

For reprint orders, please contact: reprints@futuremedicine.com

Evaluation of predictive *CYP2C19* genotyping assays relative to measured phenotype in a South African cohort

Aim: To align predicted and measured *CYP2C19* phenotype in a South African cohort. **Materials & methods:** Genotyping of *CYP2C19**2, *3, *9, *15, *17, *27 and *28 was performed using PCR-RFLP, and an activity score (AS) system was used to predict phenotype. True phenotype was measured using plasma concentrations of omeprazole and its metabolite 5'-hydroxyomeprazole. **Results:** Partial genotype-phenotype discrepancies were reported, and an adapted AS system was developed, which showed a marked improvement in phenotype prediction. Results highlight the need for a more comprehensive *CYP2C19* genotyping approach to improve prediction of omeprazole metabolism. **Conclusion:** Evidence for the utility of a *CYP2C19* AS system is provided, for which the accuracy can be further improved by means of comprehensive genotyping and substrate-specific modification.

Keywords: activity score system • *CYP2C19* • genotype–phenotype correlation • omeprazole • South Africa

Adverse drug reactions (ADRs) are debilitating. They are associated with significant morbidity and mortality and have important financial implications [1,2]. Pharmacogenetics has become an important area of research due to the fact that ADRs, at least in part, stem from interindividual genetic variability [3]. Of hospitalized patients in South Africa, 14% have been estimated to experience ADRs, and when compared with hospitals in the USA and UK, a five to tenfold higher fatality rate has been observed [4]. It has been reported that pharmacogenetics may have the potential to reduce ADRs by 10–20% and improve drug efficacy by 10–15% [5]. Since it is recognized that sub-optimal therapies are prevalent in developing countries such as South Africa, pharmacogenetics could have a positive impact.

Genetic variation in *CYP2C19* (OMIM 124020) has been associated with variable responses to numerous drugs including clopidogrel (anticoagulant), citalopram (and many other antidepressants), proguanil (prophylactic antimalarial) and omeprazole (and

other proton pump inhibitors) [6–12]. Many *CYP2C19* polymorphisms have been identified, which include 34 distinct alleles as listed by the P450 Nomenclature Committee [13]. From *in vivo* and *in vitro* data, many of these alleles have been assigned an estimated level of enzyme activity. This can be used to predict poor (PM), intermediate (IM), extensive (EM) and ultrarapid (UM) metabolizer phenotypes.

The nonfunctional *CYP2C19**2 and *3 alleles are the most commonly genotyped and have been prioritized in many *CYP2C19* studies [14]. Responsible for increased *CYP2C19* expression, *CYP2C19**17 is important for the prediction of UMs and has been found to be relatively frequent in certain populations [15–20]. The above-mentioned alleles have been prioritized based on observations in cohorts of predominantly European and Asian descent. This approach may not be sufficiently comprehensive in the context of African populations, and many African-specific alleles, such as *CYP2C19**9 [21,22], may be important in this context.

Tyren M Dodgen^{1,2,3}, Britt I Drögemöller⁴, Galen EB Wright⁴, Louise Warnich⁴, Francois E Steffens⁵, A Duncan Cromarty¹, Marco Alessandrini^{2,3} & Michael S Pepper^{*2,3,6}

¹Department of Pharmacology, University of Pretoria, Pretoria, South Africa

²Department of Immunology, Faculty of Health Sciences, University of Pretoria, PO Box 2034, Pretoria 0001, South Africa

³Institute for Cellular & Molecular Medicine, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa

⁴Department of Genetics, Faculty of Science, Stellenbosch University, Stellenbosch, South Africa

⁵Department of Statistics, Faculty of Natural & Agricultural Sciences, University of Pretoria, Pretoria, South Africa

⁶Department of Genetic Medicine & Development, Faculty of Medicine, University of Geneva, Switzerland

*Author for correspondence:

Tel.: +27 12 319 2179

Fax: +27 12 319 2946

michael.pepper@up.ac.za

South Africa is home to some of the most diverse and unique African populations [23], as illustrated by its 11 official languages and many more distinct ethnic groups and cultures [24]. This diversity is paralleled by extensive genetic diversity, which is important when the goal is to establish pharmacogenetic treatment plans in South Africa. To date, very few studies examining *CYP2C19* variation in South African populations have been published, and there is no data assessing the relationship between predicted and measured phenotype. Previously investigated cohorts include the Venda [25,26], Xhosa and Cape Mixed Ancestry (Cape Colored) populations [20,27], comprising 2.3, 17.6 and 8.9% of the South African population, respectively ([20] Statistics South Africa Census 2011 [24]). The focus of these studies was on the *CYP2C19**2 and *3 alleles. More recently, however, Drögemöller *et al.* [20] studied the Xhosa (n = 100) and Cape Colored (n = 75) populations, detecting *CYP2C19**2, *3, *9, *15, *17, as well as the novel *27 and *28 alleles (the majority of which remain uncharacterized *in vivo*).

From these studies, a number of interesting observations can be made. The predominantly Asian allele, *CYP2C19**3, was only detected in the Cape Colored population, reflecting the high level of admixture [20]. The absence of *CYP2C19**10 and an alternative link-

age disequilibrium pattern found for *CYP2C19**17, emphasized the differences in genetic composition between African-Americans and the Xhosa from South Africa. This, along with allelic frequency differences between the Venda, Xhosa and Cape Coloreds, reiterates the importance of acknowledging diversity within Africa and those of the African diaspora [20]. More comprehensive African-specific genotyping strategies should, therefore, be considered when evaluating pharmacogenetic variation in South Africans. To this end, we investigated the predicted and *in vivo* measured metabolism for *CYP2C19* in a demographically representative South African cohort, and aligned these to evaluate the utility of the predictive tool.

Materials & methods

Study subjects

The study was approved by Research Ethics Committee of the Faculty of Health Sciences, University of Pretoria (approval number 24/2007), and informed consent was obtained from all individuals prior to inclusion. This involved an explanation of the reasons for the study, the potential risks involved and the rights of the individual. A detailed document was then provided and written consent was obtained from each participant. Unrelated adult volunteers of both genders (excluding pregnant or breast-feeding females) were recruited to approximately represent the demographics of the South African population. This included 70 Black African, ten White Caucasian, ten South African Colored, and ten Indian volunteers. The sample was not chosen to compare inter-ethnic differences, but to represent that which could potentially be seen in South African clinical practice. The term 'South African Colored' is officially used to describe a complex and unique admixed group of people predominantly residing in the western Cape (where they are referred to as 'Cape Coloreds'), who have descended from many different ethnic ancestries, including European, Asian and other African populations [28,29]. At the time of sampling, the cohort had an average age of 34.4 ± 9.8 years (range 19–58), BMI of 26.1 ± 5.8 , waist circumference of 86.9 ± 13.3 cm, and 55% were female. Each of the 11 official languages were represented in the study cohort, and none of the study participants reported being on medication.

Sample collection

Peripheral blood was collected in ethylenediaminetetraacetic acid (EDTA) vacutainer tubes (Becton-Dickinson, NJ, USA) for baseline quantification of analytes during the phenotype assays and genomic DNA



 **GenomicsNet**

VOICE OF THE EXPERTS

Collaborate and share insights
through blogs and videos

Join today
www.GenomicsNet.com



Table 1. Allelic activity and CYP2C19 phenotype prediction.

Estimated metabolic potential of alleles			
Allele activity	Allelic score	Defining alleles [†]	Modified allele designation [‡]
Increased	2.0	*17	
Normal	1.0	*1+,*28	*1+,*17,*27,*28
Decreased	0.5	*9,*27	*9,*15
Absent	0	*2,*3	*2,*3
Unknown	1.0	*15	
Phenotype prediction			
Activity score		Prediction	
>2.0		Ultrarapid metabolizer	
1.5 and 2.0		Extensive metabolizer	
0.5 and 1.0		Intermediate metabolizer	
0		Poor metabolizer	
The activity score was adopted from Gaedigk <i>et al.</i> [31] to predict CYP2C19 activity.			
*Allelic activities assigned based on information provided in [30].			
*Modification based on omeprazole metabolism findings in this study.			

(gDNA) was extracted using the Genomic DNA Purification Kit (Fermentas Life Science, Lithuania) or the automated Maxwell® 16 system (Promega, WI, USA).

CYP2C19 genotyping

CYP2C19 was evaluated using PCR-RFLP analysis as described previously [20], which entailed the amplification of eight exonic regions and enzymatic digestion to identify the allele-defining SNVs for CYP2C19*2, *3, *9, *15, *17, *27 and *28. Identified alleles were named according to the Human Cytochrome P450 (CYP) Allele Nomenclature Committee's online database [30].

Phenotype prediction

The activity score (AS) concept, initially developed to predict CYP2D6 phenotype [31], was adapted for CYP2C19 phenotype prediction [32]. Activity scores for the assayed CYP2C19 alleles were based on the known enzyme activities reported in the Human Cytochrome P450 (CYP) Allele Nomenclature Committee's online database for CYP2C19 alleles [13]. Table 1 illustrates how the AS system was adapted to predict CYP2C19 metabolism, and how it was further modified to improve prediction of omeprazole metabolism.

CYP2C19 omeprazole phenotyping

CYP2C19 enzyme activity was determined in volunteers using omeprazole (OME) as the probe drug and 5'-hydroxyomeprazole (5OH) as the CYP2C19 specific metabolite. Volunteers were confirmed to be overnight-fasted or to have not eaten at least 3 h prior to commencement of sampling. Forty milligrams (40 mg) of ALTOSEC (omeprazole) was administered and

blood samples were drawn prior to dosage, and at 2, 3 and 4 h after administration. Plasma concentrations of OME and 5OH were quantified simultaneously using an automated online solid phase extraction tandem mass spectrometry (LC-MS/MS) method, as described previously by our group [33]. A metabolic ratio for plasma concentration of metabolite over probe drug ($MR_{(5OH/OME)}$) was used to indicate CYP2C19 enzyme activity. In this study, we used the 5OH/OME ratio as opposed to the traditional OME/5OH ratio. By reporting MR in this format, it is possible to present the predicted phenotype as a positive correlation, as opposed to it being inversely proportional. Therefore, the \log_{10} transformation of the ratio would then simply be a change of sign and would still be easy to compare (i.e., $\log_{10}(10/1) = 1$ and $\log_{10}(1/10) = -1$).

Statistical analysis

Statistical analyses were performed using Statistical Package for the Social Sciences version 20.0 (SPSS Inc., IL, USA). To establish the most appropriate time point for phenotypic measurement, the metabolic ratios for each time point were correlated to one another using Pearson's correlation coefficient. When comparing MRs of the various AS groups, the nonparametric Kruskal Wallis analysis of variance (ANOVA) and pairwise comparisons were performed. Correlation of MRs to AS groups was also performed with the Spearman's correlation.

Optimal $MR_{(5OH/OME)}$ antinodes/cut-offs to distinguish between PM, IM and EM were evaluated using receiver-operator characteristic (ROC). Antinode values were assigned based on obtaining 100%

specificity for PMs and by comparing sensitivity and specificity values to find the optimum for both to discriminate between IMs and EMs. Baseline characteristics (age, BMI and waist circumference) were compared with phenotype using ANOVA and gender was compared with phenotype using a Chi-square analysis.

In addition to the Kappa statistical test for conformation, a custom ‘discrepancy analysis’ was used to calculate ‘inter-rater discrepancies’ using absolute differences to test for conformation. With the four metabolizing categories being ordinal according to activity (UM > EM > IM > PM), the calculation measures the average deviation from a perfect conformation. This can also be explained as the closeness of the predicted phenotype to measured phenotype over the entire cohort. Perfect conformation would receive a value of 0, while one, two and three deviating categories would receive values of 1.0, 2.0 and 3.0, respectively. All values were added and subsequently divided by the number of individuals in the cohort (n). Sensitivity, specificity, and positive and negative predictive values were calculated as a second measure of predictive accuracy. p-values of less than 0.05 were considered to be statistically significant.

Results
CYP2C19 genotyping

The genotyping protocol adopted for this study revealed the presence of *CYP2C19**2, *9, *15, *17 and *27 (Table 2). The wildtype allele (*1) was most frequent at 36.5%, while *CYP2C19**3 and *28 were not reported in this cohort – which is in agreement with the either absent or very low frequencies reported in other populations of African descent [20,22]. Nineteen *CYP2C19* genotypes were determined to be present in the cohort (Table 2). The most frequently reported genotypes were *CYP2C19**1/*17 (17%), *1/*1 (16%), and *1/*2 and *2/*27 which were each reported in 11% of individuals.

CYP2C19 *in vivo* phenotyping with omeprazole as probe substrate

Phenotyping data from the 2-h sampling time point had a 0.849 and 0.789 correlation with the 3- and 4-h samplings, respectively. Correlation of phenotyping data between the 3- and 4-h time points was 0.910, which indicated that these would be the most appropriate time points for further investigation. Given that a 3-h sampling would be more convenient for both investigator and study participant, this time point was selected as the phenotyping time point of choice.

Table 2. CYP2C19 genotypic data.					
Allelic frequencies			Genotypic frequencies		
Allele	n	Frequency (%)	Genotype	n	Frequency (%)
*1	73	36.5	*1/*1	16	16.0
*2	40	20.0	*1/*15	2	2.0
*9	5	2.5	*1/*17	17	17.0
*15	10	5.0	*1/*2	11	11.0
*17	31	15.5	*1/*27	9	9.0
*27	41	20.5	*1/*9	2	2.0
			*2/*15	3	3.0
			*2/*17	2	2.0
			*2/*2	6	6.0
			*2/*27	11	11.0
			*2/*9	1	1.0
			*27/*27	7	7.0
			*9/*17	1	1.0
			*9/*27	1	1.0
			*15/*15	1	1.0
			*15/*17	1	1.0
			*15/*27	2	2.0
			*17/*17	3	3.0
			*17/*27	4	4.0
Total	200	100.0	–	100	100.0

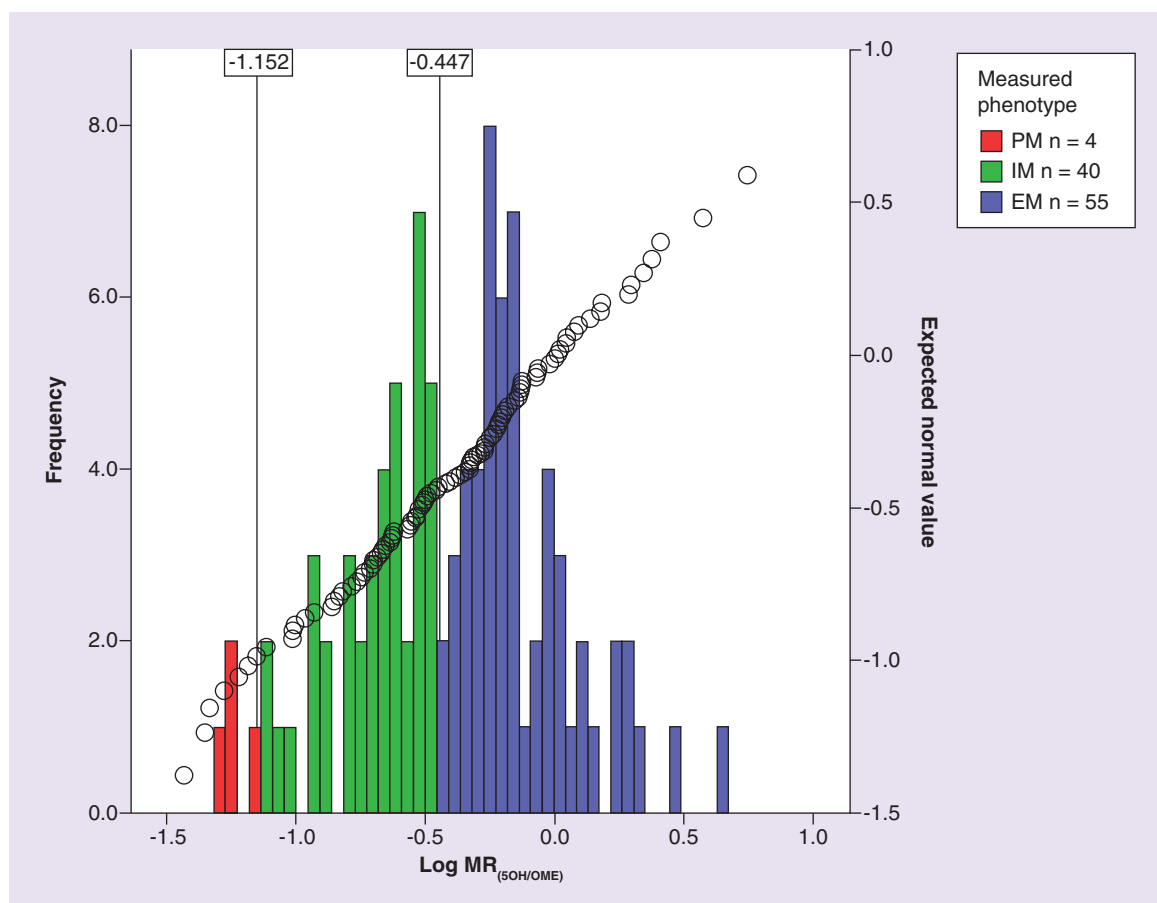


Figure 1. Phenotypic distribution of omeprazole metabolism by *CYP2C19* in a demographically representative sample of the South African population. Phenotype was measured as the metabolic ratio of the concentration of 5OH (metabolite) over OME (probe drug) in a plasma sample taken 3 h after probe drug administration. The graph has been stratified into PMs, IMs and EMs. The solid x-axis division lines represent the proposed antimodes to discriminate PM from IM and IM from EM individuals. The superimposed QQ-plot agrees with these antimodes. The number in parenthesis indicates the number of individuals in each phenotypic group. 5OH: 5'-hydroxyomeprazole; EM: Extensive metabolizer; IM: Intermediate metabolizer; MR: Metabolic ratio; OME: Omeprazole; PM: Poor metabolizer.

Figure 1 illustrates the log $MR_{(5OH/OME)}$ distribution in the sampled cohort. The modified AS in combination with ROC analysis identified a trimodal distribution (distinguishing PMs, IMs and EMs), from which phenotypic cut-offs could be generated. $MR_{(5OH/OME)} = -1.152$ (66.7% sensitivity and 100.0% specificity) was established to discriminate between PM and IM phenotypes. A second cut-off could be established at $MR_{(5OH/OME)} = -0.447$ (81.8% sensitivity and 77.3% specificity) to discriminate between IM and EM phenotypes. The superimposed QQ-plot over the histogram in Figure 1 is in agreement with the use of these antimodes as cut-offs to differentiate metabolizer status. The slight plateau (inflection point) coinciding with each antimode is evidence of agreement. No UMs were identified in this cohort. Using these cut-off points and modified AS, four individuals were identified to be PMs of *CYP2C19*, 40 IMs and 55 EMs.

Comparison between predicted & measured phenotype for *CYP2C19*

Graphic representations of the original/adopted and modified phenotypic predictions, relative to the measured phenotype, are illustrated in Figure 2. In general, there is a wide range of metabolic ratios within the different predicted phenotypic groups and fairly extensive overlapping of each of the groups. Each grouping was, however, visually distinct from the other, suggesting that the *CYP2C19* genotype (based on the investigated alleles) is a reliable measure of OME metabolism, and can hence be used to aid in the prediction of *CYP2C19* metabolic activity.

Using the original/adopted AS system, the predicted UM group represented individuals that were phenotypically measured to have metabolic activities similar to that of the EM group (Figure 2). It was, therefore, apparent that the increased metabolism allele,

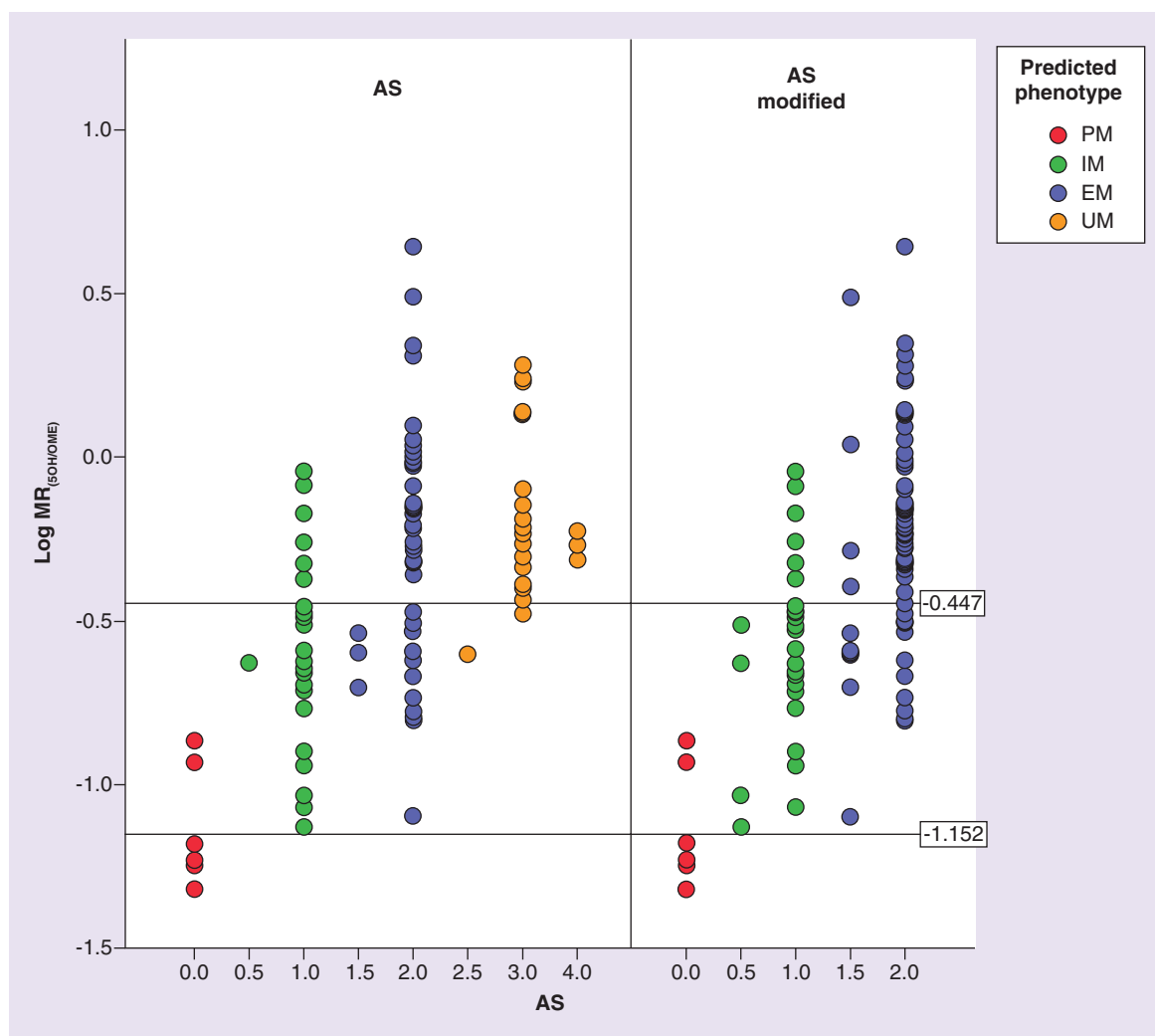


Figure 2. Comparison of predicted phenotype generated from CYP2C19 genotype (using traditional and modified prediction systems) to CYP2C19 measured phenotype. Phenotype prediction included an adapted AS system [31] and an AS system modified based on omeprazole metabolism (RFLP + AS mod.). Phenotypic groupings included PM, IM, EM and UM metabolism depending on the capability of the platform and system used for prediction. The solid line represents phenotypic cut-offs separating phenotypic PM, IM and EM from bottom to top. 5OH: 5'-hydroxyomeprazole; AS: Activity score; EM: Extensive metabolizer; IM: Intermediate metabolizer; MR: Metabolic ratio; OME: Omeprazole; PM: Poor metabolizer; UM: Ultrarapid metabolizer.

*CYP2C19**17, had no influence on OME metabolism in this cohort. For this reason, the allele was reassigned to having a normal activity in the context of OME metabolism (Table 1). In an attempt to elucidate the *in vivo* activity of *CYP2C19**15 and *27, all genotypes including these alleles were plotted against their respective MRs (Figure 3). Given the pairing of these alleles with those of known metabolic activity, it was determined that *CYP2C19**15 and *27 in all likelihood contribute to reduced and normal CYP2C19 activity, respectively. It is, therefore, recommended that these alleles be assigned AS values of 0.5 and 1.0 in the context of OME metabolism, respectively (Table 1). Although not specifically illustrated here, the decreased functionality of *CYP2C19**9 was confirmed in this study. By apply-

ing the modified AS system and aligning it with MRs, a 0.616 correlation coefficient ($p < 0.001$) was reported when performing the Spearman's correlation. This was supported by a statistically significant Kruskal Wallis ANOVA ($p < 0.001$) when comparing the AS groups, and statistically significant differences when comparing the AS = 2.0 group to AS = 0.0 ($p < 0.001$), AS = 0.5 ($p = 0.022$) and AS = 1.0 ($p < 0.001$).

Cross-correlation of the ordinal phenotypic groups allowed for further evaluation of the relationship between predicted and measured phenotype (Figure 4). The modified version of the AS system was the most accurate predictor of *in vivo* OME phenotype. Both Kappa and inter-rater discrepancy scores favored the modified AS. The modified AS was also superior when

it came to sensitivity and specificity, as well as positive and negative predictive value testing. This data, therefore, confirmed that the modified AS system presented here was a more accurate predictor of CYP2C19 phenotype when examining the metabolism of OME. Further details of this comparison are provided in [Supplementary Table 1](#) (please see online at www.futuremedicine.com/doi/full/10.2217/PGS.15.80).

Potential confounding factors for CYP2C19 metabolism

Of the additional demographic parameters that were recorded for the study cohort, none were shown to have an influence on CYP2C19 metabolism. Using a one-way ANOVA, p-values of 0.694, 0.689 and 0.624 were calculated for age, BMI and waist circumference,

respectively, indicating no statistically significant differences when stratifying the data accordingly. A Pearson's Chi-squared test was performed to investigate the potential influence of gender on OME metabolism, which reported a p-value of 0.186 and hence indicated no significant influence. Lastly, since no participant indicated that they were on medication, drug interactions are unlikely to have influence on the study results.

Discussion Phenotyping

The use of a single plasma concentration for phenotyping is more convenient for the volunteer/patient than a 4–8-h urine accumulation. It also avoids the complications associated with Phase II metabolism, preferential urinary accumulation of the metabolite

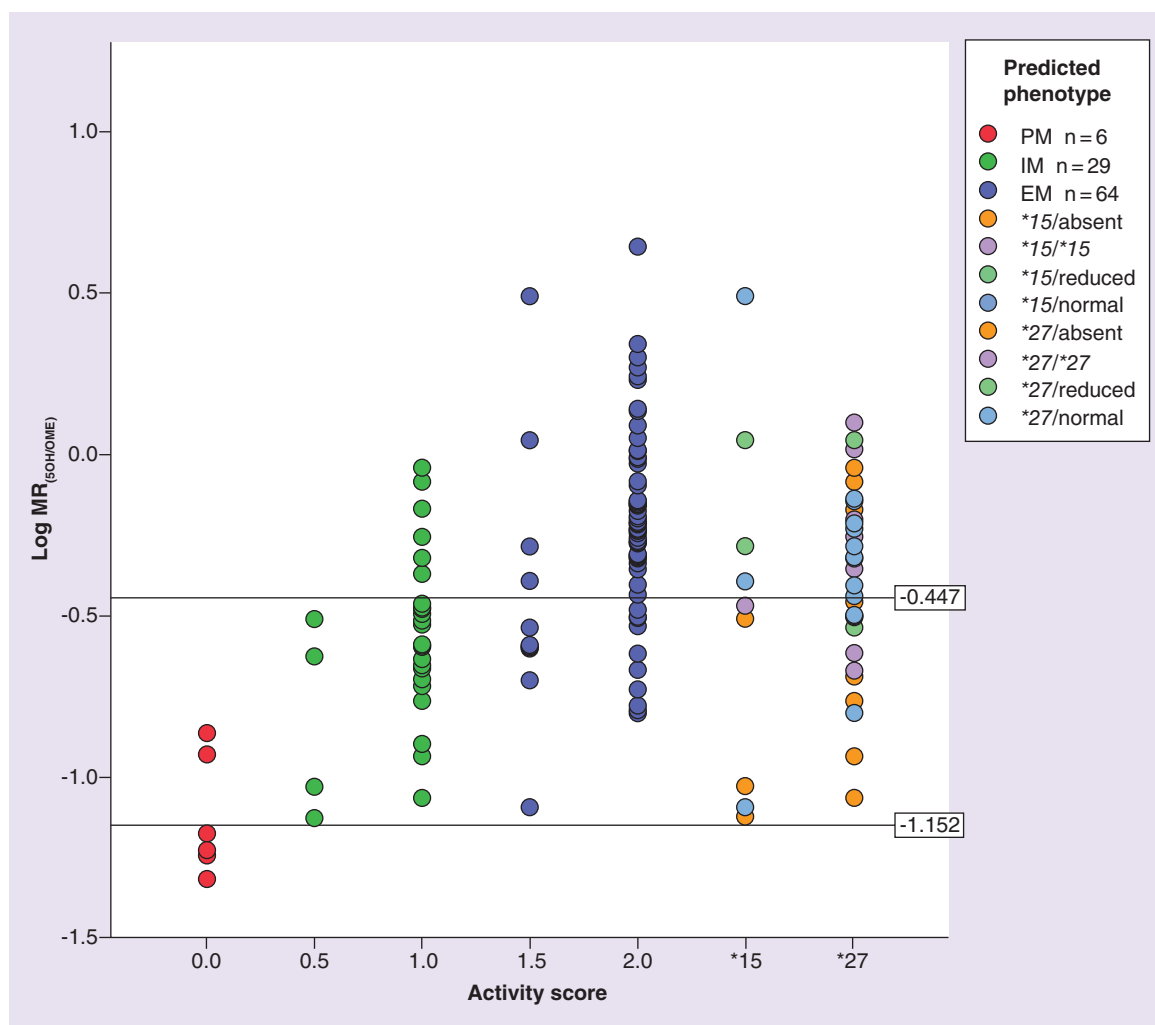


Figure 3. *In vivo* elucidation of the metabolic activity of CYP2C19*15 and *27 (unknown activity) using the comparison of predicted phenotype (as determined by the modified activity score system) to measured metabolic activity. Lines represent antimodes which indicate phenotypic cut-offs from bottom to top separating PM, IM and EM metabolism.

5OH: 5'-hydroxymepazazole; EM: Extensive metabolizer; IM: Intermediate metabolizer; MR: Metabolic ratio; OME: Omeprazole; PM: Poor metabolizer.

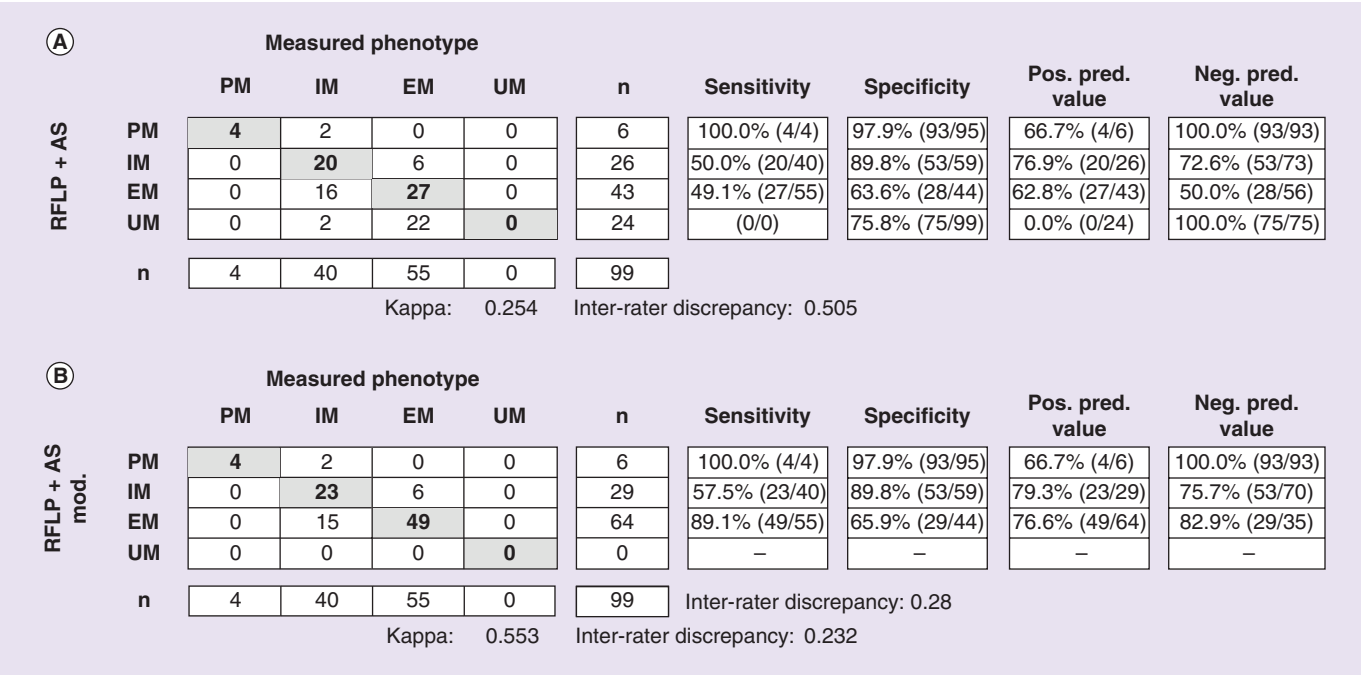


Figure 4. Evaluation of the ability to predict phenotype using CYP2C19 genotype when compared to measured phenotype (in vivo metabolic ratios) using omeprazole as a probe drug. Phenotype was predicted from (A) an adapted activity score system [31] based on known CYP2C19 allele activities [13] and (B) using a modified activity score system based on omeprazole metabolism. The grey blocks represent a perfect correlation. A diagonal line through the metabolizer group represents the inability of the prediction system to predict a specific metabolizer group. The inter-rater discrepancy statistical test was calculated to measure agreement. Sensitivity, specificity, positive predictive value (pos. pred.) and negative predictive value (neg. pred.) were calculated to measure the ability to predict phenotype. ‘—’ represents the inability to calculate the measure of performance as the metabolic class was not predicted by the system.

AS: Activity score; EM: Extensive metabolizer; IM: Intermediate metabolizer; PM: Poor metabolizer; UM: Ultrarapid metabolizer.

and variation in urinary pH [33]. By correlating 2-, 3- and 4-h data intervals, it was determined that a 3-h sampling would be the most reliable and convenient time point to use in future investigations. The 3-h time point has previously been used to phenotype CYP2C19 using OME and 5OH concentrations in plasma [34,35]. Sampling should be performed after the T_{max} , so as to avoid increasing plasma levels of the drug interfering with measurements. The T_{max} for OME is generally between 2 and 3 h, although it may vary from 30 min to 6 h, depending on various factors, including the formulation of the drug [36–40]. In this study, a capsule-based formulation of OME was used, which is known to result in an earlier T_{max} than a tablet formulation [39].

Previous studies have reported CYP2C19 phenotypic cut-off values, separating PMs from IMs/EMs, to range between $\log MR_{(5OH/OME)}$ -0.6 and -3.04 [35,41,42]. In the study by Ramsjö *et al.* [41], it was reported that the cut-off in Koreans was as high as -3.04, while it was a low -0.86 in a Swedish cohort. The cut-off value separating the PM and IM phenotypes in this study was -1.152, which is within the range of cut-off values previously reported. The disparities reported by various studies, including our own, may be due to several

factors, of which ethnicity and the associated genetic variability are likely to be important.

CYP2C19 comparison between predicted & measured phenotype

To reduce the complications associated with phenotype prediction, the numeric-based AS system was adopted in this study [31]. This approach enables a cohort to be stratified into additional groups, which may offer insight into quartiles of a cohort, while also minimizing overlapping phenotypes within a group. Stratification allows phenotypic patterns to become apparent. Caution should, however, be taken when implementing AS phenotype prediction as ethnicity, disease state (including liver and kidney function) and concomitant medication may contribute to an unpredictable phenotype [31].

A comprehensive genotyping approach would be essential to accurately predict CYP2C19 phenotype in South African individuals. To achieve this, additional alleles such as the decreased function CYP2C19*9 and *15 alleles were accounted for in this study. The former allele was of particular relevance, as it is known to be African-specific [21,22] and indeed contributed to the improved stratification of CYP2C19 phenotypes.

Even though the genotyping assay used in this study was tailored for a South African cohort, its predictive ability was weaker than expected. However, by further adapting the AS system for use of omeprazole as a probe drug (Table 1), a significant improvement in predictive power was demonstrated (Figure 4). This modification was particularly relevant for the well-studied, increased metabolism *CYP2C19**17 allele [34,35], which contrary to expectations was observed to have no influence on OME metabolism in this cohort. This may suggest that OME might not be sensitive enough as a probe drug for *CYP2C19* metabolism.

A possible confounding factor may be the role of, or the extent to which *CYP3A4* also metabolizes OME. Plasma concentrations of omeprazole sulphone, a *CYP3A4* metabolite, could have been measured to evaluate if this was indeed the case, but this was not within the scope of the project. Kearns *et al.* [18] highlight this dilemma and suggest pantoprazole (also a proton pump inhibitor) as a more selective probe drug for *CYP2C19* phenotypic analysis. Metabolism of escitalopram, an antidepressant and selective serotonin reuptake inhibitor, has also been shown to not be influenced by *CYP2C19**17, and is hence another example of where metabolism predictions may be substrate specific [17]. The role of substrate specificity has also been brought into question for another CYP enzyme, *CYP2D6*, where the reduced function *CYP2D6**17 allele metabolizes risperidone with full function, and therefore a shift in AS value of 0.5–1.0 would be appropriate when predicting risperidone metabolism [31,43].

Although this study did not validate the functionality of the 5' upstream variants, it provided, for the first time, phenotypic data for *CYP2C19**15, demonstrating reduced metabolism (Figure 2). The allele was first described in Black Africans by Blaisdell *et al.* [21], where the defining SNV, 55A>C (rs17882687), was found to be in linkage disequilibrium with 80161A>G (rs3758581). The former mutation was not reconstructed to elucidate activity, since the change is near the N-terminus, which is thought to have no influence on activity. This same variant (55A>C) was subsequently found to be in complete linkage equilibrium with -2030C>T (rs11316468) [44]. *In silico* analysis of -2030C>T, utilizing three separate programs, consistently predicted the removal of a GATA factor binding site, and although dual reporter luciferase assays did not detect a significant change in expression, the results suggested a trend towards decreased expression [20]. The role that GATA binding sites play in *CYP2C19* expression was further validated by Mwinyi *et al.* [45], and hence removal of a GATA factor binding site (due to -2030C>T) could provide an explanation for *CYP2C19**15 displaying reduced metabolic activity in this cohort.

Initially, *CYP2C19**27 appeared to have reduced phenotypic activity (Figure 2). However, once the AS was adjusted from 0.5 to 1.0 (Table 1), the Kappa score decreased and the inter-rater discrepancy value increased. This *in vivo* demonstration contradicts the *in vitro* dual luciferase results that suggest reduced expression as a result of the -1041G>A promoter variant. Unknown confounding factors may also be masking the marginally reduced activity of *CYP2C19**27. A subset of volunteers that were homozygous for this allele appeared to have an IM phenotype; however, this seemed to be sporadic as others presented with EM phenotypes. Evaluation of this allele with other substrates may result in a different metabolic activity, which may also assist in further elucidating the effects of this variant. Since this allele is frequent – 16.4% in Black South Africans [32], 10% in Xhosa and 14% in South African Coloreds [20] – it may still play an important role in improving predicted phenotype in the South African population.

Although examining a greater number of alleles and modifying the AS system according to OME did improve the predictive ability of the genotyping assays, this system remains imperfect. The wide range of metabolic ratios within each of the predicted phenotypic groups, the overlap of these groups, and the lack of good correlation (Figures 1 & 2) reveals discrepancies in the predicted-measured phenotypes of *CYP2C19*. These discrepancies could be explained, in part, by environmental factors such as diet [46], which was not considered in this study, and which could either induce or inhibit *CYP2C19* activity. Disease status, such as HIV/AIDS, has been found to affect *CYP2D6* metabolic activity [47,48] and could potentially have a similar effect on *CYP2C19* activity. Although not the focus of this study, HIV could be investigated to improve phenotype prediction and provide an explanation for the discordant results [31]. Epigenetic effects should also be considered, as these may play an important role, given that there is a confirmed CpG island in intron 1 of *CYP2C19* [3]. In addition, the potential effects of histone acetylation and microRNAs are yet to be fully understood in this context. A more rational explanation for the observed discrepancy may simply be that the PCR-RFLP method only accounts for alleles that are known to be frequently present, and hence omits important rare alleles that may be phenotypically relevant [26,49].

Conclusion

The utility of a *CYP2C19* AS system has been demonstrated in this study. Our findings highlight the potential value of applying a predictive tool of this nature in the clinical setting, and with partial modification, a good correlation between *CYP2C19* phenotype prediction and omeprazole metabolism was achieved.

Substrate- and population-specific modification of this predictive tool should thus be a consideration to improve its accuracy in future studies.

Future perspective

A comprehensive genotyping approach is essential for predicting CYP2C19 phenotype, which will require substrate specific AS prediction models in order to accurately tailor dosage regimens. Advances in genomic technologies, particularly the broadening scope of next generation sequencing applications, are expected to have a significant impact on the field of pharmacogenetics and its mainstream application in the clinical setting.

Acknowledgements

The authors thank the South African volunteers who kindly participated in this study as well as Murray Logan and Dr Chrisna Durandt for assisting with the sampling.

Financial & competing interests disclosure

Funding was provided by the Departments of Pharmacology and Immunology, University of Pretoria; the National

Research Foundation of South Africa (NRF) grant numbers FA2006032700005 and TK2006051500005; the National Health Laboratory Services of South Africa (NHLS); the South African Medical Research Council (SAMRC) Extramural Unit for Inflammation and Immunity; and Ampath Laboratories, South Africa. 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.

Ethical conduct of research

The authors state that they have obtained approval by the Human Research Ethics Committee, Faculty of Health Science at the University of Pretoria (S132/2009), and have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

Executive summary

Background

- South Africa is a developing country that will benefit from the clinical application of pharmacogenetics strategies.
- Few studies have investigated genotype and phenotype of the polymorphic and clinically relevant CYP2C19 enzyme in South African populations.

Methodology

- After genotyping 100 volunteers for CYP2C19*2, *3, *9, *15, *17, *27 and *28, phenotype was predicted by means of an activity score (AS) system.
- True phenotype was measured from plasma concentrations of omeprazole and its metabolite, 5'-hydroxyomeprazole, which were quantified simultaneously using semiautomated online solid phase extraction, coupled to tandem mass spectrometry.

Genotype–phenotype correlation of CYP2C19

- Discrepancies between the predicted and measured phenotype were reported, particularly with respect to the increased metabolism allele, CYP2C19*17, and the reduced activity allele, CYP2C19*27 – neither of which influenced omeprazole metabolism.
- Inclusion of the reduced activity alleles, CYP2C19*9 and *15, resulted in an improved phenotypic prediction in this South African cohort.
- An adapted and improved AS system is proposed, which could be further refined by incorporating a more comprehensive CYP2C19 genotyping approach such as NGS.

Concluding remarks

- Substrate-adjusted AS systems are useful and accurate strategies that can be employed for pharmacogenetic screening purposes.

References

- 1 Allison M. Is personalized medicine finally arriving? *Nat. Biotechnol.* 26(5), 509–517 (2008).
- 2 Wester K, Jonsson AK, Spigset O, Druid H, Hagg S. Incidence of fatal adverse drug reactions: a population based study. *Br. J. Clin. Pharmacol.* 65(4), 573–579 (2008).
- 3 Ingelman-Sundberg M, Sim SC, Gomez A, Rodriguez-Antona C. Influence of cytochrome P450 polymorphisms on drug therapies: pharmacogenetic, pharmacoeconomic and clinical aspects. *Pharmacol. Ther.* 116(3), 496–526 (2007).
- 4 Mehta U, Durrheim DN, Blockman M, Kredt T, Gounden R, Barnes KI. Adverse drug reactions in adult medical inpatients in a South African hospital serving a community

- with a high HIV/AIDS prevalence: prospective observational study. *Br. J. Clin. Pharmacol.* 65(3), 396–406 (2008).
- 5 Ingelman-Sundberg M. Pharmacogenetics of cytochrome P450 and its applications in drug therapy: the past, present and future. *Trends Pharmacol. Sci.* 25(4), 193–200 (2004).
 - 6 Brandt JT, Close SL, Iturria SJ *et al.* Common polymorphisms of CYP2C19 and CYP2C9 affect the pharmacokinetic and pharmacodynamic response to clopidogrel but not prasugrel. *J. Thromb. Haemost.* 5(12), 2429–2436 (2007).
 - 7 Gardiner SJ, Begg EJ. Pharmacogenetics, drug-metabolizing enzymes, and clinical practice. *Pharmacol. Rev.* 58(3), 521–590 (2006).
 - 8 Herrlin K, Yasui-Furukori N, Tybring G, Widen J, Gustafsson LL, Bertilsson L. Metabolism of citalopram enantiomers in CYP2C19/CYP2D6 phenotyped panels of healthy Swedes. *Br. J. Clin. Pharmacol.* 56(4), 415–421 (2003).
 - 9 Hoskins JM, Shenfield GM, Gross AS. Relationship between proguanil metabolic ratio and CYP2C19 genotype in a Caucasian population. *Br. J. Clin. Pharmacol.* 46(5), 499–504 (1998).
 - 10 Andersson T, Regardh CG, Lou YC, Zhang Y, Dahl ML, Bertilsson L. Polymorphic hydroxylation of S-mephenytoin and omeprazole metabolism in Caucasian and Chinese subjects. *Pharmacogenetics* 2(1), 25–31 (1992).
 - 11 Mega JL, Simon T, Collet JP *et al.* Reduced-function CYP2C19 genotype and risk of adverse clinical outcomes among patients treated with clopidogrel predominantly for PCI: a meta-analysis. *JAMA* 304(16), 1821–1830 (2010).
 - 12 Myburgh R, Hochfeld WE, Dodgen TM, Ker J, Pepper MS. Cardiovascular pharmacogenetics. *Pharmacol. Ther.* 133(3), 280–290 (2012).
 - 13 The Human Cytochrome P450 (CYP) Allele Nomenclature Page. CYP2C19 allele nomenclature. www.cypalleles.ki.se/cyp2c19.htm
 - 14 Sistonen J, Fuselli S, Palo JU, Chauhan N, Padh H, Sajantila A. Pharmacogenetic variation at CYP2C9, CYP2C19, and CYP2D6 at global and microgeographic scales. *Pharmacogenet. Genomics* 19(2), 170–179 (2009).
 - 15 Sim SC, Risinger C, Dahl ML *et al.* A common novel CYP2C19 gene variant causes ultrarapid drug metabolism relevant for the drug response to proton pump inhibitors and antidepressants. *Clin. Pharmacol. Ther.* 79(1), 103–113 (2006).
 - 16 Kurzawski M, Gawronska-Szklarz B, Wrzesniewska J, Siuda A, Starzynska T, Drozdziak M. Effect of CYP2C19*17 gene variant on *Helicobacter pylori* eradication in peptic ulcer patients. *Eur. J. Clin. Pharmacol.* 62(10), 877–880 (2006).
 - 17 Rudberg I, Mohebi B, Hermann M, Refsum H, Molden E. Impact of the ultrarapid CYP2C19*17 allele on serum concentration of escitalopram in psychiatric patients. *Clin. Pharmacol. Ther.* 83(2), 322–327 (2008).
 - 18 Kearns GL, Leeder JS, Gaedigk A. Impact of the CYP2C19*17 allele on the pharmacokinetics of omeprazole and pantoprazole in children: evidence for a differential effect. *Drug Metabol. Dispos.* 38(6), 894–897 (2010).
 - 19 Ragia G, Arvanitidis KI, Tavridou A, Manolopoulos VG. Need for reassessment of reported CYP2C19 allele frequencies in various populations in view of CYP2C19*17 discovery: the case of Greece. *Pharmacogenomics* 10(1), 43–49 (2009).
 - 20 Drögemöller BI, Wright GE, Niehaus DJ *et al.* Characterization of the genetic profile of CYP2C19 in two South African populations. *Pharmacogenomics* 11(8), 1095–1103 (2010).
 - 21 Blaisdell J, Mohrenweiser H, Jackson J *et al.* Identification and functional characterization of new potentially defective alleles of human CYP2C19. *Pharmacogenetics* 12(9), 703–711 (2002).
 - 22 Alessandrini M, Asfaha S, Dodgen TM, Warnich L, Pepper MS. Cytochrome P450 pharmacogenetics in African populations. *Drug Metabol. Rev.* 45(2), 253–275 (2013).
 - 23 Warnich L, I Drogemoller B, S Pepper M, Dandara C, EB Wright G. Pharmacogenomic research in South Africa: lessons learned and future opportunities in the rainbow nation. *Curr. Pharmacogenomics Pers. Med.* 9(3), 191–207 (2011).
 - 24 Statistics South Africa. www.statssa.gov.za/Publications/P03014/P030142011.pdf
 - 25 Dandara C, Masimirembwa CM, Magimba A *et al.* Genetic polymorphism of CYP2D6 and CYP2C19 in east- and southern African populations including psychiatric patients. *Eur. J. Clin. Pharmacol.* 57(1), 11–17 (2001).
 - 26 Matimba A, Del-Favero J, Van Broeckhoven C, Masimirembwa C. Novel variants of major drug-metabolising enzyme genes in diverse African populations and their predicted functional effects. *Hum. Genomics* 3(2), 169–190 (2009).
 - 27 Ikediobi O, Aouizerat B, Xiao Y, Gandhi M, Gebhardt S, Warnich L. Analysis of pharmacogenetic traits in two distinct South African populations. *Hum. Genomics* 5(4), 265–282 (2011).
 - 28 de Wit E, Delpont W, Rugamika CE *et al.* Genome-wide analysis of the structure of the South African coloured population in the western Cape. *Hum. Genet.* 128(2), 145–153 (2010).
 - 29 Patterson N, Petersen DC, van der Ross RE *et al.* Genetic structure of a unique admixed population: implications for medical research. *Hum. Mol. Genet.* 19(3), 411–419 (2010).
 - 30 The Human Cytochrome P450 (CYP) Allele Nomenclature Database. www.cypalleles.ki.se
 - 31 Gaedigk A, Simon SD, Pearce RE, Bradford LD, Kennedy MJ, Leeder JS. The CYP2D6 activity score: translating genotype information into a qualitative measure of phenotype. *Clin. Pharmacol. Ther.* 83(2), 234–242 (2008).
 - 32 Dodgen TM, Hochfeld WE, Fickl H *et al.* Introduction of the AmpliChip CYP450 test to a South African cohort: a platform comparative prospective cohort study. *BMC Med. Genet.* 14(1), 20 (2013).
 - 33 Dodgen TM, Cromarty AD, Pepper MS. Quantitative plasma analysis using automated online solid-phase extraction with column switching LC-MS/MS for

- characterising cytochrome P450 2D6 and 2C19 metabolism. *J. Sep. Sci.* 34(10), 1102–1110 (2011).
- 34 Tamminga WJ, Wemer J, Oosterhuis B *et al.* An optimized methodology for combined phenotyping and genotyping on CYP2D6 and CYP2C19. *Eur. J. Clin. Pharmacol.* 57(2), 143–146 (2001).
- 35 Isaza C, Henao J, Martínez JH, Arias JC, Beltrán L. Phenotype–genotype analysis of CYP2C19 in Colombian mestizo individuals. *BMC Clin. Pharmacol.* 7, 6 (2007).
- 36 Howden CW, Meredith PA, Forrest JAH, Reid JL. Oral pharmacokinetics of omeprazole. *Eur. J. Clin. Pharmacol.* (5), 641–643 (1984).
- 37 Calabresi L, Pazzucconi F, Ferrara S, Di Paolo A, Tacca MD, Sirtori C. Pharmacokinetic interactions between omeprazole/pantoprazole and clarithromycin in health volunteers. *Pharmacol. Res.* 49(5), 493–499 (2004).
- 38 Baldwin RM, Ohlsson S, Pedersen RS *et al.* Increased omeprazole metabolism in carriers of the *CYP2C19*17* allele; a pharmacokinetic study in healthy volunteers. *Br. J. Clin. Pharmacol.* 65(5), 767–774 (2008).
- 39 Thomson AB, Kirdeikis P, Lastiwka R, Rohss K, Sinclair P, Olofsson B. Pharmacokinetics and pharmacodynamics during treatment with the omeprazole 20 mg enteric-coated tablet and 20 mg capsule in asymptomatic duodenal ulcer patients. *Can. J. Gastroenterol.* 11(8), 657–660 (1997).
- 40 Sharma VK, Peyton B, Spears T, Raufman JP, Howden CW. Oral pharmacokinetics of omeprazole and lansoprazole after single and repeated doses as intact capsules or as suspensions in sodium bicarbonate. *Alimentary Pharmacol. Ther.* 14(7), 887–892 (2000).
- 41 Ramsjö M, Aklillu E, Bohman L, Ingelman-Sundberg M, Roh HK, Bertilsson L. CYP2C19 activity comparison between Swedes and Koreans: effect of genotype, sex, oral contraceptive use, and smoking. *Eur. J. Clin. Pharmacol.* 66(9), 871–877 (2010).
- 42 Kortunay S, Basci NE, Bozkurt A, Isimer A, Sayal A, Kayaalp SO. The hydroxylation of omeprazole correlates with S-mephenytoin and proguanil metabolism. *Eur. J. Clin. Pharmacol.* 53, 3–4, 261–264 (1997).
- 43 de Leon J, Susce MT, Johnson M *et al.* DNA microarray technology in the clinical environment: the AmpliChip CYP450 test for CYP2D6 and CYP2C19 genotyping. *CNS Spectrums* 14(1), 19–34 (2009).
- 44 Drögemöller DI, Hillermann-Rebello R, Niehaus D, Warnich L. *Detection of sequence diversity in the CYP2C19 gene of Xhosa South African individuals: an analytical and comparative study including in silico and functional analysis of the 5' flanking region [Master's thesis]*. University of Stellenbosch, Stellenbosch, South Africa (2010).
- 45 Mwinyi J, Hofmann Y, Pedersen RS *et al.* The transcription factor GATA-4 regulates cytochrome *P450C19* gene expression. *Life Sci.* 86, 19–20, 699–706 (2010).
- 46 Poolsup N, Li Wan Po A, Knight TL. Pharmacogenetics and psychopharmacotherapy. *J. Clin. Pharm. Ther.* 25(3), 197–220 (2000).
- 47 O'Neil WM, Gilfix BM, Markoglou N, Di Girolamo A, Tsoukas CM, Wainer IW. Genotype and phenotype of cytochrome P450 2D6 in human immunodeficiency virus-positive patients and patients with acquired immunodeficiency syndrome. *Eur. J. Clin. Pharmacol.* 56(3), 231–240 (2000).
- 48 Jones AE, Brown KC, Werner RE *et al.* Variability in drug metabolizing enzyme activity in HIV-infected patients. *Eur. J. Clin. Pharmacol.* 66(5), 475–485 (2010).
- 49 Wright GE, Niehaus DJ, Drogemoller BI, Koen L, Gaedigk A, Warnich L. Elucidation of *CYP2D6* genetic diversity in a unique African population: implications for the future application of pharmacogenetics in the Xhosa population. *Ann. Hum. Genet.* 74(4), 340–350 (2010).