**Improved prediction of endoxifen metabolism by CYP2D6 genotype in breast cancer patients treated with tamoxifen**

Short running title: Endoxifen prediction by CYP2D6 genotype

Werner Schroth, PhD 1,2, Stefan Winter, PhD 1,2, Thomas Muerdter, PhD 1,2, Diana Eccles, MD 3, Bryony Eccles, MD 3, Balram Chowbay, MD 4 , Arafat Tfayli, MD 5, Nathalie Khoueiry-Zgheib, MD6, Michel Eichelbaum, MD 1,2, Matthias Schwab, MD 1,2,7,8, Hiltrud Brauch, PhD 1,2,8

1 Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, Germany

2 University of Tuebingen, Tuebingen, Germany

3 Cancer Sciences Academic Unit and University of Southampton Clinical Trials Unit, Faculty of Medicine, University of Southampton

4 Clinical Pharmacology Laboratory, Division of Medical Sciences, Humphrey Oei Institute of Cancer Research, National Cancer Centre, Singapore

5 Hematology-Oncology Division, Department of Internal Medicine, Faculty of Medicine, American University of Beirut, Beirut, Lebanon

6 Department of Pharmacology and Toxicology, Faculty of Medicine, American University of Beirut, Beirut, Lebanon

7 Department of Clinical Pharmacology, Institute of Experimental and Clinical Pharmacology and Toxicology, University Hospital Tuebingen, Tuebingen, Germany

8 German Cancer Consortium (DKTK) and German Cancer Research Center (DKFZ), Heidelberg, Germany

Correspondence to:

Werner Schroth

Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart

Auerbachstr. 112, 70376 Stuttgart, Germany

Tel +49 711 8101 3754; Fax +49 711 859295

email: werner.schroth@ikp-stuttgart.de

**Abstract**

**Purpose:** Prediction of impaired tamoxifen (TAM) to endoxifen metabolism may be relevant to improve breast cancer treatment, e.g. via TAM dose increase. We modeled plasma endoxifen predictability depending on CYP2D6 metabolizer phenotype definition derived from genotypes.

**Methods:** CYP2D6 diplotype and metabolite plasma concentrations were assessed in 936 pre- and postmenopausal estrogen receptor (ER)-positive, TAM treated early breast cancer patients of Caucasian (N=676), Middle-Eastern Arab (N=77) and Asian (N=153) origin. Using linear modeling and five different CYP2D6 phenotype assignments, robust coefficients of determination (R2) for endoxifen (E) or metabolic ratio endoxifen/desmethyl-TAM (E/DMT) were calculated. Allele activity scores (AS) were modified with respect to the \*10 allele.

**Results:** CYP2D6 diplotypes were strongly associated with E and E/DMT independent of age (P < 10-15). Across all ethnicities, 68-82% inter-patient variability of E/DMT was explained by CYP2D6 diplotype, while plasma endoxifen was predictable by 39-58%. The commonly used codeine specific phenotype classification showed poor prediction (<20%) for both endpoints particularly in Asians (*P* < 10-9). A reduced \*10 activity slightly improved the explanatory value of metabolizer phenotype in Caucasians (*P* < 0.002).

**Conclusion:** The CYP2D6 predictive power for active drug level assessment is maximized when TAM-specific phenotype assignments and plasma E/DMT ratio are considered.

**Key words:** endoxifen, CYP2D6 polymorphism, metabolizer phenotype, tamoxifen, breast cancer

**Introduction**

Tamoxifen (TAM) is a widely prescribed antiestrogen for the control of estrogen receptor (ER)-positive breast cancer, yet its efficacy is reduced due to the development of endocrine resistance and intrinsic patient characteristics that prevent drug response. The latter has been partially attributed to a lack of TAM bioactivation towards its active metabolite, endoxifen. Pharmacological and pharmacogenetic evidence strongly support that *in vivo* endoxifen formation is mainly mediated from the main primary metabolite N-desmethyl-TAM by the cytochrome P450 2D6 (CYP2D6) enzyme*.*1,2 As distinct genetically determined functional *CYP2D6* variants are present in the general population, inter-patient variability of plasma endoxifen is expected to be predictable, at least in part.3

The *CYP2D6* polymorphism with more than 100 known alleles contributes to inter-individual differences in enzyme activities and plasma exposure of metabolized drugs and can be grouped into four CYP2D6 metabolizer phenotypes such as ultra-rapid (UM), extensive (EM), intermediate (IM) and poor (PM) metabolizers. Traditionally, these have been defined using probe substrates, however, due to probe drug differences to derive phenotypes, CYP2D6 genotyping has emerged as the method of choice to predict enzyme activity.4 Activity scores (AS) of 0, 0.5, and 1 for null (PM), reduced-function (IM), and fully-functional (EM) alleles, respectively, have been used to infer metabolizer phenotypes from diplotypes.5 Prediction of an impaired TAM metabolizer phenotype (IM, PM) with low endoxifen formation capacity is potentially important for personalized treatment decisions in breast cancer such as increasing the therapeutic TAM dose or replacing TAM with an aromatase inhibitor. Since routine therapeutic endoxifen monitoring appears not feasible in standard clinical practice, genotyping has been put forward and tested in several studies as a prospective tool to select patients for TAM dose escalation or to establish its predictive value.6-10 In the absence of standardized guidelines,4 studies used different phenotype assignments including that based on CYP2D6-dependent codeine metabolism.11 As a consequence, inaccurate endoxifen predictability from codeine-specific CYP2D6 phenotype assessment in a recent study10 led to recommendations against the use of CYP2D6 genotype to guide clinical decisions.12

While the combination of multi-locus genotypes into diplotypes based on the AS system5 appears straightforward, their attribution to a specific metabolizer phenotype has been inconsistent preventing meaningful clinical conclusions. Here, we used different metabolizer phenotype definitions to test the power of CYP2D6 diplotype and phenotype-based prediction of impaired endoxifen metabolism with the goal to provide a robust algorithm towards the standardization of CYP2D6 in personalized endocrine treatment.

**PATIENTS and Methods**

**Patients**

The genotype data and available TAM and TAM metabolite concentrations of 936 prospectively recruited, ER-positive breast cancer patients that had received adjuvant TAM treatment (20 mg/d) for at least 6 months and who had TAM plasma concentrations above 150 nM were included in this bioinformatic modeling. Patients include postmenopausal Caucasian women (N=365) derived from a German observational trial of outcome predictors in adjuvant endocrine treatment (DRKS 00000605) that were extended from Mürdter et al.13, and three ethnic groups of premenopausal Caucasian, Asian, and Middle-Eastern Arab women (N=571), previously described in Saladores et al.14

**Genotyping, phenotype definition, and plasma metabolite measurement**

CYP2D6 diplotypes were assessed by alleles predictive of metabolizer status PM (\*3, \*4, \*5, \*6, \*7), IM (\*9, \*10, \*41), EM (absence of variant alleles, or \*1, \*2, \*35) and ultra-rapid, UM (duplicated EM allele) with activity scores (AS) 0, 0.5, 1, and 2, respectively per allele5. CYP2D6 phenotypes were defined by the various diplotype groupings including a modified \*10 AS count (Table 1): the codeine metabolism-based grouping defined by the Clinical Pharmacogenetics Implementation Consortium11 (Codeine) and four possible TAM specific phenotype assignments assuming either equal IM allele activity of 0.5 (TAM16,9, TAM3) or applying a downgrade of \*10 activity from 0.5 to 0.25 as previously proposed4 (TAM4) with further downgrading \*10 containing IM genotypes into a new *slow metabolizer* (SM) group (TAM5).

Data of TAM and its metabolites N-desmethyl-TAM (DMT) and (Z)-endoxifen were taken13,14 with extended numbers of postmenopausal patients plasma that were measured by liquid chromatography tandem mass spectrometry as described.13 To account for alternative and upstream pathways of endoxifen formation from (Z)-4-hydroxy-TAM and N-desmethyl-TAM, CYP2C9\*2 and \*3 alleles exerting decreased enzyme function, as well as CYP3A5\*3 encoding a non-functional protein were genotyped as described.13,14

**Linear modeling and statistical analysis**

The effect of CYP2D6 diplotypes and phenotype classifications on square-root transformed endoxifen concentration or log-transformed metabolic ratio endoxifen/desmethyl-TAM (E/DMT) was assessed by linear modeling, incorporating CYP2C9\*2, \*3and CYP3A5\*3 genotypes as covariates of TAM metabolism. Robust coefficients of determination (R2) and 95% confidence intervals (CI) based on 10.000 bootstrap replicates of the original data were calculated using R package lmrob for the following subgroups: Caucasians premenopausal (N=311), Caucasians postmenopausal (N=365), Middle-Eastern Arabs premenopausal (N=77), and Asians premenopausal (N=153). Analysis of deviance was applied to test between linear models including one and two CYP2D6 phenotype assignments as independent variables, respectively.

**Results**

**Patient specific CYP2D6 activity**

There was a strong gene-dose effect for an association between CYP2D6 diplotype/activity score and endoxifen concentrations or metabolic ratio E/DMT across all patients (Fig 1A; P<10-15). The distribution of E and E/DMT depending on diplotype did not differ between subgroups of women younger or older than 50 years, indicating an identical TAM metabolism irrespective of age or menopause. While the median CYP2D6 activity (based on E and E/DMT ratio) increased with increasing AS, the range of phenotypic activity was larger for patients with AS ≥ 1 compared to a smaller variability in patients with severely impaired activity (AS ≤ 0.5) in which PM/PM patients were predictable with up to 9x% accuracy for low endoxifen based on a proposed threshold of 5.9 ng/mL.15 By using the E/DMT ratio the number of outliers and therefore statistical spread were reduced compared to absolute endoxifen concentrations, yet a portion of outliers that were not correlated with known functional CYP2D6 variants (grey area Fig 1A, right) imply additional factors conferring increased or reduced endoxifen plasma exposure reminiscent of an UM or PM metabolizer phenotype, respectively.

**CYP2D6 phenotype modeling**

Linear modeling across all three ethnic subgroups revealed that CYP2D6 diplotype showed the highest degree of determination for both metabolite endpoints as compared to five evaluated phenotype classifications inferred from diplotype. The contribution of CYP2D6 to the inter-individual variability was highest for diplotypes as a predictor of E/DMT with a maximum of 68% (premenopausal Caucasians) to 82% (Asians) of the variability explained. Similarly, absolute endoxifen concentrations were most optimally predicted by diplotype yet to a lesser extent of 39 to 58% (Figure 1B right and left, respectively). Of the five tested phenotype groupings derived from respective diplotypes (Table 1), TAM5 was superior in its discriminatory power for both E (34% to 52%) and E/DMT (62% to 65%). Of note, the TAM5 phenotype was adapted by a downgrade of \*10 AS from 0.5 to 0.25 and introduction of a non-classical *slow* metabolizer phenotype (SM) with activity scores halfway between IM and PM (Table 1). When compared to TAM3 as the best explanatory phenotype model without modification of \*10 activity, TAM5 was not significantly better in Asians and Middle-Eastern Arabs, however for the prediction of E/DMT, TAM5 was superior in premenopausal (*P* < 0,0001) and postmenopausal (*P* < 0.002) Caucasians. When compared to the commonly used codeine specific phenotype classification (Codeine) the latter showed poorest prediction of less than 20% for both E and E/DMT particularly in Asians, which significantly differed from TAM3 and TAM5 (*P* < 10-9). The two remaining phenotype groupings TAM1 and TAM4 showed intermediary explanatory power for both plasma endpoints, independent of whether \*10 activity was downgraded (TAM4) or not (TAM1).

**Discussion**

We re-evaluted a comprehensive data set of CYP2D6 genotypes and TAM metabolite concentrations of breast cancer patients treated with adjuvant TAM to assess the prediction of impaired TAM metabolism by CYP2D6. We applied the power of diplotype-based assignments5 to further refine the discriminatory value of metabolizer phenotype as the most intuitive concept for CYP2D6 polymorphism. To shed light on current controversies on the utility of CYP2D6 for TAM efficacy prediction16-18 standardized genotype-phenotype relationships for the validation of an association between CYP2D6 and impaired TAM metabolism are mandatory.

Currently, the extent to which CYP2D6 determines the up to 20-30 fold13 inter-patient variability of plasma endoxifen is poorly characterized. Compared to drug level monitoring of endoxifen, which requires a sophisticated analytical infrastructure, CYP2D6 genotyping requires only a standard molecular biology laboratory and is more likely implemented for routine clinical practice, if predictive. However, the lack of guidelines for the assessment of genotype-phenotype relationships led to the use of a CYP2D6 phenotype classification scheme for impaired Tam metabolism10 that was previously recommended for codeine metabolism.11 Yet, it has become increasingly clear that CYP2D6 variants may exert substrate-dependent effects,5,19,20 and therefore, diplotype specific phenotype data obtained with codeine cannot be extrapolated to other CYP2D6 substrates such as TAM. This functional discrepancy may have important clinical implications as others, based on the CYP2D6 mediated codeine metabolism inappropriately concluded that CYP2D6 has no value for the prediction of TAM metabolism.12 Our re-evaluation of existing pharmacogenetic and pharmacokinetic data challenges these findings.

We showed that plasma endoxifen prediction highly depends on the phenotypical grouping of CYP2D6 variant alleles and used plasma endpoint. CYP2D6 diplotypes predicted a maximum of endoxifen variability independent of ethnicity. This was less strong when diplotypes were collapsed into fewer class levels (phenotypes), indicating that the allele-dose dependent effects of variant alleles are maximally exploited by a semiquantitative score. Importantly, the codeine-specific phenotype grouping11 poorly predicted CYP2D6-based endoxifen formation. In particular, it did not perform in Asians (predictability <20%) most likely due to a misclassification of abundant IM/IM (\*10) diplotypes as EM. From this it follows that CYP2D6 variants act differentially on TAM and Codeine substrates, underscoring the need for substrate-specific CYP2D6 phenotype assessments.4 Within this context, the suggested extra deleterious effect on enzyme function of \*10 compared to other IM alleles4 was addressed by downgrading its AS from 0.5 to 0.25 and by placing \*10 homozygous patients together with IM/PM diplotypes into a new phenotype category of *slow* metabolizers. A moderate increase of explanatory power (TAM5) by approximately 5% compared to phenotypes that did not incorporate a \*10 downgrade (TAM3) supports the notion of an increased deleterious effect of \*10 on reduced TAM metabolism.

Notably, our study showed that the plasma endpoint closely linked to CYP2D6 activity is active metabolite-to-precursor ratio, as three quarters (68-82%) of the variability of E/DMT was explained by CYP2D6 diplotype while plasma endoxifen variability was predictable to a lesser extent (39-58%). Because the E/DMT ratio is less susceptible to metabolite fluctuations caused by non-compliance and alternate clearance pathways independent of CYP2D6, our findings indicate the necessity to consider the E/DMT endpoint when linking CYP2D6 with TAM metabolism and possibly outcome. Importantly, CYP2D6 inhibitor use in our study was less than 1% in postmenopausal Caucasians, Asians, and Arabs (unknown in premenopausal Caucasians), and other factors influencing endoxifen formation were accounted for by adjusting for CYP2C9\*2, \*3 and CYP3A5\*3 variants. Therefore, the E/DMT – based translations of diplotypes into metabolizer phenotypes TAM5 and TAM3 with or without downgrading of \*10 activity, respectively, likely provide an accurate estimate of CYP2D6-dependent endoxifen formation capacity and is superior to previous assignments such as TAM16,9 and the codeine score10. Of note, PM alleles have highest discriminatory power with up to 9x% of the E/DMT variability explained by CYP2D6, while suspect outliers not linked with CYP2D6 are more frequently seen with higher AS supporting the notion of a higher likelihood of EM and UM misclassifications. Notably, since the active drug rather than the metabolic ratio is effective in inhibiting cancer cell growth, it will be important to elucidate additional factors that influence endoxifen variability and to evaluate the clinical relevance of either plasma biomarker.

In summary, we provided an improved algorithm to predict a large quantity of variable TAM metabolism by CYP2D6 underscoring its essential role in drug bioactivation towards endoxifen independent of age and ethnicity. Model predictions must be adjusted for known covariates, should utilize phenotype classifications that quantitatively account for allele-dosage effects, and should consider E/DMT ratio as the plasma endpoint most directly linked to CYP2D6. Our improved TAM-specific CYP2D6 activity assignments may shed new light on linking the CYP2D6 genotype with endoxifen formation and TAM outcome prediction aiding the selection of patients for TAM dose increase or aromatase inhibitor treatment.

**Acknowledgments**

This work was supported by the Robert Bosch Foundation, Stuttgart, Deutsche Forschungsgemeinschaft (DFG, SCHR 1323/2-1 and MU 1727/2-1), IZEPHA (Grant 2014-07), Germany, The German Cancer Consortium (DKTK), and Bundesministerium für Bildung und Forschung (BMBF, FKZ 01EK1509A, Germany.

**References**

1. Stearns V. Active Tamoxifen Metabolite Plasma Concentrations After Coadministration of Tamoxifen and the Selective Serotonin Reuptake Inhibitor Paroxetine. CancerSpectrum Knowledge Environment. 2003;95(23):1758-1764. doi:10.1093/jnci/djg108.

2. Desta Z. Comprehensive Evaluation of Tamoxifen Sequential Biotransformation by the Human Cytochrome P450 System in Vitro: Prominent Roles for CYP3A and CYP2D6. Journal of Pharmacology and Experimental Therapeutics. 2004;310(3):1062-1075. doi:10.1124/jpet.104.065607.

3. Schultink AHMV, Zwart W, Linn SC, Beijnen JH, Huitema ADR. Effects of Pharmacogenetics on the Pharmacokinetics and Pharmacodynamics of Tamoxifen. Clinical pharmacokinetics. April 2015:1-14. doi:10.1007/s40262-015-0273-3.

4. Hicks JK, Swen JJ, Gaedigk A. Challenges in CYP2D6 phenotype assignment from genotype data: a critical assessment and call for standardization. Curr Drug Metab. 2014;15(2):218-232.

5. 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. 2008;83(2):234-242. doi:10.1038/sj.clpt.6100406.

6. Irvin WJ, Walko CM, Weck KE, et al. Genotype-Guided Tamoxifen Dosing Increases Active Metabolite Exposure in Women With Reduced CYP2D6 Metabolism: A Multicenter Study. 2011;29(24):3232-3239. doi:10.1200/JCO.2010.31.4427.

7. Kiyotani K, Mushiroda T, Imamura CK, et al. Dose-adjustment study of tamoxifen based on CYP2D6 genotypes in Japanese breast cancer patients. Breast Cancer Res Treat. 2012;131(1):137-145. doi:10.1007/s10549-011-1777-7.

8. Dezentjé VO, Opdam FL, Gelderblom H, et al. CYP2D6 genotype- and endoxifen-guided tamoxifen dose escalation increases endoxifen serum concentrations without increasing side effects. Breast Cancer Res Treat. September 2015:1-8. doi:10.1007/s10549-015-3562-5.

9. Hertz DL, Snavely AC, McLeod HL, et al. In vivoassessment of the metabolic activity of CYP2D6 diplotypes and alleles. Br J Clin Pharmacol. 2015;80(5):1122-1130. doi:10.1111/bcp.12665.

10. Fox P, Balleine RL, Lee C, et al. Dose Escalation of Tamoxifen in Patients with Low Endoxifen Level: Evidence for Therapeutic Drug Monitoring--The TADE Study. Clin Cancer Res. 2016;22(13):3164-3171. doi:10.1158/1078-0432.CCR-15-1470.

11. Crews KR, Gaedigk A, Dunnenberger HM, et al. Clinical Pharmacogenetics Implementation Consortium guidelines for cytochrome P450 2D6 genotype and codeine therapy: 2014 update. Clin Pharmacol Ther. 2014;95(4):376-382. doi:10.1038/clpt.2013.254.

12. Hertz DL, Rae JM. Individualized Tamoxifen Dose Escalation: Confirmation of Feasibility, Question of Utility. Clin Cancer Res. 2016;22(13):3121-3123. doi:10.1158/1078-0432.CCR-16-0370.

13. Mürdter TE, Schroth W, Bacchus-Gerybadze L, et al. Activity levels of tamoxifen metabolites at the estrogen receptor and the impact of genetic polymorphisms of phase I and II enzymes on their concentration levels in plasma. Clin Pharmacol Ther. 2011;89(5):708-717. doi:10.1038/clpt.2011.27.

14. Saladores P, Mürdter T, Eccles D, et al. Tamoxifen metabolism predicts drug concentrations and outcome in premenopausal patients with early breast cancer. TPJ. 2015;15(1):84-94. doi:10.1038/tpj.2014.34.

15. Madlensky L, Natarajan L, Tchu S, et al. Tamoxifen Metabolite Concentrations, CYP2D6 Genotype, and Breast Cancer Outcomes. Clin Pharmacol Ther. 2011;89(5):718-725. doi:doi:10.1038/clpt.2011.32.

16. Ratain MJ, Nakamura Y, Cox NJ. CYP2D6 Genotype and Tamoxifen Activity: Understanding Interstudy Variability in Methodological Quality. Clin Pharmacol Ther. 2013;94(2):185-187. doi:10.1038/clpt.2013.66.

17. Brauch H, Schwab M. Prediction of tamoxifen outcome by genetic variation of CYP2D6 in post-menopausal women with early breast cancer. Br J Clin Pharmacol. 2014;77(4):695-703. doi:10.1111/bcp.12229.

18. Hertz DL, Rae JM. One step at a time: CYP2D6guided tamoxifen treatment awaits convincing evidence of clinical validity. Pharmacogenomics. 2016;17(8):823-826. doi:10.2217/pgs-2016-0059.

19. Bogni A, Monshouwer M, Moscone A, et al. Substrate specific metabolism by polymorphic cytochrome P450 2D6 alleles. Toxicol In Vitro. 2005;19(5):621-629. doi:10.1016/j.tiv.2005.04.001.

20. Zhou S-F. Polymorphism of human cytochrome P450 2D6 and its clinical significance: Part I. Clinical pharmacokinetics. 2009;48(11):689-723. doi:10.2165/11318030-000000000-00000.

**Table 1.** CYP2D6 diplotype activity scores (AS), their observed frequencies, and five evaluated phenotypic groupings for the prediction of plasma endoxifen metabolizer status

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Diplotype | AS(a) | N | % | Codeine | TAM1 | TAM3 | TAM4(c) | TAM5(d) |
| EM/UM | 3 | 18 | 2.0 | UM | UM | UM | UM | UM |
| EM/EM | 2 | 300 | 33.4 | EM | EM | EM | EM | EM |
| EM/IM | 1.5 | 168 | 18.7 | EM | IM | EM | EM | EM |
|  *EM/\*10(b)* | *1.25* | *60* | *6.7* | - | - | - | IM | EM |
| EM/PM | 1 | 221 | 24.6 | EM | IM | IM | IM | IM |
| IM/IM | 1 | 68 | 7.6 | EM | IM | IM | IM | IM |
|  *\*10/\*10(b*) | *0.5* | *45* | *5.0* | - | - | - | IM | SM |
| IM/PM | 0.5 | 73 | 8.1 | IM | IM | PM | IM | SM |
|  *PM/\*10(b)* | *0.25* | *19* | *2.1* |  |  |  | PM | SM |
| PM/PM | 0 | 50 | 5.6 | PM | PM | PM | PM | PM |

UM, ultra-rapid; EM, extensive; IM, intermediate; SM, slow; PM, poor metabolizer phenotypes; AS, activity score; a) calculated as sum of allele activities for PM (0), IM (0.5), EM (1), and UM (2); b) for adjusted IM phenotype definitions \*10 AS was reduced from 0.5 to 0.25 in TAM4 and TAM5; c) reduced \*10 activity with diplotype AS of 1.5-2 (EM), 0.5-1.25 (IM), and 0-0.25 (PM); d) reduced \*10 activity and definition of a slow metabolizer (SM) group with diplotype AS of 1.25-2 (EM), 1 (IM), 0.25-0.5 (SM), and 0 (PM)

****

**Figure 1.** Plasma concentrations and explained variability of (Z)-endoxifen and metabolic ratio (Z)-endoxifen/desmethyl-TAM (E/DMT) depending on CYP2D6 **A)** Patient plasma concentrations of (Z)-endoxifen (left) and E/DMT (right) according to CYP2D6 diplotype and age in the total cohort. Concentrations are presented as boxplots with whiskers defined as 1.5 times the inter-quartile range, and extreme values outside the whiskers. Shaded areas (right) point to outliers that are unexplained by CYP2D6. **B)** Prediction of plasma endoxifen (left) and E/DMT (right) according to different CYP2D6 phenotype classifications inferred by diplotype in 3 different ethnicities. Robust coefficients of determination (R2) are indicated by symbols, bars represent 95% confidence intervals (CI) based on 10.000 bootstrap replicates of the original data. Symbols are referring to Caucasians premenopausal (pre, white circle, N=311), Caucasians postmenopausal (post, black circle, N=365), Arabs (triangle, N=77), Asians (diamond, N=153).