

OECD TG 445A: human hepatic CYP induction enters standardised chemical safety testing

The 2026 OECD guideline introduces a functional CYP induction assay using differentiated human hepatic cells, supported by new validation evidence and explicit limitations.

PharmaToxAI · 2026-10-19

A new human-relevant metabolic endpoint

On 2 July 2026, OECD adopted TG 445A, a standardised method for measuring induction of CYP1A2, CYP3A4 and CYP2B6 using differentiated human hepatic cells. CYP2B6 is associated with greater uncertainty than the other two enzymes. The test can support mechanistic interpretation of chemical disposition but does not independently establish toxicity or a clinical drug interaction [1].

Induction can reduce exposure to a parent compound while increasing the formation of metabolites. CYP1A2 is associated primarily with AhR, CYP3A4 with PXR and CAR, and CYP2B6 particularly with CAR. These pathways overlap and do not imply a unique receptor mechanism.

The method

The validated workflow is based on differentiated cryopreserved HepaRG cells. At least six concentrations are tested during a 48-hour induction exposure, with medium renewal at approximately 24 hours. Phenacetin, bupropion and midazolam serve as probe substrates. LC-MS measures acetaminophen, hydroxybupropion and 1'-hydroxymidazolam. Product formation is normalised to cellular protein [1].

Three independent experiments with different cell batches are required, along with appropriate controls and replicate wells. A positive within-run response requires at least two consecutive concentrations producing two-fold induction and a plausible concentration-response pattern. Borderline responses have separate interpretation rules.

Validation and uncertainty

An EURL ECVAM multicentre validation study found HepaRG more suitable for routine testing than primary human hepatocytes [2]. The guideline reports original pharmaceutical-set interlaboratory reproducibility of 80% for CYP1A2 and CYP3A4, versus 66% for CYP2B6. These are reproducibility statistics, not estimates of predictive sensitivity or specificity [1].

Two linked 2026 studies expanded the evaluation to industrial and agricultural chemicals. Reproducibility was reasonable but weaker for some non-pharmaceutical compounds, especially borderline inducers. mRNA measurements helped interpret activity but were not validated as replacements for the functional endpoint [3,4]. A separate study found enzyme-specific induction patterns among six candidate metabolism-disrupting chemicals [5]; this does not establish adverse outcomes in humans.

Applicability

Only soluble, non-cytotoxic concentrations can be interpreted. The method uses a starting concentration associated with at least 90% fractional survival. DMSO is generally limited to 0.1% because solvent can affect induction. Nonspecific binding and differences in protein binding between serum-free culture and human plasma complicate exposure extrapolation. The validation set consisted of defined single-constituent chemicals; applicability to mixtures remains unestablished [1].

TG 445A does not validate CYP inhibition, Phase II metabolism, metabolite identification or intrinsic hepatic clearance. Other hepatic systems are not automatically validated without the prescribed proficiency testing.

Integration with computational models

The concentration producing a two-fold induction response (F2), when estimable, may inform kinetic modelling only with additional exposure and mechanistic information. A clinical drug interaction model also requires enzyme turnover, unbound exposure, pathway contributions and other parameters. ICH M12 provides a separate regulatory framework for pharmaceutical drug interactions [6].

A robust evidence record should retain chemical identity, concentrations, cell batch, cytotoxicity, raw metabolite measurements, normalisation, controls and interpretation rules. These details determine whether the data can support quantitative in vitro-in vivo extrapolation.

How the assay's decision procedure should be read

The TG distinguishes an experimental response from a conclusion across runs. The two-fold threshold is not based on a hypothesis test alone; it is a pragmatic criterion developed against human-relevant reference information. A single elevated well or isolated high-concentration response is insufficient for an unqualified positive conclusion. The interpretation procedure accounts for whether responses are consecutive, plausible and replicated across independent cell batches.

This matters when results are converted into datasets. Treating all responses above a numeric threshold as identical binary labels discards borderline and equivocal calls, uncertainty and the experimental context that produced them. For modelling, the underlying concentration-response and quality-control information are more informative than a single label.

The original validation involved a limited number of chemicals selected partly because human induction information was available. OECD TG 445A also reports the augmentation-study reproducibility separately. These data should not be pooled indiscriminately into a universal accuracy estimate, since the chemical selections, laboratory comparisons and endpoints differ [1-4].

Practical constraints of human hepatic models

HepaRG cells offer reproducible access to a metabolically competent human-derived system, but they are not interchangeable with primary human hepatocytes. The original validation compared both systems and found greater variability in the primary-hepatocyte arm. The adopted method therefore rests principally on HepaRG evidence, while permitting other systems only after they demonstrate correct classification of the prescribed proficiency chemicals [1,2].

Several design choices protect interpretability. Reference inducers must produce the expected response; basal CYP activity must be adequate; and each experimental series must include vehicle controls. Independent cell batches, rather than replicate wells alone, are needed to address batch-to-batch variation. A cell-viability assay is required to establish the non-cytotoxic concentration range before induction is evaluated.

These controls do not eliminate all sources of bias. Strongly lipophilic compounds can adsorb to plastics, volatile compounds can be lost during incubation, and unstable substances may change composition over the exposure period. Serum-free medium facilitates standardisation but also means that comparisons with total human plasma concentrations are misleading when binding is substantial. Measured or estimated unbound exposure should be considered when attempting in vitro-in vivo interpretation [1].

For a complex formulation, a positive result can reflect one constituent or an interaction between constituents. A negative result is equally difficult to interpret if solubility, instability or assay interference has prevented meaningful cellular exposure. The guideline notes that applicability to mixtures was not established during validation; compositional characterisation remains necessary before extrapolating from ingredient data to finished products.

Why a functional endpoint was prioritised

The relationship between CYP mRNA and activity is context-dependent. Transcriptional induction, enzyme synthesis, degradation, substrate competition and direct inhibition can occur on different timescales. A compound may therefore induce expression and simultaneously inhibit the measured catalytic reaction.

Quartermain and colleagues investigated mRNA alongside functional CYP measurements in the supplementary validation. The transcript data were useful as supporting evidence, but the authors did not validate mRNA as a substitute for the enzyme-activity decision procedure [4]. This distinction is relevant to transcriptomic NAM platforms: a pathway enrichment score or elevated CYP transcript can motivate further testing, but it should not be reported as a validated TG 445A induction result.

Conversely, the activity assay does not establish which transcription factor was activated. Receptor assays, time-course experiments and targeted perturbations may be needed when the mechanistic question concerns PXR, CAR or AhR specifically.

Regulatory scope and modelling implications

An OECD Test Guideline standardises the performance and interpretation of a specified test. It does not by itself confer universal acceptance of every downstream application. For chemical safety, TG 445A may provide an evidence stream for integrated approaches to testing and assessment. For pharmaceutical development, ICH M12 addresses the broader question of clinically relevant drug-drug interactions [6].

A PBPK model can incorporate enzyme induction only when the required quantitative parameters and assumptions are justified. The F2 concentration is potentially useful as a point on an induction concentration-response curve, but it does not describe the entire response, enzyme turnover or the relationship between an in vitro nominal concentration and a tissue unbound concentration. Modelling without these distinctions risks giving a precise-looking prediction unsupported by the data.

The most defensible integrated workflow preserves four layers: **experimental measurements, assay interpretation, exposure and kinetic modelling**, and **the final context-specific risk conclusion**. Each layer has its own uncertainties.

Agreement between a predicted clinical outcome and an observed one, when available, should be evaluated separately from the reproducibility of the underlying cell assay.

Conclusions

TG 445A establishes a human-relevant functional endpoint with a reproducible decision procedure. Its contribution is the quality and traceability of CYP induction evidence, not a standalone prediction of human risk. CYP2B6 uncertainty, borderline responses and chemical-domain limitations remain important research priorities.

References

1. OECD (2026). *Test No. 445A: Determination of Cytochrome P450 (CYP) Induction using Differentiated Human Hepatic Cells*. DOI.
2. Bernasconi C et al. (2019). Validation of in vitro methods for human cytochrome P450 enzyme induction: Outcome of a multi-laboratory study. *Toxicology in Vitro* 60:212–228. DOI.
3. Jacobs MN et al. (2026). HepaRG CYP induction test method augmentation part 1: The GOLIATH Study. *ALTEX* 43:680–695. DOI.
4. Quartermain E et al. (2026). HepaRG CYP induction test method augmentation part 2: Study supported by mRNA analysis. *ALTEX* 43:696–710. DOI.
5. Person E et al. (2026). Induction of human cytochrome P450 enzyme activities by metabolism disrupting chemicals in the hepatic cell line HepaRG. *Toxicology* 519:154300. DOI.
6. ICH (2024). *M12 Drug Interaction Studies*. FDA guidance.

Online edition: <https://pharmatoxai.ar/whitepapers/oecd-tg-445a-human-hepatic-cyp-induction-2026/>