Evaluation and Application of Population Pharmacokinetic Models for Identifying Delayed Methotrexate Elimination in Patients With Primary Central Nervous System Lymphoma

Junjun Mao¹, Qing Li²,³, Pei Li², Weiwei Qin¹⁺, Bobin Chen²,³⁺ and Mingkang Zhong¹

¹Department of Pharmacy, Huashan Hospital, Fudan University, Shanghai, China, ²Department of Hematology, Huashan Hospital, Fudan University, Shanghai, China, ³Department of Hematology, Huashan Hospital North, Fudan University, Shanghai, China

Objective: Several population pharmacokinetic (popPK) models have been developed to determine the sources of methotrexate (MTX) PK variability. It remains unknown if these published models are precise enough for use or if a new model needs to be built. The aims of this study were to 1) assess the predictability of published models and 2) analyze the potential risk factors for delayed MTX elimination.

Methods: A total of 1458 MTX plasma concentrations, including 377 courses (1–17 per patient), were collected from 77 patients who were receiving high-dose MTX for the treatment of primary central nervous system lymphoma in Huashan Hospital. PopPK analysis was performed using the NONMEM® software package. Previously published popPK models were selected and rebuilt. A new popPK model was then constructed to screen potential covariates using a stepwise approach. The covariates were included based on the combination of theoretical mechanisms and data properties. Goodness-of-fit plots, bootstrap, and prediction-corrected visual predictive checks were used to determine the stability and predictive performance of both the published and newly built models. Monte Carlo simulations were conducted to qualify the influence of risk factors on the incidence of delayed elimination.

Results: Among the eight evaluated published models, none presented acceptable values of bias or inaccuracy. A two-compartment model was employed in the newly built model to describe the PK of MTX. The estimated mean clearance (CL/F) was 4.91 L h⁻¹ (relative standard error: 3.7%). Creatinine clearance, albumin, and age were identified as covariates of MTX CL/F. The median and median absolute prediction errors of the final model were -10.2 and 36.4%, respectively. Results of goodness-of-fit plots, bootstrap, and prediction-corrected visual predictive checks indicated the high predictability of the final model.

Conclusions: Current published models are not sufficiently reliable for cross-center use. The elderly patients and those with renal dysfunction, hypoalbuminemia are at higher risk of delayed elimination.

Keywords: methotrexate, population pharmacokinetics, external evaluation, delayed elimination, Monte carlo simulations
1 INTRODUCTION

High-dose methotrexate (HD-MTX, ≥ 1 g m⁻²) is the base therapy for the treatment of various lymphoid malignancies, such as acute lymphoblastic leukemia (Sakura et al., 2018) and non-Hodgkin’s lymphoma (Reiter et al., 1999), especially for treating primary or secondary central nervous system lymphoma (Plotkin et al., 2004; Zhu et al., 2009). As an antifolate inhibitor of dihydrofolate reductase, MTX may cause the depletion of purines and thymidylate, which inhibits DNA synthesis, leading to cell death (Baram et al., 1987).

Following intravenous infusion, approximately 60% of MTX binds to plasma protein (Maia et al., 1996). As a small polar molecule, MTX elimination is highly correlated with renal function. Approximately only 10% of MTX is excreted as unchanged drug in the bile, whereas the majority is eliminated as unchanged drug through the kidneys within 24 h (Csordas et al., 2013).

Both the pharmacokinetic (PK) of MTX, which exhibit wide inter-individual variability (IIV), and its exposure are directly related to efficacy and toxicity (Evans et al., 1986; Evans et al., 1998). Patients experiencing delayed MTX elimination have been reported to be at an elevated risk of toxicity, such as nephrotoxicity, myelotoxicity, mucositis, neurological complications, and other adverse effects, which may lead to significant morbidity and delays in treatment (Howard et al., 2016). To prevent this systemic toxicity, supportive care, such as fluid hydration, urine alkalization, and leucovorin rescue, is conducted as HD-MTX is administered (Widemann and Adamson, 2006). Post-dose therapeutic drug monitoring is routinely performed to maintain MTX plasma concentrations within the cytotoxic range for leukemic cells and, below those associated with toxicity (Wall et al., 2000; Paci et al., 2014).

Delayed MTX elimination is defined as plasma MTX concentrations above 50 μmol L⁻¹ at 24 h, above 5 μmol L⁻¹ at 48 h, or above 0.2 μmol L⁻¹ at 72 h (Hospira, 2017). As MTX-induced nephrotoxicity correlates with clearance (CL), delayed elimination is closely related to acute kidney injury (Huang et al., 2020). Thus, if patients with delayed MTX CL can be identified, allowing for the implementation of a personalized dosage regimen before chemotherapy, concentration-related toxicity may be avoided (Ramsey et al., 2018). Although some oncology guidelines recommend MTX dosage based on the patient’s body size (i.e., body surface area or actual body weight) (Gurney and Shaw, 2007; Griggs et al., 2012), this strategy is not always suitable for clinical practice, particularly for obese patients (Gallais et al., 2020; Pai et al., 2020). Therefore, determining the risk factors for delayed MTX elimination is essential.

Compared to conventional PK analysis, population pharmacokinetics (popPK) is a superior approach in facilitating the understanding and quantification of PK variability (Ette and Williams, 2004). PopPK models combined with maximum posterior Bayesian estimation can be used to guide dosing regimen individualization.

A few popPK models have been constructed to assess the sources of MTX PK variability. However, whether these popPK models can be extrapolated to other clinical centers remains unknown. In addition, inconsistencies and differences in study design, research purpose, and population properties have been noted in some of the published popPK models with regard to model structure, parameter estimates, selected covariates, and their functional forms on PK parameters (Mao et al., 2018; Cheng et al., 2020). As the published models were developed based on specific populations, selecting an appropriate model to guide precision dosing in clinical practice is challenging.

Whether the published models are sufficiently precise for use in patients with primary central nervous system lymphoma, or whether a new popPK model based on our center is needed, remains unknown. Therefore, we summarized and assessed the predictability of published HD-MTX popPK models in adult patients with lymphoid malignancies. In addition, a new popPK model was constructed to investigate the effects of physiological parameters on the distribution and elimination of MTX. The predictability of this model was also compared with previously published models. Finally, the most suitable model was applied to identify the population with delayed elimination.

2 MATERIALS AND METHODS

2.1 Patient Data Collection

Data from 77 adults (49 men and 28 women) who were diagnosed with primary central nervous system lymphoma and received HD-MTX (>1 g m⁻²) for treatment between June 2011 and November 2016 were retrospectively collected.

To protect against MTX-induced renal dysfunction, hydration and alkalization (urine pH > 7) were achieved 12 h prior to initiating MTX therapy (Hospira, 2017). Serial plasma MTX levels were measured at 24, 48, and 72 h until the plasma concentration was ≤0.2 μmol L⁻¹ (Dupuis et al., 2008). The leucovorin rescue was initiated and repeated every 6 h after 24 h from the MTX infusion until the MTX concentration was lower than ≤0.2 μmol L⁻¹.

MTX concentrations in plasma were determined using a well-validated enzyme-multiplied immunoassay (EMIT) using the SYVA Viva-Emit 2000 Kit (Siemens Healthcare Diagnostics, Newark, DE, United States). The limit of detection was 0.3 μmol L⁻¹, and the calibration concentrations ranged from 0.3 μmol L⁻¹ to 2,600 μmol L⁻¹.

Demographic covariates, including age, sex, weight, body surface area (BSA), and concomitant medication with benzimidazoles and corticoids, were included in the database. Clinical covariates, such as serum creatinine (Scr) levels, were recorded before each MTX infusion and in tandem with MTX plasma samples. According to the standard of ‘drug interaction score’ (Benz-de Bretagne et al., 2014), benzimidazoles and corticoids were assigned scores of 2 and 1, respectively.

The Ethics Committee of Huashan Hospital approved the study protocols. Written informed consent was obtained from all volunteers.
2.2 Population Pharmacokinetic Models

Our study consisted of the following steps:

Step 1: PopPK models of HD-MTX in adult patients with lymphoid malignancies were selected and reviewed.

Step 2: A new popPK model was constructed based on our dataset.

Step 3: The predictability of published popPK models and that of the newly built model was evaluated.

Step 4: The most suitable model was applied to identify the population with delayed elimination.

2.2.1 Review of Published popPK Studies on HD-MTX

A systematic review of popPK studies on HD-MTX published in English before 31 December 2020 was performed using PubMed, Web of Science, and Embase. According to the Preferred Reporting Items for Systematic Reviews and Meta-analyses statement, the relevant identification, screening, and assessment were conducted (Moher et al., 2015).

The inclusion criteria for published studies were as follows: 1) studied population: adult patients with lymphoid malignancies; 2) treatment: HD-MTX; 3) PK analysis using NONMEM® software package; and 4) language: English. Studies were excluded if they 1) lacked the required details for external evaluation and re-estimate or 2) overlapped with other data or cohorts. Reference lists of the identified reports were also screened.

The demographic characteristics and following popPK parameters were collected from each identified study: apparent clearance (CL/F), apparent volume of distribution (V/F), and corresponding between-subject variability and residual variability.

2.2.2 Development of a Model Based on Our Dataset

NONMEM® software package (version 7.4; ICON Development Solutions, Ellicott City, MD, United States) with Piramis® 2.9 as an interface for Perl Speaks NONMEM (PsN; version 4.9.0) was used for popPK analysis (Keizer et al., 2013). R software (version 3.5.0, http://www.r-project.org/) was used to construct the visualizations for output and model evaluations.

Based on a literature review and visual data inspection, the concentration-time profile of MTX was described by two-compartment models. Model variability and random effects were classified as one of three types of errors: between-subject variability (BSV), inter-occasion variability (IOV), and residual unexplained variability (RUV). BSV was assumed to be log-normally distributed and an estimate for all parameters. IOV was assumed to be the same for all occasions (Karlsson and Sheiner, 1993). RUV was described by testing proportional and combined proportional as well as additive structures. The first-order conditional estimation method including h-ε interaction (FOCE-I) was used for the model (Beal et al., 1989).

The evaluated covariates included demographic and pathophysiological data, as well as concomitant medications (Table 1). Age, body size, hematocrit (HCT), albumin (ALB), creatinine clearance (CrCl), and concomitant medications were evaluated as possible covariates of MTX PK. CrCl was estimated using the Cockcroft-Gault equation (Cockcroft and Gault, 1976).

As the most frequently identified covariate, the effect of CrCl on MTX PK CL/F was tested first. The other covariates were screened according to a previous study and their clinical relevance (Mao et al., 2018). Each co-administered drug was assigned a “drug interaction score (DIS)” list (Supplementary Table S1), as presented previously (Benz-de Bretagne et al., 2014), and was considered individually by testing its effect on PK parameters as categorical variable.

The influence of continuous covariates was explored using the linear, exponential, and power function models, whereas that of categorical variables was described using a shift model. After considering the most frequently identified covariates in the model, the remaining covariates were screened using a stepwise approach primarily based on objective function value (OFV) (Beal et al., 1989), and parameter precision. Error estimates were also considered.

The likelihood ratio tests at a significance level of $p < 0.05$ ($\Delta$OFV >3.84) and $p < 0.001$ ($\Delta$OFV >10.83) were performed in forward inclusion and backward elimination procedures, respectively. Moreover, the clinical meaning of parameters with a significant reduction of model variability in covariate selection was also considered. In the modeling process, the condition numbers were calculated to avoid over-parameterization, accepting no more than 1,000 as the criterion (Owen and Fiedler-Kelly, 2014).

2.2.3 Implementation of Published Models

Published models were rebuilt and fixed parameters were reported in each study. Prediction- and simulation-based...
diagnostics were then used to evaluate the predictive performance of the published models (Zhao et al., 2016; Mao et al., 2018). If specific continuous covariates were missing, the median of the dataset or the model population was imputed. The data were assumed to be in the negative category (e.g., not receiving concomitant omeprazole) if categorical covariates were not available.

2.3 Model Evaluation
The predictability of published popPK models and the newly built model was evaluated by prediction- and simulation-based diagnostics. To compare the accuracy and precision of model predictive performance, prediction-based prediction error (PE, Eq. 1), median prediction error (MDPE), and median absolute prediction error (MAPE) were calculated and estimated (Sheiner and Beal, 1981).

$$\text{PE} (\%) = \left( \frac{\text{PRED} - \text{OBS}}{\text{OBS}} \right) \times 100$$ (1)

The percentages of PE within 20% (F20) and 30% (F30) were used as the combination index of both the accuracy and precision.

The goodness-of-fit plots were examined for model evaluation. The model stability and precision of parameter estimates were assessed using the bootstrap method (Ette et al., 2003). By random sampling with replacement in Perl modules, 2000 bootstrap datasets were generated (Ette, 1997). The final popPK model was compared with each of the bootstrap datasets to obtain 95% confidence intervals (CI) for all model parameters.

The predictability of the candidate model was evaluated using prediction-corrected visual predictive checks (pcVPcS) with 2000 simulations (Bergstrand et al., 2011). The 95% CI for the median, and the 5th and 95th percentiles of the simulations were calculated and compared with the observations, binning automatically.

2.4 Model Application
Using parameter estimates from the most suitable model, Monte Carlo simulations were performed. The objective of this study was to determine the influence of covariates on the incidence of delayed MTX elimination. The proportion of patients with MTX concentrations ≤0.2 μmol L\(^{-1}\) at 72 h was analyzed, as 3 g m\(^{-2}\) was administered for standard patients (with BSA 1.6 m\(^{2}\)) with different covariate levels (2.5th percentile, median, and 97.5th percentile). After simulating 1,000 hypothetical individuals, the concentration profiles were obtained in each scenario.

3 RESULTS

3.1 Patients and Data Collection
Patient demographic and physical characteristics are presented in Table 1. Data from 77 patients, covering 377 courses (1–17 per patient), and 1458 MTX plasma concentrations were available for analysis. The doses administered to patients ranged from 2 to 15.8 g. MTX dosage was transformed into molar equivalents by dividing them by the molecular weight (MTX: 222 g mol\(^{-1}\), http://chem.nlm.nih.gov/chemidplus/). For 92.3% of the treatment cycles, patients had short infusions, ranging from 1 to 4 h, whereas the remaining patients had long infusions, ranging up to 28.25 h.

In total, 567 concentrations were below the limit of quantification (LOQ), among which 236 were the second samples below the LOQ in the same treatment cycle. The M6 method was used to handle samples below the LOQ of 0.3 μmol L\(^{-1}\) as Gallais et al. done previously (Beal, 2001; Gallais et al., 2020). For concentrations under the lower LOQ (LLOQ), the first measurement in each continuous series was set to LLOQ/2, with the following measurements being treated as missing values. More complex approaches such as the M3/M4 method (likelihood estimation) did not improve model predictability, and thus were not evaluated further. The description of the sampling points and samples below the LOQ is provided in Supplementary Table S2.

3.2 Population Pharmacokinetic Models
3.2.1 Review of Published popPK Studies on HD-MTX
Eight eligible popPK studies (Faltaos et al., 2006; Min et al., 2009; Simon et al., 2013; Benz-de Bretagne et al., 2014; Nader et al., 2017; Mei et al., 2018; Gallais et al., 2020; Yang et al., 2020) were identified during the literature review for further analysis. The screening process is presented in Text S1. The details of each study are summarized in Table 2. Among them, seven were single-center studies, whereas one was conducted at two centers (Benz-de Bretagne et al., 2014). Additionally, four studies were primarily conducted in France (Faltaos et al., 2006; Simon et al., 2013; Benz-de Bretagne et al., 2014; Gallais et al., 2020), three in China (Min et al., 2009; Mei et al., 2018; Yang et al., 2020), and one in Qatar (Nader et al., 2017). Furthermore, seven studies in the analysis had a small sample size of less than 1,000 concentrations (Faltaos et al., 2006; Min et al., 2009; Simon et al., 2013; Benz-de Bretagne et al., 2014; Nader et al., 2017; Mei et al., 2018; Yang et al., 2020). Four bioassay methods, HPLC, EMIT, TDx, and HEI, were used in seven studies (Table 2). Bioassay information was not provided in one study (Nader et al., 2017).

The covariates of the published final CL/F models included age, Scr, CrCL, HCT, BSA, change in basal urinary coproporphyrin I/coproporphyrin I + III ratio value at the time of hospital discharge (MTX pre-administration [DP3]), co-administration with at least one drug of score 2 (SCO2), and ABCB2 genotype. Scr, CrCL, and age were the most frequently identified covariates in the final models and were reported in three, two, and two studies, respectively (Supplementary Table S3 and Supplementary Table S4). Body size was screened in all studies, and three included it in the volume of distribution, whereas one included it in CL/F and Q/F. Moreover, ABCC2 polymorphisms were screened in two studies, and it was included in the final model in one study (Simon et al., 2013).

3.2.2 Population Pharmacokinetic Model Development
Data was described using a two-compartment PK structural model (ADVAN3, TRANS4 subroutine) with linear elimination. The exponential model provided the best result.
| Study (publication year) | Country (single/multiple sites) | Number of samples/Patients (M/F) | Dosage regimen | Sampling schedule | Bio-assay | PK parameters and formula | BSV% (IDV%) | Residual error | Evaluation |
|--------------------------|---------------------------------|---------------------------------|----------------|------------------|----------|--------------------------|----------------|----------------|------------|
| Faltaos et al. (2006)    | France (Single)                 | 496/51 (28/23)                 | 1–8 g/m², 1–6 h, i.v. | C₂₀/C₄₀ and other two samples<sup>a</sup> | EMIT     | 7.1×(AGE/62)⁻⁰·₂² ×(SCR/67)⁻⁰·₄₃ | 22.0            | 46.0%          | GOF, Bootstrap, NPDE, RMSE |
|                          |                                 |                                 |                |                  |          | V₁/F                     | 25.1            |                |            |
|                          |                                 |                                 |                |                  |          | Q/F                      | 0.15            |                |            |
|                          |                                 |                                 |                |                  |          | V₂/F                     | 2.7             |                |            |
|                          |                                 |                                 |                |                  |          | V₃/F                     |                |                |            |
| Min et al. (2009)        | China (Single)                  | 400/82 (60/22)                 | 1.5–9 g, 24 h, i.v. | Before and 6, 12, 18,24,30,36,44,50,56,68,74,80,92 h after infusion | TDx      | 7.45×[1 + 0.22×(0.89-SCR/100)]  | 50.6            | 42.3%          | GOF, NPDE, RMSE, Cross-over validation |
|                          |                                 |                                 |                |                  |          | V₁/F                     | 25.9×[1−0.0099×7+ (86-WT)] | 22.5 |        |            |
|                          |                                 |                                 |                |                  |          | Q/F                      | 0.333           |                |            |
|                          |                                 |                                 |                |                  |          | V₂/F                     | 9.23            |                |            |
| Simon et al. (2013)      | France (Single)                 | 496/50 (27/23)                 | 1–8 g/m², 1–6 h, i.v. | At the end of infusion and 8–12,24,48, 72 h until <0.03 μmol/L | EMIT     | 3.99×(1.63, if ABC22 CT or TT) + 1.91×(SCR/89) | 28.74           | 44.4%          | GOF, Bootstrap, VPC |
|                          |                                 |                                 |                |                  |          | V₁/F                     | 19.0×(1.63, if ABC22 CT or TT) | 36.74 |        |            |
|                          |                                 |                                 |                |                  |          | Q₁/F                     | 0.1             |                |            |
|                          |                                 |                                 |                |                  |          | Q₂/F                     | 1.58            |                |            |
|                          |                                 |                                 |                |                  |          | Q₃/F                     | /               |                |            |
|                          |                                 |                                 |                |                  |          | V₄/F                     | 0.023           |                |            |
|                          |                                 |                                 |                |                  |          | /                        | 1.98            |                |            |
| Bretagne et al. (2014)   | France (Multiple)               | 363/81 (46/35)                 | 1–8 g/m², 3–24 h, i.v. | C₂₀/C₄₀/C₇₀, then q24 h until <0.2 μmol/L | EMIT     | 7.05×(CRCL/91.6)⁻⁰·₁⁶ ×0.90×SC02 | 23.0            | 41.7%          | GOF, Bootstrap, NPDE |
|                          |                                 |                                 |                |                  |          | V₁/F                     | 23.5            |                |            |
|                          |                                 |                                 |                |                  |          | Q₁/F                     | 0.13            |                |            |
|                          |                                 |                                 |                |                  |          | V₂/F                     | 3.01            |                |            |
| Nader et al. (2017)      | Qatar (Single)                  | 530/37 (31/6)                  | 0.5–7 g/m², 4–6 h or 24 h, i.v. | q12 h or q24 h until <0.05 μmol/L | NA       | 15.7×h/10.32<sup>0.85</sup> | 34.9            | 33.4%          | GOF, Bootstrap, VPC |
|                          |                                 |                                 |                |                  |          | V₁/F                     | 79.2×(WT/69.3)<sup>0.29</sup> | / |                |            |
|                          |                                 |                                 |                |                  |          | Q₁/F                     | 0.97            |                |            |
|                          |                                 |                                 |                |                  |          | V₂/F                     | 51.4            |                |            |
| Mei et al. (2018)        | China (Single)                  | 701/98 (53/45)                 | 0.9–5.4 g/m², 1.3–8.2 h, i.v. | C₂₀/C₄₀/C₇₀/C₆₀ | HPLC     | 6.67×(SCR/68.1)⁻⁰·₄⁶×(BSA/1.75)<sup>1.1⁷</sup> | 40.0            | 3.02 μmol/L | GOF, Bootstrap, VPC |
|                          |                                 |                                 |                |                  |          | V₁/F                     | 24.46×(AGE/57.16)<sup>0.8¹</sup> | 42.7 |        |            |
|                          |                                 |                                 |                |                  |          | Q₁/F                     | 0.047           |                |            |
|                          |                                 |                                 |                |                  |          | V₂/F                     | 1.32            |                |            |
| Yang et al. (2020)       | China (Single)                  | 852/91 (64/27)                 | 1–3 g/m², 24 h, i.v. | NA               | HPLC     | 6.03×(CRCL/115.1)<sup>0.4⁴¹</sup> | 51.6            | 0.32 μmol/L | GOF, Bootstrap, VPC |
|                          |                                 |                                 |                |                  |          | V₁/F                     | 20.7            |                |            |
|                          |                                 |                                 |                |                  |          | Q₁/F                     | 0.074×(BSA/1.65)<sup>0.43</sup> | 65.6 |        |            |
|                          |                                 |                                 |                |                  |          | V₂/F                     | 3.76            |                |            |
| Gallais et al. (2020)    | France (Single)                 | 1,179/328 (180/133)<sup>2</sup> | 1–8 g/m², 0.5–36 h, i.v. | C₂₀/C₄₀, then q24 h until <0.2 μmol/L | HPLC     | 8.3×(AGE/50)<sup>0.3¹⁷⁷</sup> | 23.0            | 34.0%          | GOF, Bootstrap, VPC |
|                          |                                 |                                 |                |                  |          | V₁/F                     | 27.4            |                |            |
|                          |                                 |                                 |                |                  |          | Q₁/F                     | 0.15 (fixed) |                |            |
|                          |                                 |                                 |                |                  |          | V₂/F                     | 3.1×(WT/70.4)<sup>0.4⁰³</sup> | 38.0 |        |            |

<sup>a</sup>Two supplementary samples: at the end of infusion and between 8 and 12 h from the beginning of the infusion. Eventual follow up plasma level determination was done at 72 h, 96 h or more.

<sup>b</sup>Correlation is CL ~ VC, 0.78.
for the residual variability from the results of the OFVs and the distribution of residuals in the diagnostic plots. The parameter estimates and associated precisions of the base model are presented in Table 3.

Mechanistic plausibility was mainly considered as a potential covariate incorporated into the base model. As MTX is primarily eliminated unchanged by renal excretion, the effect of CrCl on MTX PK CL/F was tested first. The OFV substantially declined when CrCl was included exponentially (ΔOFV = 97.6, \( p < 0.001 \)), indicating a significant improvement in the model. A decrease (18.0%) in CL/F was observed as the CrCL decreased from 90 ml min\(^{-1}\) to 60 ml min\(^{-1}\).

The pathophysiological factors influencing MTX protein binding were then investigated empirically. ALB and HCT were included in the model to assess which was more suitable for describing the change in protein binding in the MTX PK process. The condition number of the final model was 83.6. Shrinkage analysis for CL showed a mean shrinkage of 24.9% and ε-shrinkage of 24.9%.

Moreover, the influence of morphological characteristics, such as body weight, BSA, and body mass index (BMI), on PK disposition parameters was tested based on allometric scaling theory (West et al., 1997; Anderson and Holford, 2009). However, no significant differences were observed. The presence of DIS \( \geq 2 \) was added to the CL/F to test the influence of drug-drug interaction factors on the MTX PK process. However, the drop in the OFV was 3.6 and no significant differences were observed. The step-by-step covariate screening procedure is shown in Supplementary Table S5 and the correlation between BSVs is presented in Supplementary Figure S1.

The additional estimation of the IOV for CL, which was estimated to be 24.7%, significantly improved model predictions (ΔOFV = 679.9, \( p < 0.001 \)), which suggests that elimination parameters vary across MTX courses.

The parameters of the final model are listed in Table 3. The final model with CL/F covariates was described as follows:

\[
\frac{CL}{F} = 4.91 \times (\text{CrCL} / 98)^{0.49} \times (\text{ALB} / 40)^{0.85} \times (0.89, \text{if age } > 60)
\]

In the final model, all retained covariates caused a significant increase in OFV upon removal. The condition number of the final model was 83.6. Shrinkage analysis for CL showed a mean η-CL shrinkage of 25.5% and ε-shrinkage of 24.9%.

### 3.3 Model Evaluation

The predictive performance of published models was evaluated in the evaluation population dataset using the prediction- and...

---

**TABLE 3** Parameter estimates for the base model, final model and bootstrap procedure.

| Parameters | Base model | Final model | Bootstrap of final model |
|------------|------------|-------------|-------------------------|
| Objective function value | 4306.6 | 3455.9 | / | / |
| Cl/F (L h\(^{-1}\)) | 4.8 | 4.91 | 3.7 | / |
| Vc/F (L) | 20.9 | 18.4 | 3.8 | / |
| Vp/F (L) | 0.09 | 0.063 | 9.8 | 0.073 |
| Covariate effect on Cl/F | | 2.14 | 14.1 | / | 2.17, 1.59–2.77 |
| CrCl | / | 0.49 | 22.6 | / | 0.50, 0.29–0.69 |
| ALB | / | 0.35 | 50.6 | / | 0.35, 0.031–0.72 |
| Age | / | 0.89 | 9.4 | / | 0.90, 0.72–1.06 |
| Between-subject variability | | | | |
| Cl/F (%) | 39.2 | 20.9 | 25.5 | 25.5, 20.3 |
| Vc/F (%) | 36.9 | 19.6 | 26.9 | 20.9 |
| Vp/F (%) | 62.3 | 40.6 | 18.1 | 19.4 |
| V/F (%) | 44.2 | 30.4 | 17.5 | 47.7 |
| Inter-occasion variability | | | | |
| IOV on Cl | / | 24.7 | 20.4 | 43.8 | 24.2, 5.6–35.4 |
| Residual variability | | | | |
| Proportional (%) | 58.3 | 40.1 | 6.3 | 11.5 |

### TABLE 4 | Results of the prediction-based metrics.

| Models | MDPE | MAPE | F\(_{20}\) | F\(_{30}\) |
|--------|------|------|-----------|-----------|
| Published models | | | | |
| Faltas et al. (2006) | –35.4 | 47.7 | 18.5 | 27.4 |
| Min et al. (2009) | –29.9 | 43.0 | 20.4 | 31.7 |
| Simon et al. (2013) | –53.1 | 54.2 | 15.8 | 25.1 |
| Bretagne et al. (2014) | –27.9 | 40.2 | 23.5 | 36.5 |
| Nader et al. (2017) | –68.4 | 70.6 | 11.0 | 17.0 |
| Mei et al. (2018) | –32.7 | 50.5 | 19.2 | 29.9 |
| Yang et al. (2020) | –28.8 | 37.0 | 26.3 | 41.3 |
| Gallais et al. (2020) | –45.7 | 47.6 | 16.0 | 28.3 |
| Newly built model | Final model | –10.2 | 36.4 | 29.7 | 42.3 |

F\(_{20}\), the percentages of prediction errors within 20%; F\(_{30}\), the percentages of prediction errors within 30%; MAPE, median absolute prediction error; MDPE, median prediction error.
simulation-based diagnostics described above. Supplementary Figure S2 shows the goodness-of-fit plots for all published models. The values of bias and inaccuracies of the published models are presented in Table 4. All models presented an underestimation of concentrations with large inaccuracies. The predictive performance of the predicted HD-MTX concentration-time profiles, as revealed by the pcVPC of the published models, was highly variable (Supplementary Figure S3). None of the models presented acceptable bias or inaccuracy values.

The goodness-of-fit plots for the newly built model are presented in Figure 1 showing no structural bias. The pcVPCs of the final model are depicted in Figure 2. The data below the LOQ were included in the model evaluation to assess the capability of the final model to describe these points. The simulated data corresponded well with the observed data, except for data below the LOQ around 24 h and more than 100 h after the dose. However, as only 18 samples and 10 samples below the LOQ in these periods were included in our study, there were no significant model misspecifications. Model stability was also confirmed with consistent bootstrapped parameter estimates differing by no more than 15% from its corresponding estimate in the final model (Table 3). The MDPE and MAPE were -10.2% and 36.4%, respectively. The relatively low values of MDPE and MAPE further confirmed the high prediction accuracy of the final model.

3.4 Model Application

The characteristics of the involved covariates in the simulation are listed in Supplementary Table S6. The results of the Monte Carlo simulation are presented in Supplementary Table S7. The predicted time course of MTX concentration in the scenarios with median covariate levels is presented in Figure 3. Based on the simulation, elderly patients with renal dysfunction and hypoalbuminemia have a higher incidence of delayed MTX elimination. The percentage decreased from 64.7% to 27.6% when the CrCL decreased from 98 ml min$^{-1}$ to 46.3 ml min$^{-1}$, and the percentage decreased from 64.7% to 54.8% when ALB decreased from 39 g L$^{-1}$ to 29 g L$^{-1}$.

4 DISCUSSION

To the best of our knowledge, this is the first comprehensive analysis of the predictability of published HD-MTX popPK models. Prediction-based metrics and simulation-based diagnostics were applied to assess the accuracy and precision.
of the published models. According to our results, the published models were insufficient for cross-center use. Thus, constructing a new popPK model based on an independent dataset is a priority.

Whether the assumed description of the popPK model is in agreement with the population it is intended to treat, depends on two aspects: a well-developed model with reliable a priori PK parameter distribution (mixed and random) and center-related factors that may result in unexplained inter-study variability (Laporte-Simitsidis et al., 2000). The center-related factors, such as study design, ethnic differences, assay methods, and modeling strategies, may influence model predictability (Mao et al., 2018; Mao et al., 2020). In this study, the inclusion of older patients with primary central nervous system lymphoma, as well as inconsistencies in the assay methods, may be the primary causes of the poor cross-center predictability observed.

Although the published models were inadequate for cross-center use, the information included in these models can be utilized to guide the building of new models. In previous studies, we found that theory-based modeling is helpful to improve model predictability (Mao et al., 2020; Mao et al., 2021). Unlike empirical stepwise covariate selection, theory-based covariate selection allows the incorporation of relationships linking parameters and covariates based on a fundamental understanding of PK processes rather than only on the available data, and may improve model predictability (Danhof et al., 2008). Therefore, we conducted modeling based on the combination of theoretical mechanisms and data properties in this study.

According to published studies, CrCL was the most frequently identified covariate. This was consistent with the PK characteristics of MTX. MTX and its major metabolite, 7-OH-MTX, are mainly eliminated by glomerular filtration and active secretion (Shitara et al., 2006). After analyzing HD-MTX serum samples and urine samples, it was revealed that MTX renal clearance is 83% of the total clearance and CrCL, as an index of glomerular filtration rate, has a great effect on renal clearance but not on non-renal clearance (Fukuhara et al., 2008). In the present study, as no urine samples were included, the influence of CrCL was added to the total clearance. As the CrCL doubling decreased from 120 ml min$^{-1}$ to 60 ml min$^{-1}$, the CL/F decreased by 28.8%, which is consistent with the results obtained using published models (17.1%–50%), and in patients with CrCL increase from 120 ml min$^{-1}$–180 ml min$^{-1}$, the CL/F increased by 22.0%, which is close to the results obtained using published models (11.6%–19.6%) (Simon et al., 2013; Benz-de Bretagne et al., 2014; Yang et al., 2020).

Alterations in HCT or hemoglobin may affect MTX binding and thus influence the PK process. Hypoalbuminemia is reportedly associated with a significantly increased time for MTX clearance (Reiss et al., 2016). Therefore, the reduction of ALB is a risk factor for delayed MTX elimination in HD-MTX
monotherapy (Kataoka et al., 2021). In the final model, a reduction in ALB from 50.0 g L$^{-1}$ to 24.0 g L$^{-1}$ may result in a 22.7% decrease in MTX CL/F. The change in MTX CL/F may have clinical effects on MTX exposure and treatment response. Moreover, malignant cachexia or liver metastases induced by hypoalbuminemia may increase the half-life of MTX PK, which may be associated with unanticipated toxicity (Li and Gwilt, 2002).

FIGURE 3 | Simulated time-concentration profiles of MTX under conditions involving different covariate levels. (A,B) For patients ≤60 years old; (C,D) for patients >60 years old.
The CL/F was decreased by 11.0% in elderly patients, which may also be a risk factor for delayed elimination. The incidence of delayed clearance of MTX is higher in older patients and during the first cycle of treatment (Bacci et al., 2003). After comparing the PK of total and free MTX in rheumatoid arthritis patients, it has been shown that elderly patients have a longer elimination half-life of free and total MTX (Bressolle et al., 1997). For elderly patients, altered metabolic functions are a result of the complex processes of aging-related physiological changes in the functional reserve of multiple systems and organs. Furthermore, renal excretion is reduced (up to 50%) in approximately two-thirds of elderly patients, which can potentially cause the delayed elimination of MTX (Klotz, 2009).

The estimated PK parameters in our final model were consistent with those in previous reports, except that the CL/F was lower than that in published studies. This phenomenon may be owing to the influence of drug-drug interactions. In our study, the proportion of samples with DIS ≥2 was more than 85.4%. Co-administration, such as that of non-steroidal anti-inflammatory drugs, β-lactams, and proton pump inhibitors, has been reported to be an important factor that influences MTX elimination (Iven and Brach, 1990; Suzuki et al., 2009; Kawase et al., 2016), and may inhibit the excretion of MTX. A reduction in CL/F by 0.929, as at least one score 2 drug was used (Benz-de Bretagne et al., 2014).

Previous studies have demonstrated that variability exists between MTX therapy cycles (Fukuhara et al., 2008; Min et al., 2009; Kim et al., 2012; Mei et al., 2018). In this study, we estimated course-to-course variability with the inclusion of IOV terms in the random effects model. IOV may arise from variable renal function owing to repeated chemotherapy, drug-drug interactions, or other environmental factors (Breedveld et al., 2004; El-Sheikh et al., 2007; Leveque et al., 2011).

No gene information was available for analysis in this study. The influence of polymorphisms in genes encoding transporters or enzymes during the MTX PK process was inconsistent between different studies, which may partly be accounted for by the different types of patients. The sample size did not have sufficient power to detect a significant association between the target single nucleotide polymorphisms (SNPs) and the MTX PK properties, and the different frequencies of SNPs owing to racial differences (Simon et al., 2013; Lui et al., 2018; Yang et al., 2020). We will consider this point in a future study.

In relative terms, body size and composition were less influential in the model of MTX CL (Pai et al., 2020). A statistically significant reduction in the IIV of CL in only 15% of anticancer drugs is associated with BSA. Moreover, the relative reduction in the variability of CL is between 15% and 35% (Baker et al., 2002). Among the eight published models, only one model included the influence of body size on CL. It is also worth mentioning that body size does not take into account abnormal body habitus such as cachexia or morbid obesity (Gurney and Shaw, 2007; Gallais et al., 2020). In our study, only two patients had a BMI ≥30. Therefore, body size was not a risk factor identified in this study.

Certain limitations were noted in our study. First, this study relied on routine MTX monitoring data that was retrospectively collected and did not include information regarding aspects of supportive care, such as fluid hydration and urine pH, which may have impacted MTX CL and influenced model predictability. Second, MTX concentrations were measured using multiple analytical methods between studies. No conversion of immunoassays may introduce uncertainty into the external evaluation results. Third, no information on MTX metabolites was available for this study. The inclusion of MTX metabolite information in the popPK model could provide more insights into MTX disposal. Therefore, this should be the focus of future studies.

Identifying ways to obviate MTX delays could facilitate toxicity prediction, rational dose adjustments, and theoretically improve treatment outcomes (Crews et al., 2004). According to our results, the currently published models are not sufficiently reliable for cross-center use; the newly built model provides better predictability. The elderly patients and those with renal dysfunction, hypoalbuminemia are at higher risk of delayed elimination.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Clinical Research Ethics Committee of Huashan Hospital, Fudan University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JM, BC, and WQ participated in the study design; JM, QL, PL, WQ, BC, and MZ implemented and conducted the study; JM, QL, and WQ performed the study and analyzed the data. JM drafted the manuscript, which was revised and approved by all authors.

FUNDING

This work was supported in part by a grant from the Shanghai Shenhkang Clinical Innovation Project (No. SHDC12020112) and Shanghai “Rising Stars of Medical Talent” Youth Development Program (No.2019-72) supported by Shanghai Municipal Health Commission.

ACKNOWLEDGMENTS

We would like to thank Editage (www.editage.cn/) for English language editing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2022.817673/full#supplementary-material
