low	Low	High	Low	Low	Low	[27]
Ford	2009	Low	Low	Low	Low	Low	Low	Unclear	Low	[28]
Ganiere-Monteil	2004	Low	Low	Low	Low	High	Low	Low	Low	[58]
Gazouli	2012	Low	Low	Low	Low	Low	Low	Low	Low	[29]
Hindorf	2012	Low	Low	Low	Low	Unclear	Low	Low	Low	[30]
Jorquera	2012	Low	Low	Low	Low	Low	Low	Low	Low	[38]
Langley	2002	Low	Low	High	Low	Low	Low	Low	Low	[59]
Larussa	2012	Low	Low	Low	Low	Low	Low	Low	Low	[31]
Lennard	2013	Low	Low	Low	Low	Low	Low	Low	Low	[32]
Liang	2013	Low	Low	Low	Low	Low	Low	Low	Low	[14]
Loennechen	2001	Low	Low	Low	Low	Low	Low	Low	Low	[23]
Ma	2006	Low	Low	Low	Low	Unclear	Low	Low	Low	[39]
Marinaki	2003	Low	Low	Low	Low	Unclear	Unclear	Low	Low	[60]
Milek	2006	Low	Low	Low	Low	Low	Low	Low	Low	[33]

Study	Year	Domain 1 – patient selection		Domain 2 – index test		Domain 3 – reference test		Domain 4 – flow and timing	Domain 5 – genomics	Ref.
		Risk of bias	Concern for applicability	Risk of bias	Concern for applicability	Risk of bias	Concern for applicability	Risk of bias	Risk of bias	
Oselin	2006	Low	Low	Low	Low	High	Low	Low	Low	[34]
Schaeffeler	2004	Low	Low	Low	Low	Low	Low	Low	Low	[62]
Schwab	2002	Low	Low	Low	Low	Low	Low	Low	Low	[64]
Serpe	2009	Low	Low	Low	Low	Unclear	Low	Low	Low	[35]
Spire-Vayron de la Moureyre	1998	Unclear	Low	Low	Low	Low	Low	Low	Low	[67]
Spire-Vayron de la Moureyre	1998	Unclear	Low	Low	Low	Low	Low	Low	Low	[24]
von Ahsen	2005	Low	Low	Low	Unclear	Unclear	Low	Low	Low	[65]
Wennerstrand	2013	Low	Low	Low	Low	Low	Low	Low	Low	[16]
Winter	2007	Low	Low	Low	Low	Low	Low	Low	Low	[10]
Wusk	2004	Low	Low	Low	Low	High	Low	Low	Low	[11]
Xin	2009	Low	Low	Low	Low	Low	Low	Low	Low	[36]
Yates	1997	Low	Low	Low	Low	Low	Low	Low	High	[66]
Zhang	2007	Low	Low	Low	Low	Unclear	Low	Low	Low	[37]
Phenotype–phenotype or genotype–genotype studies (n = 6)										
Chowdhury	2007	Unclear	Low	Low	Low	Low	Low	Low	Low	[71]
Kim	2013	Unclear	Low	Low	Low	Low	Low	Low	Low	[75]
Lu	2005	Low	Low	Low	Low	Unclear	Low	Low	Low	[76]
Ma	2003	Low	Low	Low	Low	Low	Low	Low	Unclear	[70]

Study	Year	Domain 1 – patient selection		Domain 2 – index test		Domain 3 – reference test		Domain 4 – flow and timing	Domain 5 – genomics	Ref.
		Risk of bias	Concern for applicability	Risk of bias	Concern for applicability	Risk of bias	Concern for applicability	Risk of bias	Risk of bias	
Roman	2012	Unclear	Low	Low	Low	Low	Low	Low	Low	[15]
Schaeffeler	2008	Low	Low	Low	Low	Low	Low	Low	Low	[74]

Table 2

Design characteristics of high-quality studies.

Author	Year	Primary objective	Inclusion criteria	Age group	Disease group	Ethnicity	Number included	Ref.
Phenotype–genotype studies (n = 30)								
Ben Salah	2013	Investigate TPMT activity distribution and allele frequency of common alleles	Not specified	Not specified	Other	Tunisian	88	[13]
Fakhoury	2007	Study correlations between TPMT genotype and enzyme activity	Children diagnosed with ALL; enrolled in two consecutive European trials	Pediatric	ALL	European	118	[25]
Fangbin	2012	Role of phenotype and genotype in predicting leukopenia	Patients with steroid- dependent disease, frequent relapses, on remission maintenance and postoperative prophylaxis	Adult	IBD	Chinese Han nationality; lived in Henan Province, Peoples Republic of China	499	[26]
Ford	2006	Compare new method phenotype (whole blood) with old method (RBC lysate) and genotype	Routine samples collected over 4-week period	Not specified	Not specified	Not specified	402	[27]
Ford	2009	Examine phenotype–genotype concordance to investigate effectiveness as QA tool	All consecutive routinely collected samples	Not specified	Not specified	Not specified	Not specified	[28]
Ganiere- Monteil	2004	Investigate the impact of age on TPMT activity by comparing TPMT activity (pheno and geno) in healthy young Caucasians from birth (cord blood) to adolescence with adult Caucasians	Patients with IBD; taking AZA or 6-MP for at least 3 months or experienced adverse events with these drugs; dose between 0.3–2.5 mg/kg	Mix of adult and pediatric	Otherwise healthy	Caucasian	468	[58]
Gazouli	2012	Examine sensitivity and specificity of TPMT genotyping for TPMT enzymatic activity	Patients with diagnosis of IBD; patients using AZA or 6-MP >3 months or adverse event during treatment; dosage range specified	Pediatric	IBD	Not specified	108	[29]
Hindorf	2012	Investigate the correlation between TPMT genotype and phenotype; analyze the results from a clinical and practical perspective	Unselected and consecutive TPMT phenotype and genotype determinations sent to the study site	Not specified	IBD	Not specified	7195	[30]
Jorquera	2012	Study the TPMT activity and genotype in Chilean subjects	Healthy persons; older than 18 years; unrelated	Adult	Otherwise healthy	Spanish, Chilean	200	[38]
Langley	2002	Determine whether the phenotypes or genotypes correlate with clinical outcomes for AZA therapy	Patients attending the autoimmune liver disease outpatients' clinic	Mix of adult and pediatric	Other	Not specified	53	[59]

Author	Year	Primary objective	Inclusion criteria	Age group	Disease group	Ethnicity	Number included	Ref.
Larussa	2012	Investigate TPMT genotype and phenotype status in southern Italian IBD patients	Patients with Crohn's or UC	Adult	IBD	Caucasian, Italian	51	[31]
Lennard	2013	Investigate phenotype–gentoype TPMT concordance in children with ALL	Patients diagnosed with ALL in time frame specified, at treatment centers in the UK and Ireland	Pediatric	ALL	UK and Ireland (English, Irish)	1117	[32]
Liang	2013	Investigate the relationship between TPMT enzymatic activity and genetic variation in TPMT with AZA clinical efficacy, especially in prevention of rejection and safety in HTX recipients	Heart transplant recipients at Mayo Clinic; treated with AZA	Adult	Organ transplant	Not specified	93	[14]
Loennechen	2001	Identify TPMT mutant alleles in a Saami population to develop genotype tests for prediction of TPMT activity	Patients >18 years old	Adult	Patients admitted to a cardiology center	Caucasian, Saami	260	[23]
Ma	2006	Investigate the relationship between the TPMT gene polymorphisms and its enzymatic activity	Healthy blood donors; cord blood; patients with leukemia	Mix of adult and pediatric	ALL	Chinese	630	[39]
Marinaki	2003	Establish frequencies of genetic modifiers of TPMT activity in Asian residents of the United Kingdom	Patients originating from India and Pakistan attending an IBD clinic	Not specified	IBD	Originating from India and Pakistan; Caucasian	85	[60]
Milek	2006	Determine the frequency of clinically significant, low-activity TPMT alleles	Unrelated healthy volunteers	Not specified	Otherwise healthy	Slovenian	95	[33]
Oselin	2006	Develop and validate an HPLC method with UV detection to determine TPMT activity in human erythrocytes using 6-MP as a substrate	Volunteers; Estonian	Adult	Otherwise healthy	Estonian	99	[34]
Schaeffeler	2004	Sensitivity, specificity, PPV, and NPV for TPMT genotyping	No regular drug use with the exception of oral contraceptives and/or vitamins.	Adult	Otherwise healthy	Caucasian, German	1214	[62]
Schwab	2002	Whether AZA-related serious side-effects can be explained by TPMT polymorphism using both pheno and genotyping	Patients with IBD from Department of Gastroenterology at University Hospital Tubingen; on AZA therapy at present or previously	Adult	IBD	Caucasian	93	[64]
Serpe	2009	Elucidate the impact of genotype, age, gender on TPMT phenotype by comparing the activity of the enzyme among infants,	Healthy, unrelated, Italian–Caucasian adults; newborn, Italian–Caucasian babies, children or adolescents	Mix of adult and pediatric	Otherwise healthy	Italian–Caucasian	943	[35]

Author	Year	Primary objective	Inclusion criteria	Age group	Disease group	Ethnicity	Number included	Ref.
		children, adolescents and adults						
Spire- Vayron de la Moureyre	1998	Describe and demonstrate the usefulness of a new SSCP procedure to assay simultaneously for known mutations within TPMT, and to detect new ones	Selected from previously phenotyped individuals; healthy volunteers or patients	Not specified	Otherwise healthy	European	35	[67]
Spire- Vayron de la Moureyre	1998	Overall mutational spectrum of TPMT gene	Unrelated, European, volunteers or patients starting AZA therapy	Not specified	Not specified	European	191	[24]
von Ahsen	2005	Analyze AZA tolerance in relation to ITPA and TPMT mutation status and TPMT activity	>18 years; active Crohn's disease; prednisone treatment >300 mg during the last 4 weeks or a relapse within 6 months after steroid pulse therapy	Adult	IBD	Caucasian	71	[65]
Wennerstrand	2013	Investigate the fluctuation in TPMT enzyme activity from the time of diagnosis until after the end of maintenance treatment	Children starting their treatment per NOPHO ALL- 2000 study protocol	Pediatric	ALL	Scandinavian (Norway, Sweden, Finland)	53	[16]
Winter	2007	To determine if screening for TPMT status predicts side-effects to AZA in patients with IBD	Patients with IBD; no history of treatment with thiopurine drugs	Not specified	IBD	Not specified	130	[10]
Wusk	2004	Phenotype–genotype comparison of TPMT; develop a new screening strategy for patients prior to taking thiopurine drugs	Unrelated healthy volunteers; patients with IBD	Not specified	IBD	German	240	[11]
Xin	2009	Whether AZA-treated serious side effects can be explained by the TPMT polymorphism using both phenotype and genotype tests in adult patients with renal transplantation on AZA therapy	Renal transplant recipients treated with AZA presently or previously	Not specified	Organ transplant	Not specified	150	[36]
Yates	1997	Establish frequencies of the genetic modifiers of TPMT activity in an Asian population resident in the United Kingdom	Volunteer blood donors; children with ALL being treated or referred for evaluation	Mix of adult and pediatric	ALL	Caucasian	48	[66]
Zhang	2007	Phenotype–genotype comparison of the TPMT enzyme and develop a new screening strategy for patients prior to taking thiopurine drugs	Patients with chronic renal failure; no blood transfusion within 1 month prior to study	Adult	Other	Not specified	278	[37]

Author	Year	Primary objective	Inclusion criteria	Age group	Disease group	Ethnicity	Number included	Ref.
Genotype–genotype studies (n = 6)								
Chowdhury	2007	Study compared three methods of genotyping – conventional vs microchip RFLP, and used TaqMan as the 'gold standard'. Also tested new steps in AS-PCR-CE and portable microchip CE, but these were not tested against the others	Patients with IBD; undergoing thiopurine immunosuppression	Not specified	IBD	Not specified	80	[71]
Kim	2013	Develop and validate a new AS-PCR for TPMT genotyping	Not specified	Not specified	Requiring AZA or mercaptopurine	Not specified	244	[75]
Lu	2005	Test feasibility of genotyping using APEX	Patients with β-thalassemia and random selection of patients for TPMT screening (healthy blood donors and children with ALL)	Children and adult	β-thalassemia + patient selected for TPMT screening, also healthy volunteers		200	[76]
Ma	2003	To confirm and study the Chinese TPMT gene polymorphism; to compare and discuss the methodology for SNP tests; to find the best way and most suitable way to test the TPMT polymorphisms	ALL patients who were admitted inpatients by the Hematology Department of Beijing Children Hospital	Mix of adult and pediatric	ALL + healthy blood donors, cord blood	Chinese	630	[70]
Roman	2012	To validate a TPMT genotyping method by comparing it with a conventional PCR approach	Adult white patients from the Hospital Universitario de la Princesa (Spain) for whom genotyping was requested	Adult	For whom TPMT genotyping was requested – GE, derm, rheu, neph, inter med, hemato	White	111	[15]
Schaeffeler	2008	Establishment and application of a novel assay, called iPLEX, for detection of all functional relevant 22 TPMT allelic variants	Healthy unrelated volunteers; Korean, Ghanians	Not specified	Not specified	German (white)	586	[74]

6-MP: 6-mercaptopurine; ALL: Acute lymphoblastic leukemia; APEX: Arrayed primer extension technology; AS-PCR: Allele-specific PCR; AZA: Azathioprine; CE: Capillary electrophoresis; IBD: Inflammatory bowel disease; ITPA: Inosine triphosphatase; NPV: Negative predictive value; PPV: Positive predictive value; QA: Quality assurance; RBC: Red blood cell; RFLP: Restriction fragment length polymorphism; SSCP: Single strand conformational polymorphism; TPMT: Thiopurine S-methyltransferase; UC: Ulcerative colitis; UV: Ultraviolet.

Table 3

Genotype and phenotype laboratory methods for high-quality phenotype–genotype studies (n = 30).

Study	Year	Amplification/genotype method	Population	Polymorphisms tested	Phenotype method	Ref.
Ben Salah	2013	PCR; AS-PCR; RFLP	Other	*TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c*	HPLC	[13]
Fakhoury	2007	PCR; AS-PCR	European	*TPMT*2, TPMT*3a, TPMT*3c*	HPLC	[25]
Fangbin	2012	PCR; RFLP	Chinese	*TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c*	Not specified	[26]
Ford	2006	ARMS; AS-PCR; PCR	Not specified	*TPMT*2, TPMT*3*	HPLC	[27]
Ford	2009	ARMS; AS-PCR; PCR	Not specified	*TPMT*2, TPMT*3a, TPMT*3c*	HPLC	[28]
Ganiere-Monteil	2004	PCR; AS-PCR	Caucasian	*TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c*	HPLC	[58]
Gazouli	2012	PCR; RFLP	Not specified	*TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c*	RC	[29]
Hindorf	2012	Pyrosequencing	Not specified	*TPMT*2, TPMT*3a, TPMT*3c*; those with phenotype under 9.0 were further investigated on exons 3–10	RC	[30]
Jorquera	2012	PCR; RFLP	Other	*TPMT*1, TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c*	HPLC	[38]
Langley	2002	PCR; RFLP	Not specified	*TPMT*3a, TPMT*3b, TPMT*3c*	RC	[59]
Larussa	2012	PCR; RFLP	Caucasian	*TPMT*2, TPMT*3b, TPMT*3c*	Competitive micro-well immunoassay	[31]
Lennard	2013	PCR; RFLP	Other	*TPMT*3a, TPMT*3b, TPMT*3c*	HPLC	[32]
Liang	2013	PCR; TaqMan	Not specified	*TPMT*2, TPMT*3a, TPMT*3c*	Not specified	[14]
Loennechen	2001	PCR; AS-PCR; RFLP	Caucasian	*TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c, TPMT*6*	RC	[23]
Ma	2006	PCR; RFLP	Chinese	*TPMT*2, TPMT*3a, TPMT*3c*	HPLC	[39]
Marinaki	2003	PCR; RFLP	Caucasian	*TPMT*2, TPMT*3a, TPMT*3c*	RC	[60]
Milek	2006	PCR; RFLP, TaqMan	Other	*TPMT*2, TPMT*3b, TPMT*3c*	HPLC	[33]
Oselin	2006	PCR; RFLP	Other	*TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c, TPMT*3D, TPMT*8*	HPLC	[34]
Schaeffeler	2004	PCR; DHPLC	Caucasian	*TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c, TPMT*3D*	RC	[62]
Schwab	2002	DHPLC	Caucasian	*TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c, TPMT*3D*	Not specified	[64]
Serpe	2009	AS-PCR; PCR; RFLP	Other	*TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c*	Not specified	[35]
Spire-Vayron de la Moureyre	1998	PCR-SSCP; Direct sequencing	European	*TPMT*1, TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c, TPMT*1S, TPMT*1A, TPMT*7, TPMT *3d*	RC	[67]
Spire-Vayron de la Moureyre	1998	PCR-SSCP; Direct sequencing	European	*TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c, TPMT*3D, TPMT*4, TPMT*5, TPMT*6, TPMT*7*	RC	[24]
von Ahsen	2005	Not specified	Caucasian	*TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c*	RC	[65]

Study	Year	Amplification/genotype method	Population	Polymorphisms tested	Phenotype method	Ref.
Wennerstrand	2013	Pyrosequencing	Other	*TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c, TPMT*3D*	RC	[16]
Winter	2007	PCR; RFLP		*TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c*	Mass spectrometry	[10]
Wusk	2004	PCR; sequencing	German	*TPMT*2, TPMT*3b, TPMT*3c*	HPLC	[11]
Xin	2009	AS-PCR; PCR; RFLP	Not specified	*TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c*	HPLC	[36]
Yates	1997	PCR; RFLP	Caucasian	*TPMT*1, TPMT*2, TPMT*3a, TPMT*3c*	RC	[66]
Zhang	2007	PCR; RFLP	Not specified	*TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c*	HPLC	[37]

ARMS: Multiplex amplifications refractory mutation; AS-PCR: Allele-specific PCR; DHPLC: Denaturing high performance liquid chromatography; RC: Radiochemical method; SSCP: Single strand conformation polymorphism.

Table 4

Genotype laboratory methods for high-quality genotype–genotype studies (n = 6).

Study	Year	Index method	Reference method(s)	Polymorphisms tested	Ref.
Chowdhury	2007	Microchip RFLP	Conventional RFLP and AS-PCR; integrated microchip PCR and AS-PCR; TaqMan SNP genotyping	*TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c*	[71]
Kim	2013	AS-PCR	PCR	*TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c*	[75]
Lu	2005	APEX	ARMS-PCR; PCR-RFLP	*TPMT*3b, TPMT*3c, TPMT*6*	[76]
Ma	2003	PCR + DHPLC	PCR + RFLP; PCR + SNaPshot sequencing with direct DNA sequencing; AS-PCR	*TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c, TPMT*3d*	[70]
Roman	2012	LightSNiP	Traditional PCR and Sanger sequencing	*TPMT*2, TPMT*3b, TPMT*3c*	[15]
Schaeffeler	2008	MALDI-TOF MS	DHPLC	*TPMT*2, TPMT*3a, TPMT*3c, TPMT*9, TPMT*11, TPMT*16, TPMT*17, TPMT*18, TPMT*20, TPMT*22*	[74]

APEX: Arrayed primer extension technology; AS-PCR: Allele-specific PCR; DHPLC: Denaturing high performance liquid chromatography; MS: Mass spectrometry; TMPT: Thiopurine s-methyltransferase.

Table 5

Diagnostic test performance calculated from 2 × 2 tables for presence of homozygous deficient genotype.

Study	Year	Calculated sensitivity	Calculated specificity	Calculated PPV	Calculated NPV	Ref.
TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c						
Ben Salah	2013	100.0%	100.0%	100.0%	100.0%	[13]
Fangbin	2012	†	100.0%	†	100.0%	[26]
Ganiere-Monteil	2004	100.0%	100.0%	100.0%	100.0%	[58]
Gazouli	2012	33.3%	98.9%	85.7%	88.1%	[29]
Serpe	2009	100.0%	100.0%	100.0%	100.0%	[35]
von Ahsen	2005	†	100.0%	†	100.0%	[65]
Winter	2007	0.0%	100.0%	†	99.2%	[10]
Xin	2009	†	100.0%	†	100.0%	[36]
Zhang	2007	†	100.0%	†	100.0%	[37]
TPMT*2, TPMT*3a, TPMT*3c						
Fakhoury	2007	100.0%	100.0%	100.0%	100.0%	[25]
Ford	2009	†	†	†	†	[28]
Hindorf	2012	86.0%	100.0%	100.0%	99.9%	[30]
Liang	2013	†	100.0%	†	100.0%	[14]
Ma	2006	†	100.0%	†	100.0%	[39]
Marinaki	2003	†	100.0%	†	100.0%	[60]
TPMT*2, TPMT*3b, TPMT*3c						
Larussa	2012	16.7%	97.8%	50.0%	89.8%	[31]
Milek	2006	†	100.0%	†	100.0%	[33]
Wusk	2004	†	.	†	†	[11]
TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c, TPMT*3d						
Schaeffeler	2004	100.0%	100.0%	100.0%	100.0%	[62]
Schwab	2002	100.0%	100.0%	100.0%	100.0%	[64]
Wennerstrand	2013	†	100.0%	†	100.0%	[16]
TPMT*3a, TPMT*3b, TPMT*3c						
Langley	2002	0.0%	100.0%	†	98.1%	[59]

Study	Year	Calculated sensitivity	Calculated specificity	Calculated PPV	Calculated NPV	Ref.
Lennard	2013	†	†	†	100.0%	[32]
TPMT*2, TPMT*3						
Ford	2006	100.0%	100.0%	100.0%	100.0%	[27]
TPMT*1, TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c						
Jorquera	2012	†	100.0%	†	100.0%	[38]
TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c, TPMT*6						
Loennechen	2001	†	100.0%	†	100.0%	[23]
TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c, TPMT*3d, TPMT*8						
Oselin	2006	†	†	†	†	[34]
TPMT*1, TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c, TPMT*1S, TPMT*1A, TPMT*7, TPMT *3d						
Spire-Vayron de la Moureyre	1998	100.0%	100.0%	100.0%	100.0%	[67]
TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c, TPMT*3D, TPMT*4, TPMT*5, TPMT*6, TPMT*7						
Spire-Vayron de la Moureyre	1998	100.0%	100.0%	100.0%	100.0%	[24]
TPMT*1, TPMT*2, TPMT*3a, TPMT*3c						
Yates	1997	100.0%	100.0%	100.0%	100.0%	[66]

†Unable to calculate.

NPV: Negative predictive value; PPV: Positive predictive value.

Table 6

Diagnostic test performance calculated from 2 × 2 tables for presence of (homozygous or) heterozygous deficient genotype.

Study	Year	Calculated sensitivity	Calculated specificity	Calculated PPV	Calculated NPV	Ref.
TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c						
Ben Salah	2013	100.0%	100.0%	100.0%	100.0%	[13]
Fangbin	2012	38.5%	100.0%	100.0%	97.3%	[26]
Ganiere-Monteil	2004	92.7%	100.0%	100.0%	99.3%	[58]
Gazouli	2012	52.2%	100.0%	100.0%	73.8%	[29]
Serpe	2009	13.4%	98.3%	78.8%	70.3%	[35]
von Ahsen	2005	†	100.0%	†	75.8%	[65]
Winter	2007	64.7%	100.0%	100.0%	95.0%	[10]
Xin	2009	29.2%	100.0%	100.0%	88.1%	[36]
Zhang	2007	36.8%	100.0%	100.0%	95.6%	[37]
TPMT*2, TPMT*3a, TPMT*3c						
Fakhoury	2007	29.3%	97.5%	85.7%	72.6%	[25]
Ford	2009	†	†	†	†	[28]
Hindorf	2012	69.5%	98.8%	89.5%	95.5%	[30]
Liang	2013	60.0%	98.7%	90.0%	92.8%	[14]
Ma	2006	67.7%	99.8%	95.5%	98.4%	[39]
Marinaki	2003	55.6%	100.0%	100.0%	95.0%	[60]
TPMT*2, TPMT*3b, TPMT*3c						
Larussa	2012	22.2%	97.0%	80.0%	69.6%	[31]
Milek	2006	50.0%	97.6%	75.0%	93.1%	[33]
Wusk	2004	†	†	†	†	[11]
TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c, TPMT*3d						
Schaeffeler	2004	86.8%	99.4%	94.9%	98.4%	[62]
Schwab	2002	100.0%	96.6%	62.5%	100.0%	[64]
Wennerstrand	2013	17.4%	100.0%	100.0%	59.6%	[16]
TPMT*3a, TPMT*3b, TPMT*3c						
Langley	2002	66.7%	90.9%	60.0%	93.0%	[59]

Study	Year	Calculated sensitivity	Calculated specificity	Calculated PPV	Calculated NPV	Ref.
Lennard	2013	†	†	†	92.2%	[32]
TPMT*2, TPMT*3						
Ford	2006	80.6%	98.1%	80.6%	98.1%	[27]
TPMT*1, TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c						
Jorquera	2012	83.3%	99.5%	93.8%	98.4%	[38]
TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c, TPMT*6						
Loennechen	2001	95.8%	100.0%	100.0%	99.6%	[23]
TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c, TPMT*3d, TPMT*8						
Oselin	2006	†	†	†	†	[34]
TPMT*1, TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c, TPMT*1S, TPMT*1A, TPMT*7, TPMT *3d						
Spire-Vayron de la Moureyre	1998	83.3%	94.1%	93.8%	84.2%	[67]
TPMT*2, TPMT*3a, TPMT*3b, TPMT*3c, TPMT*3D, TPMT*4, TPMT*5, TPMT*6, TPMT*7						
Spire-Vayron de la Moureyre	1998	54.5%	94.3%	66.7%	90.9%	[24]
TPMT*1, TPMT*2, TPMT*3a, TPMT*3c						
Yates	1997	96.3%	100.0%	100.0%	95.6%	[66]

†Unable to calculate.

NPV: Negative predictive value; PPV: Positive predictive value.