Comparing basic static models for predicting clinical CYP3A4 induction risk

Rachel Upcott Gill, Katie Plant, Lara Ward, Catherine Pearson

Abstract

Induction of cytochrome P450 enzymes is associated with an increased risk of clinical drug-drug interactions (DDIs), particularly those interactions involving CYP3A. The risk of a clinical DDI presents both a safety and efficacy concern, particularly when considering increasing polypharmacy in an aging population. As clinical DDI studies are costly and time consuming, it is important to ensure predictions of *in vivo* induction potential from *in vitro* data are as accurate as possible. In addition to fold change and other static models for risk assessment, correlation methods such as calculation of the relative induction score (RIS) can be used to predict the magnitude of CYP3A4 clinical induction risk from *in vitro* data. This involves calibrating the RIS values for a set of known inducers against their clinically measured induction (AUC ratio) within each hepatocyte donor. Cyprotex has generated RIS data sets for multiple hepatocyte donors.
To understand the accuracy of RIS predictions compared to other basic methods including R_3, six additional test compounds with clinically available CYP3A4 induction data were assessed, and models evaluated on number of false negatives and false positives and correlation of predicted AUCR with observed AUCR for quantitative prediction of clinical induction risk.

Methods assessed

Basic equations	R_3 (d value =1)
	R_3 (d value corrected)
Correlation methods	Relative Induction Score (RIS)
	$I_{max.\,u}/EC_{50}$

Table 2: Methods used to assess prediction of clinical induction risk

Compound	Clinical DDI observed	TN/TP
Bosentan	Moderate inducer	TP
Dexamethasone	No Induction	TN
Efavirenz	Moderate inducer	TP
Rifaximin	No Induction	TN
Tasimelteon	No Induction	TN
Teriflunomide	No Induction	TN

Table 1: Clinical categories of compounds selected for analysis [1,2,3,4,5,6]. TP = True Positive, TN = True Negative

Introduction

Assessing clinical induction risk remains a vital part of the drug discovery process in order to assess the possibility of DDIs with co-administered drugs. Extensive *in vitro* CYP induction experiments are required as part of the regulatory guidelines, and as a result, maximising their efficiency and predictability is desired. Cyprotex's CYP induction assay is designed to meet regulatory guidance, and enable clinical risk prediction utilising basic methods such as R_3 calculations, alongside validated RIS calibration curves. In order to assess the accuracy of these methods, six compounds were selected, focusing on compounds that were either clinical non inducers (defined as a clinical AUCR of >0.8) or those demonstrating moderate induction (clinical AUCR of 0.2-0.5) *in vivo* (Table 1). To further assess methods of risk prediction, R_3 values were calculated both without a scaling factor (d=1) and calculating this value using the observed and predicted magnitude of rifampicin induction in each experiment[7]. Whilst the recent M12 guidance[8] highlights a preference for using a set of known inducers to provide the scaling factor, our approach aligns with previous publications[7] and demonstrates an example of an approach that can be utilised to incorporate data from each individual experiment.
To assess performance of correlation methods, in addition to the validated RIS curve, calibration curves were generated using $I_{max,u}/EC_{50}$ values, using the same set of known inducers as the RIS analysis. All methods were assessed using both mRNA and activity end points.

Methods

CYP induction methods followed validated protocols and regulatory guidelines. Assessment of induction using mRNA analysis was performed via qRT-PCR on an Biosystems QuantStudio™ 7 Real Time PCR system. Assessment of induction utilising a catalytic activity end point used midazolam as a CYP3A4 substrate and LC-MS/MS methods for detection of metabolite. All experiments utilised cryopreserved hepatocytes. All compounds were analysed in three separate hepatocyte donors each of which had been previously validated for RIS analysis.

Data analysis

To calculate fold induction using the activity end point, 1-hydroxymidazolam concentration for each test compound replicate was calculated and compared to vehicle control wells. For mRNA assessment, relative fold mRNA expression was determined based on the threshold cycle (C_T) data of target gene relative to endogenous control for each reaction, and normalised to vehicle control using the $2^{-\Delta\Delta CT}$ method[9]. Where appropriate, EC_{50} and E_{max} values were determined from nonlinear regression analysis using the equation below:

$$\text{Fold increase} = E_{min} + \frac{E_{max} - E_{min}}{\left(1 + \left(\frac{EC_{50}}{X}\right)^{slope}\right)}$$

For further analysis, R_3 values were calculated using the equation below:

$$R_3 = {}^{1}\Big/\left[1 + \left(\frac{(d \times E_{max} \times 10^{*} \times I_{max,u})}{(EC_{50} + 10^{*} \times I_{max,u})}\right)\right]$$

*safety factor 10, 5, 2 and 0 used for further calculations

To calculate the d value, the following equation was used:

$$d = \left(\frac{1}{f_m\left(\frac{E_{max}[I]}{EC_{50}+[I]}\right)}\right)\left(\frac{1}{AUCR} - 1\right)$$

To calculate RIS values for each compound, the following equation was used:

$$RIS = \frac{(E_{max} \times I_{max,u})}{(EC_{50} + I_{max,u})} \qquad \text{where } I_{max,u} = I_{max} \times f_u$$

Using previously validated RIS calibration curves, RIS values were plotted and % decrease in AUC change determined. >20% decrease in AUC was used as a cut-off for potential clinical induction.

To assess the performance of each model, the percentage of False Negatives (FN), False Positives (FP) and the Geometric Mean Fold Error (GMFE) were calculated for the median and worst case donor. Median and worst case donors were selected based on E_{max} and EC_{50} values across both end points. The equation used to calculate the GMFE is as follows:

$$GMFE = 10^{Mean(|\log\frac{predictedDDI}{ObservedDDI}|)}$$

			Correlation methods		Basic calculations - R3					
					R3 (d=1)			R3 (d=corrected)		
			RIS	I_{max}/EC_{50}	10 x	5 x	2 x	0 x	10 x	5 x	2 x	0 x
FN	mRNA	Worst case	0%	0%	0%	0%	0%	0%	0%	0%	0%	50%
		Median	0%	0%	0%	0%	0%	0%	0%	0%	0%	50%
	Activity	Worst case	0%	0%	0%	0%	0%	50%	0%	0%	0%	50%
		Median	0%	0%	0%	0%	0%	50%	0%	0%	0%	50%
FP	mRNA	Worst case	100%	75%	100%	75%	50%	25%	50%	50%	25%	25%
		Median	25%	75%	100%	25%	25%	25%	25%	25%	25%	0%
	Activity	Worst case	50%	75%	50%	50%	25%	25%	50%	50%	25%	0%
		Median	25%	75%	25%	25%	25%	25%	25%	25%	25%	0%
GMFE	mRNA	Worst case	1.23	1.50	2.20	1.63	1.25	1.26	1.20	1.20	1.29	1.35
		Median	1.21	1.47	1.66	1.37	1.13	1.23	1.30	1.24	1.34	1.40
	Activity	Worst case	1.33	1.59	1.62	1.37	1.26	1.26	1.40	1.28	1.21	1.26
		Median	1.24	1.64	1.30	1.21	1.25	1.32	1.19	1.18	1.30	1.37

Table 3: FN (%), FP (%) and GMFE values calculated using mRNA and catalytic activity end points for the selected methods of induction risk prediction.

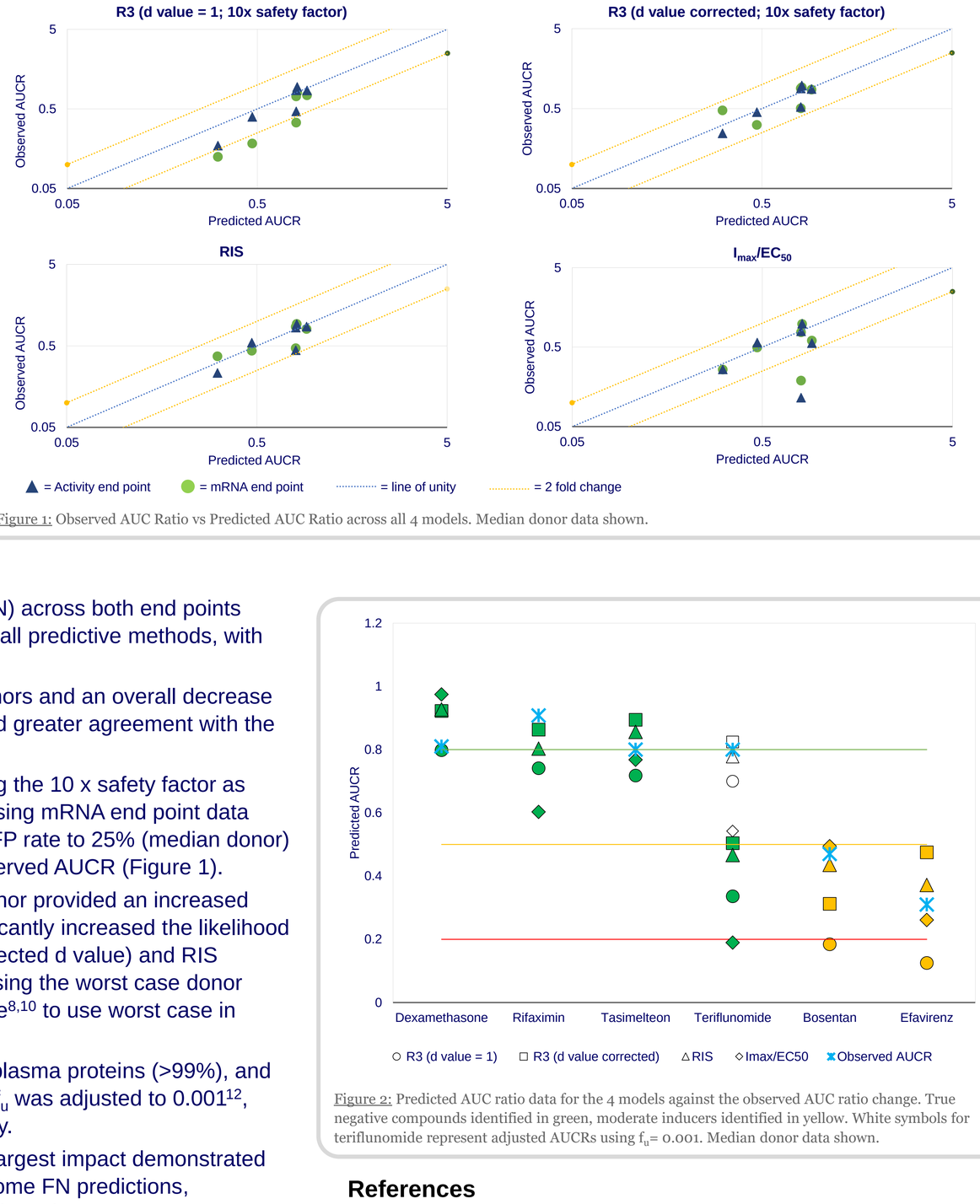

Figure 1: Observed AUC Ratio vs Predicted AUC Ratio across all 4 models. Median donor data shown.

Results

- Both correlation methods and R_3 calculations (using 10 x safety factor[10]) demonstrated no false negatives (FN) across both end points and using median or worst case donors (Table 3). The false positive (FP) rate was significantly higher across all predictive methods, with R_3 (d=1) and the $I_{max,u}/EC_{50}$ correlation method demonstrating the highest rate of FPs.
- In comparison to $I_{max,u}/EC_{50}$, the RIS model demonstrated an improved GMFE across both endpoints and donors and an overall decrease in the FP rate. When comparing the observed and predicted AUCRs (Figure 1), the RIS method demonstrated greater agreement with the line of unity and all predicted AUCRs were within 2 fold of the observed values.
- Analysis of the R_3 equation with and without a corrected d value had a significant impact on predictions. Using the 10 x safety factor as recommended in the FDA guidance[10] and recent M12 guidance[8] the R_3 (d=1) resulted in a FP rate of 100% using mRNA end point data (Table 3). Using the calculated d value significantly increased the accuracy of the prediction, decreasing the FP rate to 25% (median donor) and improving the GMFE, in addition to improving predictions for all compounds to be within 2 fold of the observed AUCR (Figure 1).
- Comparison of the worst case and median donors across the four methods demonstrated that the median donor provided an increased accuracy in prediction (improved GMFE values and reduction in FPs, Table 3). Median donor data also significantly increased the likelihood of the correct category (no, low, moderate or high inducer) being predicted. When using mRNA data, R_3 (corrected d value) and RIS methods predicted the category correctly for 5 out of 6 compounds compared to 3 and 1 respectively when using the worst case donor (data not shown). This is a similar approach to literature[11] and while may improve prediction, regulatory advice[8,10] to use worst case in donor selection in order to proceed cautiously and limit possible false negatives is still relevant.
- Teriflunomide was incorrectly identified as a positive inducer in all methods. Teriflunomide is highly bound to plasma proteins (>99%), and as a result, $I_{max,u}$ was calculated using f_u=0.01 as recommended by regulatory guidance. However, when the f_u was adjusted to 0.001[12], predicted AUCR ratios increased (Figure 2) to 0.824 and 0.777 for R_3 (corrected d value) and RIS respectively.
- Alteration of the safety factor used in R_3 (d value corrected) calculations decreased the rate of FPs, with the largest impact demonstrated without a safety factor included. No FPs were observed when utilizing mRNA data, however this also led to some FN predictions, highlighting the necessity to include a safety factor in order to prevent under prediction of induction risk. The data shown suggests a 2x safety factor may provide a suitable compromise (no FNs but a decrease in FP rate) agreeing with other recent literature findings[13].

Figure 2: Predicted AUC ratio data for the 4 models against the observed AUC ratio change. True negative compounds identified in green, moderate inducers identified in yellow. White symbols for teriflunomide represent adjusted AUCRs using f_u= 0.001. Median donor data shown.

Conclusion

- Median donor provided an improved prediction across multiple methods compared the worst case donor, regardless of endpoint.
- Assessment of correlation methods demonstrates that RIS provides a more accurate DDI prediction relative to the $I_{max,u}/EC_{50}$ method.
- Restriction of fu to 0.01 can contribute to a significant over prediction in clinical induction risk for highly bound compounds.
- Calculating d values for R_3 calculations significantly improves prediction of AUCR, and whilst using a 10 x safety factor may lead to over prediction of clinical implications it does ensure a reduced risk of false negatives.

References

1) Bosentan product label, EMA
2) Villikka, Kivisto & Neuvonen, (1998) *Pharmacol Toxicol* 83(3):135-8
3) Efavirenz product label, EMC
4) Terilfunomide product label, NDA 202992
5) Rifaximin product label, FDA 4196955
6) Tasimelteon product label, Reference ID: 3627825
7) Vermet et al., (2016) DMD 44:50–60
8) ICH; Drug Interaction Studies M12 (2022)
9) Livak KJ and Schmittgen TD., (2001) Methods. 25, 402-408.
10) FDA: *In Vitro* Drug Interaction Studies (2020)
11) Kenny et al., (2018), DMD 46,1285-1303
12) Rakhila et al, (2011) J Pharm Biomed, 15;55(2):325-31
13) Ramsden & Fullenwider., (2022) Eur J Drug Metab Pharmacokinet, 47(4):467-482

Corporate Headquarters: Evotec SE, Manfred Eigen Campus, Essener Bogen 7, 22419 Hamburg, Germany **Email:** enquiries@cyprotex.com **Web:** www.cyprotex.com www.evotec.com