Title: Obesity alters endoxifen plasma levels in young breast cancer patients: A pharmacometric simulation approach.

Authors: Anna Mueller-Schoell\textsuperscript{1,2}, Lena Klopp-Schulze\textsuperscript{1}, Werner Schroth\textsuperscript{3,4}, Thomas Mürdter\textsuperscript{3,4}, Robin Michelet\textsuperscript{1}, Hiltrud Brauch\textsuperscript{3,4,17}, Wilhelm Huisingsa\textsuperscript{5}, Markus Joerger\textsuperscript{6}, Patrick Neven\textsuperscript{7}, Stijn L.W. Koolen\textsuperscript{8}, Ron H.J. Mathijsse\textsuperscript{9,10}, Diana Eccles\textsuperscript{9,10}, Sylvia Chen\textsuperscript{11}, Balram Chowbay\textsuperscript{11,12,13}, Arafat Tfayli\textsuperscript{14}, Nathalie K. Zgheib\textsuperscript{15}, Matthias Schwab\textsuperscript{3,16,17}, Charlotte Kloft\textsuperscript{1}

\textsuperscript{1}Department of Clinical Pharmacy and Biochemistry, Institute of Pharmacy, Freie Universitaet Berlin, Berlin, Germany
\textsuperscript{2}Graduate Research Training Program PharMetrX, Germany
\textsuperscript{3}Dr. Margarete Fischer-Bosch Institute of Clinical Pharmacology, Stuttgart, Germany
\textsuperscript{4}University of Tübingen, Tübingen, Germany
\textsuperscript{5}Institute of Mathematics, University of Potsdam, Potsdam, Germany
\textsuperscript{6}Department of Medical Oncology and Haematology, Cantonal Hospital, St. Gallen, Switzerland
\textsuperscript{7}Vesalius Research Center – VIB, University Hospitals Leuven, KU Leuven-University of Leuven, Leuven, Belgium
\textsuperscript{8}Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands
\textsuperscript{9}Cancer Sciences Academic Unit and University of Southampton Clinical Trials Unit, Faculty of Medicine, University of Southampton, UK
\textsuperscript{10}University Hospital Southampton NHS Foundation Trust, Southampton, UK
\textsuperscript{11}University Clinical Pharmacology Laboratory, Division of Medical Sciences, Humphrey Oei, Institute of Cancer Research, National Cancer Centre Singapore, Singapore
\textsuperscript{12}Centre for Clinician Scientist Development, Office of Clinical Sciences, Duke-NUS Graduate Medical School Singapore, Singapore
\textsuperscript{13}SingHealth Clinical Pharmacology Core, Academia, Singapore Health Services, Singapore
\textsuperscript{14}Hematology-Oncology Division, Department of Internal Medicine, Faculty of Medicine, American University of Beirut, Lebanon
\textsuperscript{15}Department of Pharmacology and Toxicology, Faculty of Medicine, American University of Beirut, Lebanon
\textsuperscript{16}Departments of Clinical Pharmacology, Pharmacy and Biochemistry, University Tübingen, 72076 Tübingen, Germany
\textsuperscript{17}German Cancer Consortium (DKTK) and of German Cancer Research Center (DKFZ), Heidelberg, Germany
Corresponding author: 
Charlotte Kloft 
Freie Universitaet Berlin 
Institute of Pharmacy 
Department of Clinical Pharmacy & Biochemistry 
Kelchstr. 31 
12169 Berlin 
Tel.: 030 838 50656 
email: charlotte.kloft@fu-berlin.de

Conflict of interest statement: 
CK and WH report grants from an industry consortium (AbbVie Deutschland GmbH & Co. KG, Boehringer Ingelheim Pharma GmbH & Co. KG, Grünenthal GmbH, F. Hoffmann-La Roche Ltd, Merck KGaA and Sanofi) for the PharMetrX program. CK reports grant for the Innovative Medicines Initiative-Joint Undertaking (‘DDMoRe’). CK and RM report grants from the Federal Ministry of Education and Research within the Joint Programming Initiative on Antimicrobial Resistance Initiative (JPIAMR), all outside the submitted work.

Funding information: 
This work was in parts supported by the Robert Bosch Stiftung, the Bundesministerium für Bildung und Forschung (BMBF; 01EK1509A), and the Deutsche Forschungsgemeinschaft (DFG; MU 1727/2-1 and SCHR 1323/2-1), Germany.

Keywords: tamoxifen, endoxifen, age, body weight, pharmacokinetics, modelling, simulation, individualised dosing
ABSTRACT

Endoxifen is the most important metabolite of the prodrug tamoxifen. High interindividual variability in endoxifen steady-state concentrations (CSS,min ENDX) is observed under tamoxifen standard dosing breast cancer patients that do not reach endoxifen concentrations above a proposed therapeutic threshold of 5.97 ng/mL may be at higher recurrence risk. In this investigation, 10 clinical tamoxifen studies were pooled (n Patients=1388) to investigate influential factors on CSS,min ENDX using nonlinear mixed-effects modelling. Age and body weight were found to significantly impact CSS,min ENDX in addition to CYP2D6 phenotype. Compared to post-menopausal patients, pre-menopausal patients had a 30% higher risk for subtarget CSS,min ENDX at tamoxifen 20 mg per day. In treatment simulations for distinct patient subpopulations, young overweight patients had a 3.1-13.8-fold higher risk for subtarget CSS,min ENDX compared to elderly low-weight patients. Considering ever-rising obesity rates and the clinical importance of tamoxifen for pre-menopausal patients, this subpopulation may benefit most from individualised tamoxifen dosing.

INTRODUCTION

Tamoxifen treatment for 5 to 10 years is widely used in pre-menopausal- and an option in post-menopausal oestrogen receptor positive breast cancer patients. During its use for more than 40 years, a 5-year adjuvant tamoxifen treatment has been proven to effectively reduce breast cancer recurrence by around 30% in the 15 first years after therapy start. Tamoxifen is extensively metabolised and considered to be the pro-drug to its 100-fold more active metabolite endoxifen. Several polymorphic enzymes such as CYP2D6, CYP2C9, CYP2C19, CYP3A5, sulfotransferases (SULTs) and UDP-glucuronosyltransferases (UGTs) are involved in tamoxifen metabolism and consequently large interindividual variability (IIV) in endoxifen minimum concentrations at steady-state (CSS,min ENDX) has been observed under tamoxifen standard dosing (20 mg once daily (Q.D.))7–9. CYP2D6 is especially important for endoxifen formation and patients with impaired or no CYP2D6 activity have shown an increased risk for subtarget CSS,min ENDX8–11. Regarding a putative therapeutic threshold concentration, Madlensky et al. reported that patients with CSS,min ENDX<5.97 ng/mL had a 26% higher breast cancer recurrence rate compared to patients with CSS,min ENDX above this threshold9. This target concentration was later supported by Saladores et al. for pre-menopausal patients8. Other studies failed to find the described relationship between CSS,min ENDX and/or CYP2D6 and treatment outcome12–14, which might in part be due to heterogeneous patient populations, study designs DNA source used for CYP2D6 genotype determination and insufficient power to detect the relationships16,17. Accordingly, the efficacy of breast cancer tamoxifen treatment may be influenced by the target threshold (CSS,min ENDX >5.97 ng/mL), however, non-genetic factors beyond CYP2D6 drug-metabolism that influence the PK of tamoxifen and endoxifen may play a role. Of those, a positive correlation between patient age and tamoxifen concentrations has been described in literature18–20 and was later quantified and found to be clinically relevant in a pharmacokinetic analysis using nonlinear mixed-effects modelling (NLME)21. Furthermore, increasing body weight or BMI have been associated with not only decreased concentrations of tamoxifen and its primarily lipophilic metabolites but also worse clinical outcome. However, the impact of body weight on CSS,min ENDX has never been quantified.
In this work, we applied mathematical modelling and simulations to quantify the influence of age and body weight on $C_{SS,\text{min}}^{ENDX}$ in patients treated with adjuvant tamoxifen and report a patient subpopulation at risk for subtarget endoxifen concentrations.

**METHODS**

**Clinical study database**

A large tamoxifen clinical study dataset was compiled by pooling data from 10 clinical studies. Studies 1-6 (referred to as ‘development dataset’ that previously was pooled at the Freie Universitaet Berlin, Germany) and studies 7-10 (referred to as ‘evaluation dataset’ that previously was pooled at the Dr. Margarete Fischer-Bosch Institute of Clinical Pharmacology in Stuttgart, Germany) are described in detail elsewhere. All studies were conducted in accordance with the ethical standards of the Declaration of Helsinki and had been approved by the respective ethics committees.

The pooled dataset comprised demographic, pharmacokinetic (PK) and pharmacogenetic data and tamoxifen and endoxifen steady-state (SS) plasma concentrations in 1388 female breast cancer patients receiving 20 mg (n=1373) or 40 mg (n=15) tamoxifen once daily (Q.D.) (Table 1). Patients receiving strong CYP2D6 inhibitors or CYP3A4 inducers and patients that had not yet reached SS were excluded from the development dataset (n=16) prior to pooling.

According to the Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for CYP2D6 and Tamoxifen Therapy, patients were assigned CYP2D6 activity scores (AS) based on their CYP2D6 diplotype. Genotype-predicted phenotype (gX) assignment was as follows: (i) AS of 0 refers to poor metabolisers (gPM), (ii) AS of 0.5-1 refers to intermediate metabolisers (gIM) and (iii) AS of ≥1.5 refers to normal metabolisers (gNM, including ultrarapid-metabolisers (AS>2)). For patients with missing genotype information (n=39, 2.81%) the CYP2D6 wildtype (AS=2) as most frequent CYP2D6 AS was imputed.

Menopausal status had not been reported in the development dataset and was imputed based on the age-menopause relationship observed in the evaluation dataset: Age densities for pre- and post-menopausal patients in the evaluation dataset were plotted and the intersection of both densities (52 years) was used as cut-off value. All patients with missing menopausal status <52 years old were classified pre-menopausal while patients ≥52 years old were classified post-menopausal.

The development dataset included Caucasian (n=433) and African (n=2) patients, while the evaluation dataset included pre- and post-menopausal Caucasians (n=681) and pre-menopausal Africans (n=12), Middle-Eastern Arabs (n=77), Asians (n=153) and Indians (n=12). For patients without reported ethnicity (n=14, 1.01%), Caucasian ethnicity, as the most frequent, was imputed.

**Joint parent-metabolite pharmacokinetic model of tamoxifen and endoxifen, and external model evaluation**

The parsimonious joint parent-metabolite NLME pharmacokinetic model of tamoxifen and endoxifen developed using the development dataset was externally evaluated using the evaluation dataset. A one-compartment model - parameterised in terms of relative clearances (CL/F) and relative volumes of distributions (V/F), with first-order absorption with lag time for tamoxifen - was linked to an endoxifen one-compartment model via a linear first-order formation process (CL23/F). Elimination of tamoxifen and of endoxifen (CL20/F and
CL30/F, respectively) were both described as linear first-order processes. Parameter values for endoxifen apparent clearance (CL30/F) and endoxifen apparent volume of distribution (V\text{ENDX}/F) were adopted from a study in which endoxifen had been administered as a single compound\textsuperscript{32}. Interindividual (IIV) variability parameters were estimated for both tamoxifen clearance and endoxifen formation (CL20/F and CL23/F, respectively), whereas interoccasion variability was not considered, as only one PK sample per patient was available in the evaluation dataset. CYP2D6 AS and age as significant covariates on endoxifen formation and tamoxifen clearance, respectively, were implemented as proportional and power functions, respectively.

Based on the final estimates using the development dataset, tamoxifen and endoxifen concentrations were predicted for patients of the evaluation dataset and compared with observed concentrations. Mean absolute prediction errors (MAPE) and mean prediction errors (MPE) were calculated to assess precision and bias, respectively\textsuperscript{33}. Finally, model parameters, except absorption parameters which were fixed to the estimates obtained during model development as the evaluation dataset contained C\text{SS.min} only, were re-estimated using the pooled dataset and compared with previously estimated parameters using the development dataset.

**Extensive covariate analysis and final model development**

Patient characteristics were pre-selected for the extensive covariate analysis based on physiological plausibility, previous literature reports and sufficient information in the pooled clinical study dataset. A relationship between increasing age and decreasing tamoxifen clearance had been reported in the literature\textsuperscript{20} and was supported by our previous analysis using the development dataset\textsuperscript{21}. To evaluate this relationship in the pooled dataset and test for differences between both datasets, C\text{SS.min ENDX} were compared between pre- and post-menopausal patients receiving 20 mg tamoxifen Q.D. in the development (n=435) and evaluation (n=935) dataset, respectively. Based on expected PK differences between ethnicities\textsuperscript{34}, C\text{SS.min ENDX} were additionally compared between pre-menopausal patients of different ethnicities in the evaluation dataset. To evaluate the contribution of varying CYP2D6 phenotype frequencies to the observed differences between ethnicities, C\text{SS.min ENDX} were compared in pre-menopausal patients of different ethnicities stratified for CYP2D6 phenotype.

To assess if differences between patient subpopulations were statistically significant, nonparametric Wilcoxon tests were performed. The extensive covariate analysis was based on the pooled dataset and the original PK base model. Age, CYP2D6 AS together with newly selected covariates body weight as proportional or power function and ethnicity as categorical, function were tested for significance on model parameters endoxifen formation (CL23/F), endoxifen clearance (CL30/F) and tamoxifen clearance (CL20/F) using a stepwise covariate model-building approach\textsuperscript{35}. Significance criteria of 3.84 (α=0.05) and 7.88 (α=0.005) points change in objective function value were applied in the inclusion and exclusion steps, respectively. Finally, goodness-of-fit (GOF) plots were created to assess model performance.

**Treatment simulations for different patient subpopulations**

Applying the updated joint parent-metabolite PK model with its final parameter estimates, treatment simulations were performed to investigate the impact of age and body weight on
achieving target $C_{SS,min}^{ENDX}$ under tamoxifen standard dosing. In two separate simulation study set-ups, 14 large virtual patient populations (n=10,000) with CYP2D6 AS frequencies extrapolated from the pooled dataset and different age and body weight ranges or combinations thereof, respectively, were generated (Table 2):

Study set-up 1: endoxifen subtarget concentrations for subpopulations with different age and body weight distributions

Study set-up 1 (SU1) was based on the observed distributions of age and body weight in the pooled dataset. Achievement of target $C_{SS,min}^{ENDX}$ was compared between patients with low or high covariate values (<1st quartile and >3rd quartile, respectively) and patients with covariate values in the interquartile range (IQR) of the covariate value distribution in the pooled dataset (‘reference subpopulation’; in total: 7 patient populations, Table 2). Specifically, for each virtual patient, an age and body weight value were sampled independently with replacement from the respective section (e.g. <1st quartile) of the covariate value distribution in the pooled dataset.

Study set-up 2: endoxifen subtarget concentrations for subpopulations with extreme age and body weight values

In study set-up 2 (SU2), achievement of target $C_{SS,min}^{ENDX}$ was compared between virtual patients with minimum or maximum covariate values and patients with median covariate values in the pooled dataset (‘reference subpopulation’; in total: 7 patient populations, Table 2).

To account for parameter uncertainty, a large number of simulations (n=10,000) using bootstrapped parameter sets were performed for each subpopulation and the 50th, 5th and 95th percentiles were used to determine medians and 90% confidence intervals (CIs), respectively, of (i) the fraction of patients of the respective subpopulation at risk for $C_{SS,min}^{ENDX}$ target non-attainment, (ii) the absolute change in risk compared to the reference subpopulation in SU1 and SU2, respectively, (iii) the relative change in risk compared to the reference subpopulation in SU1 and SU2, respectively and (iv) the number needed to treat/harm (NNT/NNH), defined as 1 divided by the absolute change in risk, compared to the reference subpopulation in SU1 and SU2, respectively.

Finally, for the two respective patient populations that had shown the highest risks to miss target $C_{SS,min}^{ENDX}$ in SU1 and SU2, $C_{SS,min}^{ENDX}$ at alternative daily tamoxifen doses of 40 mg and 60 mg were simulated and medians and 90% CIs of the fractions of patients at risk to miss target $C_{SS,min}^{ENDX}$ were calculated.

RESULTS

External model evaluation

The original joint-parent metabolite population PK model of tamoxifen and endoxifen performed well for the external evaluation dataset: MPEs indicated a low bias for tamoxifen (-13.9 ng/mL) and a minimal bias for endoxifen (-0.923 ng/mL). Precision was acceptable for both tamoxifen and endoxifen as indicated by MAPEs <8% (7.62% and 6.29%, respectively)36. After parameter re-estimation using the pooled dataset, all fixed (structural and covariate) parameter estimates remained comparable except the tamoxifen clearance $CL_{20/F}$ for a typical (AS 2, median age 55 years) patient (development dataset: 6.51 L/h
(2.4% RSE), pooled dataset: 5.08 L/h (1.1% RSE)) and the exponent for the typical age effect on the tamoxifen clearance (development dataset: -0.844 (10.0%), pooled dataset: -0.148 (24.0%)). Furthermore, estimated IIV values on CL20/F and CL23/F were slightly lower (40.4% vs. 41.5% and 46.1% vs. 49.2%, respectively) for the pooled dataset compared to the development dataset.

Extended covariate analysis and final model development
A significant difference between $C_{\text{SS, min ENDX}}$ of pre- (n=67) and post-menopausal (n=368) patients was observed in the development dataset (97.4 % Caucasians, Table S1): while 29.9% of pre-menopausal patients showed subtarget $C_{\text{SS, min ENDX}}$<5.97 ng/mL, it was only 20.1% of post-menopausal patients (Table 3). Conversely in the evaluation dataset, with 18.8% and 18.0% of patients with subtarget $C_{\text{SS, min ENDX}}$ (Table 3), there was no difference in $C_{\text{SS, min ENDX}}$ between pre- (n=568) and post-menopausal (n=367) patients (Table S1). However, after stratifying patients in the evaluation dataset for their ethnicity, a highly significant difference between $C_{\text{SS, min ENDX}}$ in pre- and post-menopausal Caucasians became apparent (Tables 3 and S1). Furthermore, there were large differences between $C_{\text{SS, min ENDX}}$, ascending from pre-menopausal African, Caucasian, Middle-Eastern Arab and Asian to Indian patients (Table S1). Indians, Asians and Middle-Eastern Arabs showed the lowest number of patients at risk for subtarget $C_{\text{SS, min ENDX}}$ (0%, 5.8% and 13.0%, respectively) while Africans and Caucasians showed the highest (50.0% and 26.1%, respectively). Of note, relative risk reductions due to transition from pre- to post-menopause were 32.8% in the development dataset (n=433 Caucasians, n=2 Africans), 4.26% for the evaluation dataset without stratification for ethnicity (n=935) and 31.0% for Caucasians in the evaluation dataset (n=681) (no further analysis was possible as no data from post-menopausal patients of other ethnicities were available). Upon stratification for CYP2D6 phenotype, the observed differences in $C_{\text{SS, min ENDX}}$ between pre-menopausal patients of different ethnicities remained.

Further exploratory data and graphical analyses revealed a correlation between body weight and ethnicity in the evaluation dataset. Body weight was highest in pre-menopausal Middle-Eastern Arabs, followed by Caucasians, Africans, Indians and Asians (Table S2). Furthermore, patients of ethnicities with low body weights demonstrated a lower risk for subtarget $C_{\text{SS, min ENDX}}$ compared to patients of ethnicities with high body weights (Figure S1). Subsequently, both ethnicity and body weight were taken forward and tested for significance on CL20/F, CL23/F and CL30/F in the extended covariate analysis.

Covariate relationships of CYP2D6 AS on CL23/F (categorical), age and body weight on CL20/F (both power functions) and ethnicity on CL20/F (categorical) were all found to be significant in univariate analyses. Including ethnicity on CL20/F in addition to body weight, however, did not further improve model predictions. Due to the stronger physiological plausibility, body weight remained and ethnicity was excluded as covariate on CL20/F in the final model. Thus, the updated full covariate model (schematic representation in Figure 1) included three covariate relationships: CYP2D6 AS on CL23/F and age and body weight on CL20/F (final parameter estimates and their relative standard errors (RSEs) in Table 4). The population estimate for the power exponent of age was -0.17 (RSE: 21%), thus the tamoxifen clearance was estimated to moderately decrease with increasing age. In contrast, the population estimate for the power exponent of body weight was 0.284 (RSE: 19%), indicating a moderately increasing clearance with increasing body weight. With RSE≤28%, all model parameters were estimated with good precision. GOF plots showed good model performance in predicting observed individual tamoxifen and endoxifen concentrations (Figure S2).
Treatment simulations for different patient subpopulations

Study set-up 1: endoxifen subtarget concentrations for subpopulations with different age and body weight distributions

Up to 3.1-fold differences in reaching target $C_{SS,\text{min \text{ENDX}}}$ were observed between different patient subpopulations in SU1 (Figure 2, Table S3): heavy young patients (<40 years, >76 kg) showed the highest risk to miss target $C_{SS,\text{min \text{ENDX}}}$ (36.9%, 90% CI: 34.6%-39.2%), while light elderly patients (>65 years, <60 kg) showed the lowest risk (12.1%, 90% CI: 10.8%-13.4%). gIM were most sensitive to changes in covariate values: while the NNH for heavy young gNM and gPM was 8 and 9, respectively, it was 5 in gIM (Table S4). This means, that every 5th, every 8th and every 9th heavy young gIM, gNM and gPM, respectively, would not reach target $C_{SS,\text{min \text{ENDX}}}$ compared to patients of the respective CYP2D6 phenotype category with body weight and age within the IQR.

Study set-up 2: endoxifen subtarget concentrations for subpopulations with extreme age and body weight values

The patterns observed in SU1 were expectedly even stronger in SU2 (Table 2): Quantitatively, up to 13.8-fold differences in therapeutic target attainment were observed between heavy young (22 years and 150 kg) and light elderly (95 years and 39 kg) patients (70.6% (90% CI: 66.2%-75.1%) vs. 5.10%, (90% CI: 4.18%-6.22%) of patients at risk, respectively) (Figure 3, Table S5). NNH were again lowest in heavy young patients (2 for gNM and gIM, 6 for gPM) (Table S6).

In both study set-ups, the impact of body weight on endoxifen $C_{SS,\text{min \text{ENDX}}}$ was more pronounced than the impact of age as displayed by e.g. the lower relative risk increase in young (Median: +13.0%; 90% CI: 6.50% - 19.4%) compared to heavy patients (Median: +58.1%; 90% CI: 49.8% - 66.8%) when compared to patients in SU1.

As heavy young patients showed the highest risk to miss target $C_{SS,\text{min \text{ENDX}}}$ in both study-set-ups, $C_{SS,\text{min \text{ENDX}}}$ target attainment at 40 mg and 60 mg tamoxifen Q.D. was assessed for this subpopulation in both SU1 and SU2.

In SU1, 40 mg tamoxifen Q.D. were sufficient to reduce the fraction of patients at risk for subtarget $C_{SS,\text{min \text{ENDX}}}$ from 36.9% to 10.6% (90% CI: 9.44-11.8%). This fraction varied substantially between CYP2D6 phenotypes (3.06% for gNM, 14.2% for gIM and 62.0% for gPM). In SU2, 32.2% (90% CI: 27.9% - 36.9%) of patients receiving 40 mg tamoxifen Q.D. were still at risk for subtarget $C_{SS,\text{min \text{ENDX}}}$ when compared to patients in SU1.

At 60 mg tamoxifen Q.D., 4.10% (90% CI: 3.48% - 4.77%) and 15.8% (90% CI: 13.2% - 18.9%) of patients were still at risk for subtarget $C_{SS,\text{min \text{ENDX}}}$ in SU1 and SU2, respectively. When stratified for CYP2D6 phenotype, 0.600% of gNM, 4.63% of gIM and 36.2% of gPM showed subtarget $C_{SS,\text{min \text{ENDX}}}$ in SU1 while it was 5.83% of gNM, 22.3% of gIM and 74.4% of gPM in SU2.

DISCUSSION

We identified young overweight breast cancer patients as a subpopulation at increased risk for not achieving clinically relevant endoxifen levels ($C_{SS,\text{min \text{ENDX}}} > 5.97 \text{ ng/mL}$) during adjuvant tamoxifen treatment. This finding is of potential clinical relevance since premenopausal breast
cancer patients highly depend on the efficacy of tamoxifen for the control of their disease given that ovarian function suppression in combination with an aromatase inhibitor can only be considered for a small portion of high-risk patients. Therefore, every effort needs to be made to increase tamoxifen efficacy particularly in those patients with an intact CYP2D6 function for the sufficient metabolisation of tamoxifen to endoxifen.

The strength of our study is its large cohort size of 1388 pre- and post-menopausal tamoxifen-treated breast cancer patients with a wide ranging body weight of 39 to 150 kg. This allowed us to reliably identify and quantify the influence of body weight on the target endoxifen steady state concentration in addition to the impact of genetically determined factors.

Using treatment simulations to investigate $C_{\text{SS, min ENDX}}$ in different patient subpopulations, young overweight patients were identified as the subpopulation at highest risk for subtarget $C_{\text{SS, min ENDX}}$. The design of SU1 was chosen to consider the ‘real-world’ variability of the covariate distributions and to decrease potential bias of the simulation results due to extreme values observed in the pooled dataset. In contrast, SU2 aimed to assess ultimate best- and worst-case scenarios, as could be expected considering the covariate values observed in the pooled dataset. The large number of 10,000 patients for each subpopulation was used to represent the distribution of CYP2D6 phenotypes observed in ‘real-world’ populations and allowed the generation of sufficient numbers of virtual patients with rare CYP2D6 genotypes (e.g., AS=0: 5.56%) in each subpopulation. Furthermore, it allowed to represent the high inter-individual variation observed in real-world data. The large number of simulations with bootstrapped parameter sets for each subpopulation allowed us to additionally determine confidence intervals for the fractions of patients at risk.

The large size of our study dataset allowed us to revise and update the relationship between increasing age and decreasing tamoxifen clearance previously described by us$^{21}$. At first sight, this relationship was far less pronounced in the current evaluation data set compared to the development dataset indicated by a higher (less negative) power exponent in the covariate relationship of tamoxifen clearance and age. This can be explained by different body weight distributions in the two datasets. While body weight was similar in pre- and post-menopausal patients in the development dataset (Table S2), it was significantly lower in pre-menopausal compared to post-menopausal patients in the evaluation dataset ($p<0.001$). The latter might be explained by differences in ethnicities and cultural background. Especially pre-menopausal patients from India and Asia had a lower body weight compared to individuals of Caucasian decent who were the only ethnic group in postmenopausal patients.

Thus, the opposing influences of low body weight and young age on the tamoxifen clearance could have masked each other in the evaluation dataset. In support of this hypothesis, relative risk reductions due to the transition from pre- to post-menopause were similar in Caucasian patients of both datasets (32.8% in the development dataset, 31.0% in the evaluation dataset). Mechanistic explanations for our finding of decreased tamoxifen and endoxifen concentrations in patients with high body weight include either (i) an increased clearance due to increased body weight causing an increased liver size and function$^{37}$ or (ii) an increased distribution of the more lipophilic compound tamoxifen into fat tissue (logP-values: 7.1, 6.7 and 6.3 for tamoxifen, N-desmethyltamoxifen and endoxifen, respectively$^{38-40}$). Decreased concentrations of tamoxifen’s lipophilic metabolite N-desmethyltamoxifen in patients with high body mass indices (BMI) compared to patients with low BMIs have been reported before$^{8}$ and no influence of body weight on endoxifen formation (CL23/F) and endoxifen
Clearance (CL30/F) was determined in our extended covariate analysis, supporting the latter hypothesis. Further efforts should also be made to investigate the reason for increased tamoxifen and endoxifen concentrations in post-menopausal compared to pre-menopausal patients. Contributing factors could be a decrease in CYP3A4 activity with increasing age due to decreasing liver volumes, resulting in reduced clearance rates \(^41\), the influence of oestrogen concentrations \(^7,42\), or differences in drug adherence \(^43\).

Our dose escalation simulations for young overweight patients clearly demonstrated that 40 mg tamoxifen Q.D. were more adequate for gIMs, reducing the number of patients at risk for subtarget \(C_{SS,\text{min ENDX}}\) to 14.2% in SU1. However, 44.7% of young overweight gIMs were still at risk in SU2. Moreover, 40 mg and even 60 mg tamoxifen Q.D. were not enough to reduce the number of young overweight gPMs with subtarget \(C_{SS,\text{min ENDX}}\) to less than 36.2% and 74.4% in SU1 and SU2, respectively. From this it follows that other treatment options like aromatase inhibitors with ovarian function suppression should be used for young overweight gPMs and obese gIMs, which is an alternative supported by prospective clinical data \(^44\).

It is important to note, that while covariates CYP2D6 AS, body weight and age could explain important general trends within the population, the IIV in both tamoxifen clearance (39.9% CV, RSE: 3%) and in endoxifen formation (46% CV, RSE: 3%) remained high. Thus, individual \(C_{SS,\text{min ENDX}}\) may deviate from the predictions for typical patients. We therefore strongly advocate to use a model-informed precision dosing framework to identify personalised tamoxifen doses for \(C_{SS,\text{min ENDX}}\) target attainment \(^21\): based on a patient’s CYP2D6 AS, age and body weight, the developed model can be applied to guide initial dose selection and, if needed, dose refinement upon availability of measured \(C_{SS,\text{min ENDX}}\) based on therapeutic drug monitoring. In this respect, it should be mentioned that the endoxifen therapeutic target threshold used in this study is yet controversial. However, a recent report from a prospective clinical trial that suggested no relationship between CYP2D6 genotype or \(C_{SS,\text{min ENDX}}\) and treatment outcome \(^14\) has been provoking large criticism with regards to applied methods \(^17,45,46\) and low statistical power \(^16\). Thus, a properly designed and well-powered prospective clinical trial \(^16\) is needed to assess the relationship between CYP2D6 genotype or \(C_{SS,\text{min ENDX}}\) and breast cancer outcome. Provided the threshold or a similar clinical concentration cut-off point will be confirmed, a patient’s CYP2D6 genotype, body weight and age should be considered in an individualised dose selection process in order to reach therapeutic endoxifen levels. In conclusion, young overweight breast cancer patients treated with tamoxifen present a vulnerable patient subpopulation with the risk not to achieve therapeutically relevant endoxifen concentrations. Considering that obesity rates are on the rise globally and the relevance of tamoxifen for pre-menopausal patients, this subpopulation may benefit from individualised tamoxifen such as dose adjustments in order to increase their chances of clinical benefit.

Study Highlights

- **What is the current knowledge on the topic?**

Large interindividual variability in concentrations of tamoxifen’s most active metabolite endoxifen is observed during standard breast cancer tamoxifen treatment. Minimal steady-state endoxifen concentrations have been suggested below which the
risk for breast cancer recurrence and mortality is increased. The influence of age and body weight on endoxifen concentrations is not well established.

- **What question did this study address?**
  What is the quantitative impact of age and body weight on the pharmacokinetics (PK) of tamoxifen and endoxifen beyond the patients’ genetically determined CYP2D6 tamoxifen metaboliser capacity?

- **What does this study add to our knowledge?**
  Age and body weight contribute to the PK of tamoxifen and endoxifen in that young and overweight patients are at increased risk to not achieve sufficient endoxifen concentration.

- **How might this change drug discovery, development, and/or therapeutics?**
  Obese pre-menopausal patients may benefit from individualised tamoxifen dosing, particularly in the case of an intact genetically determined tamoxifen drug metabolism. If their CYP2D6 function is impaired, alternative endocrine treatment of ovarian function suppression combined with AI should be considered.

**Acknowledgements**

The authors thank the High-Performance Computing Service of ZEDAT at Freie Universitaet Berlin (https://www.zedat.fu-berlin.de/HPC/Home) for computing time.

**Author Contributions**

AM wrote the manuscript, AM, AT, BC, CK, DE, EC, HB, MJ, MS, NZ, PN, RM, SC, SK, TM, WH and WS designed and performed the research and AM and LK analysed the data. All authors contributed to manuscript revision, read and approved the submitted version.

**References**


22. Antunes, M. V. et al. CYP3A4* 22 is related to increased plasma levels of 4-hydroxytamoxifen and partially compensates for reduced CYP2D6 activation of tamoxifen. Pharmacogenomics 16, 601–617 (2015).


**Figure Legends**

**Figure 1.** Schematic representation of the joint tamoxifen and endoxifen pharmacokinetic model and the implemented covariate relationships. CL30/F: relative clearance of endoxifen; CL20/F: relative clearance of tamoxifen; CL23/F: relative formation of endoxifen; CYP2D6 AS: CYP2D6 activity scores as ordered categorical covariate from 0 to ≥2 in increments of 0.5; ENDX: endoxifen compartment with VENDX/F; kₐ: absorption rate constant; Gut: tamoxifen dose in gut compartment; TAM: central tamoxifen compartment with VTAM/F; tlag: lag time; bold: estimated parameters (other parameters fixed to values from literature). kₐ, tlag, VTAM/F: fixed to estimates using the development dataset (with rich sampling data).

**Figure 2.** Patients at risk of subtarget endoxifen concentrations across patient subpopulations in study-setup 1 (see main text for further explanation) as observed in 1 of the 1000 stochastic
Simulated minimum steady-state concentrations of endoxifen (ENDX min. conc. at SS) in 7 different patient populations with covariate characteristics as indicated on the right. Dashed horizontal line: endoxifen target threshold; boxes: interquartile range (IQR), including median; whiskers: range from hinge to lowest/highest value within 1.5 IQR; points: data outside whiskers.

**Figure 3.** Patients at risk of subtarget endoxifen concentrations across patient subpopulations in study-setup 2 (see main text for further explanation) as observed in 1 of the 1000 stochastic simulations. Simulated minimum steady-state concentrations of endoxifen (ENDX min. conc. at SS) in 7 different patient populations with covariate characteristics as indicated on the right. Dashed horizontal red line: endoxifen target threshold; boxes: interquartile range (IQR), including median; whiskers: range from hinge to lowest/highest value within 1.5 IQR; points: data outside whiskers.

**Supplementary Information Titles**

**Table S1.** Endoxifen steady-state minimum concentrations across different subpopulations in the development and evaluation dataset.

**Table S2.** Body weights across different subpopulations in the development and evaluation dataset.

**Table S3.** Number of patients at risk for subtarget \( C_{SS,min} \) ENDX and absolute and relative risk changes for different patient subpopulations in SU1.

**Table S4.** Number of patients at risk for subtarget \( C_{SS,min} \) ENDX and absolute and relative risk changes for different patient subpopulations in SU1, stratified for CYP2D6 phenotype.

**Table S5.** Number of patients at risk for subtarget \( C_{SS,min} \) ENDX and absolute and relative risk changes for different patient subpopulations in SU2.

**Table S6.** Number of patients at risk for subtarget \( C_{SS,min} \) ENDX and absolute and relative risk changes for different patient subpopulations in SU2, stratified for CYP2D6 phenotype.

**Figure S1.** Endoxifen steady-state observed minimum concentrations [ng/mL] versus body weight in the pooled dataset, stratified for ethnicity.

**Figure S2.** Goodness-of-fit plots for the updated full joint parent-metabolite pharmacokinetic model of tamoxifen and endoxifen using the pooled dataset.
Table 1. Clinical study and population characteristics of the development, evaluation and pooled dataset at baseline.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Development dataset</th>
<th>Evaluation dataset</th>
<th>Pooled dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td>452</td>
<td>936</td>
<td>1388</td>
</tr>
<tr>
<td>Age [years] Median (range)</td>
<td>64 (25-95)</td>
<td>48 (22-84)</td>
<td>55 (22-95)</td>
</tr>
<tr>
<td>Body weight [kg] Median (range)</td>
<td>70 (42-150)</td>
<td>66 (39-144)</td>
<td>67 (39-150)</td>
</tr>
<tr>
<td>Frequency of CYP2D6 phenotypes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gNM</td>
<td>53.5%</td>
<td>54.0%</td>
<td>53.8%</td>
</tr>
<tr>
<td>gIM</td>
<td>34.5%</td>
<td>39.4%</td>
<td>37.8%</td>
</tr>
<tr>
<td>gPM</td>
<td>5.53%</td>
<td>5.56%</td>
<td>5.55%</td>
</tr>
<tr>
<td>(according to 33)</td>
<td>6.42%</td>
<td>1.07%</td>
<td>2.81%</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>97.4% Caucasian</td>
<td>72.4% Caucasian</td>
<td>80.6% Caucasian</td>
</tr>
<tr>
<td>0.44% African</td>
<td>1.28% African</td>
<td>0.87% African</td>
<td></td>
</tr>
<tr>
<td>2.21% n.r.</td>
<td>16.4% Asian</td>
<td>11.0% Asian</td>
<td></td>
</tr>
<tr>
<td>8.23% Middle-Eastern Arab</td>
<td>1.28% Indian</td>
<td>0.87% Indian</td>
<td></td>
</tr>
<tr>
<td>0.43% n.r.</td>
<td>1.01% n.r.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Menopausal status</td>
<td>60.0% Pre-menopausal</td>
<td>41.0% Pre-menopausal</td>
<td>32.7%</td>
</tr>
<tr>
<td>100% n.r.</td>
<td>39.0% Post-menopausal</td>
<td>26.3% Post-menopausal</td>
<td></td>
</tr>
<tr>
<td>Treatment setting</td>
<td>41.6% Adjuvant</td>
<td>100% Adjuvant</td>
<td>81.0% Adjuvant</td>
</tr>
<tr>
<td>13.1% Neo-Adjuvant</td>
<td>4.25% Neo-Adjuvant</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22.1% Primary metastatic</td>
<td>7.20% Primary metastatic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21.5% Metastatic</td>
<td>6.99% Metastatic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.7% n.r.</td>
<td>0.58% n.r.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PK sampling design</td>
<td>Sparse &amp; dense</td>
<td>Sparse</td>
<td>Sparse &amp; dense</td>
</tr>
</tbody>
</table>

\textit{gNM, gIM, gPM}: genotype-predicted CYP2D6 normal (including ultrarapid), intermediate and poor metaboliser; \textit{n.r.}: not reported; PK: pharmacokinetic(s).
Table 2: Covariate values used for simulated populations in 7 different patient subpopulations using study-setups 1 and 2 (see main text for detailed explanations of study set-ups 1 and 2).

<table>
<thead>
<tr>
<th>Subpopulation</th>
<th>Age [years]</th>
<th>Body weight [kg]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Study-setup 1</td>
<td>Study-setup 2</td>
</tr>
<tr>
<td>Heavy young</td>
<td>22-39 (&lt;Q1)</td>
<td>22 (Min.)</td>
</tr>
<tr>
<td>Young</td>
<td>22-39 (&lt;Q1)</td>
<td>22 (Min.)</td>
</tr>
<tr>
<td>Heavy</td>
<td>40-65 (IQR)</td>
<td>55 (Med.)</td>
</tr>
<tr>
<td>IQR/Median (Reference)</td>
<td>40-65 (IQR)</td>
<td>55 (Med.)</td>
</tr>
<tr>
<td>Elderly</td>
<td>66-95 (&gt;Q3)</td>
<td>95 (Max)</td>
</tr>
<tr>
<td>Light</td>
<td>40-65 (IQR)</td>
<td>55 (Med.)</td>
</tr>
<tr>
<td>Light elderly</td>
<td>66-95 (&gt;Q3)</td>
<td>95 (Max)</td>
</tr>
</tbody>
</table>

IQR: Interquartile range, Max: Maximum, Med: Median, Min: Minimum, Qx.: Quartile with x=1-3.
Table 3: Fraction of pre- and post-menopausal patients at risk for sub-therapeutic 
\( \text{CSS}_{\text{min}} \) \(<5.97\) ng/mL in the development and evaluation dataset.

<table>
<thead>
<tr>
<th></th>
<th>Dev. dataset (all) (n=435)</th>
<th>Eval. dataset (all) (n=935)</th>
<th>Eval. dataset (African) (n=12)</th>
<th>Eval. dataset (Arab) (n=77)</th>
<th>Eval. dataset (Asian) (n=153)</th>
<th>Eval. dataset (Caucasian) (n=681)</th>
<th>Eval. dataset (Indian) (n=12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-menopausal (%)</td>
<td>29.9% (n=67)</td>
<td>18.8% (n=568)</td>
<td>50% (n=12)</td>
<td>13.0% (n=77)</td>
<td>5.88% (n=153)</td>
<td>26.1% (n=314)</td>
<td>0% (n=12)</td>
</tr>
<tr>
<td>Post-menopausal (%)</td>
<td>20.1% (n=368)</td>
<td>18.0% (n=367)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>18.0% (n=367)</td>
<td>-</td>
</tr>
<tr>
<td>Absolute change in risk (%)</td>
<td>-9.8% (n=368)</td>
<td>-0.8% (n=367)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Relative change in risk (%)</td>
<td>-32.8% (n=367)</td>
<td>-4.26% (n=367)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-31.0% (n=367)</td>
<td>-</td>
</tr>
</tbody>
</table>

\( \text{CSS}_{\text{min}} \): Endoxifen minimum concentrations at steady-state; Dev. dataset: Development dataset; Eval. dataset: Evaluation dataset.
**Table 4:** Final parameter estimates for the updated joint parent-metabolite population pharmacokinetic model of tamoxifen and endoxifen using the pooled dataset (n=1388 patients).

<table>
<thead>
<tr>
<th>Parameter [unit]</th>
<th>Estimate</th>
<th>RSE, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_a$ [1/h]</td>
<td>1.08</td>
<td>Fixed</td>
</tr>
<tr>
<td>$t_{lag}$ [h]</td>
<td>0.442</td>
<td>Fixed</td>
</tr>
<tr>
<td>$V_{TAM,C/F}$ [L]</td>
<td>912</td>
<td>Fixed</td>
</tr>
<tr>
<td>CL30/F [L/h]</td>
<td>5.10</td>
<td>Fixed</td>
</tr>
<tr>
<td>$V_{ENDX/F}$ [L]</td>
<td>400</td>
<td>Fixed</td>
</tr>
<tr>
<td>CL20/F [L/h]</td>
<td>5.07</td>
<td>1</td>
</tr>
<tr>
<td>CL20/F_Age**</td>
<td>-0.17</td>
<td>21</td>
</tr>
<tr>
<td>CL20/F_Body weight**</td>
<td>0.284</td>
<td>19</td>
</tr>
<tr>
<td>CL23/F [L/h]</td>
<td>0.459</td>
<td>2</td>
</tr>
<tr>
<td>CL23/F_AS: 0*</td>
<td>-0.759</td>
<td>2</td>
</tr>
<tr>
<td>CL23/F_AS: 0.5*</td>
<td>-0.598</td>
<td>4</td>
</tr>
<tr>
<td>CL23/F_AS: 1*</td>
<td>-0.347</td>
<td>6</td>
</tr>
<tr>
<td>CL23/F_AS: 1.5*</td>
<td>-0.16</td>
<td>18</td>
</tr>
<tr>
<td>CL23/F_AS: 2.5-3*</td>
<td>0.302</td>
<td>28</td>
</tr>
<tr>
<td>IIV CL20/F</td>
<td>0.148</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>(39.9% CV)</td>
<td></td>
</tr>
<tr>
<td>IIV CL23/F</td>
<td>0.192</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>(46.0% CV)</td>
<td></td>
</tr>
<tr>
<td>RUV tamoxifen</td>
<td>0.0295</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>(17.3% CV)</td>
<td></td>
</tr>
<tr>
<td>COV$_{RUVtam-RUVendx}$</td>
<td>0.0228</td>
<td>7.28</td>
</tr>
<tr>
<td>RUV endoxifen</td>
<td>0.037</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>(19.4% CV)</td>
<td></td>
</tr>
</tbody>
</table>

*AS*: CYP2D6 activity score; $CL20/F$: relative tamoxifen clearance; $CL23/F$: relative endoxifen formation; $CL30/F$: apparent endoxifen clearance; IIV: interindividual variability; $k_a$: absorption rate constant; RUV: residual unexplained variability; RSE: relative standard error=(standard error/estimate)•100; $t_{lag}$: absorption lag time; $V_{TAM/F}$: tamoxifen apparent volume of distribution; $V_{ENDX/F}$: endoxifen apparent volume of distribution; *: implemented as fractional change covariate model (detailed description in\textsuperscript{23}; **: implemented as power covariate model (detailed description in\textsuperscript{23}).