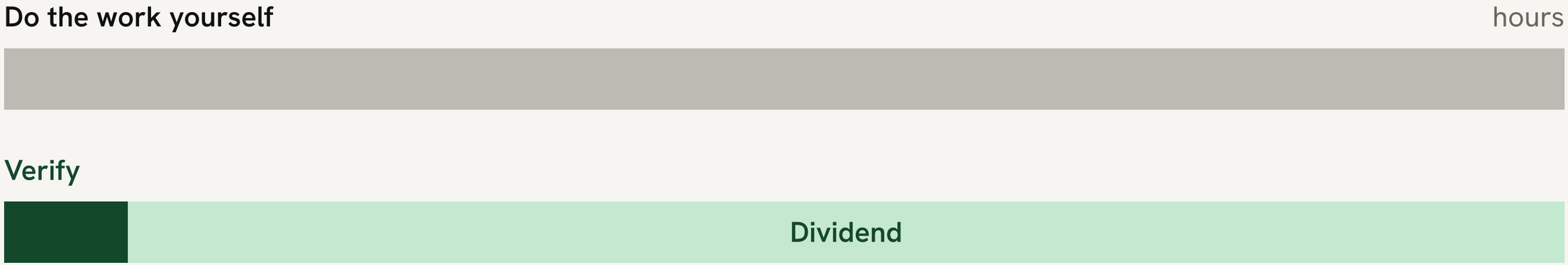
The Verification Dividend

Mattson Thieme, PhD
Co-Founder, Human Chemical Co.

JIFSAN-CFS^3 Annual Symposium · October 22, 2026

An AI answer is worth what it saves you after you check it.

If checking costs as much as doing, the dividend is zero.



Everything depends on three questions.

1 What is the claim?

2 Why was it made?

3 Is it correct?

How fast you can answer these is the price of the tool.

The worked example

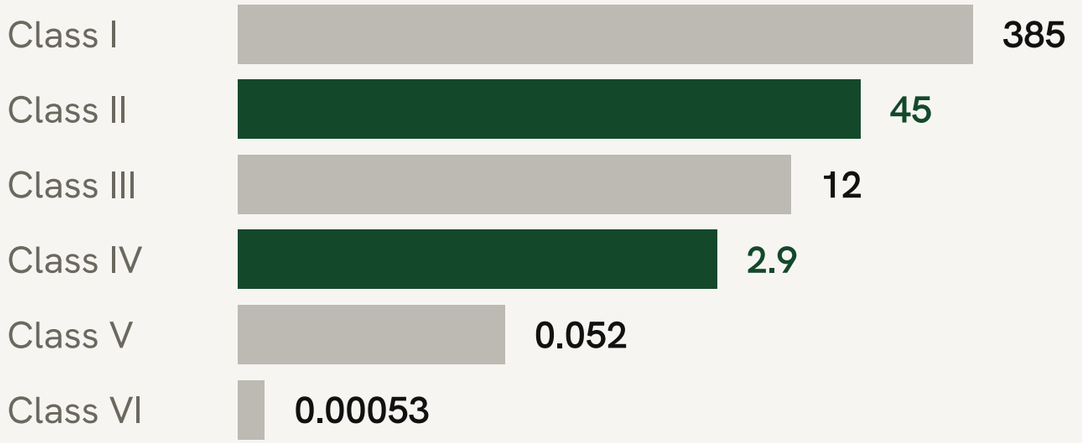
FDA's Expanded Decision Tree

A structure-based screen for chronic oral hazard. Each class carries a threshold of toxicological concern.

47 questions

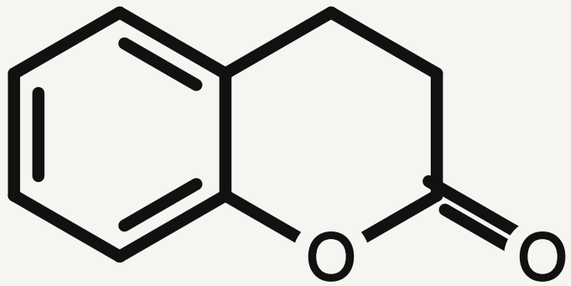
142 sub-questions

6 classes



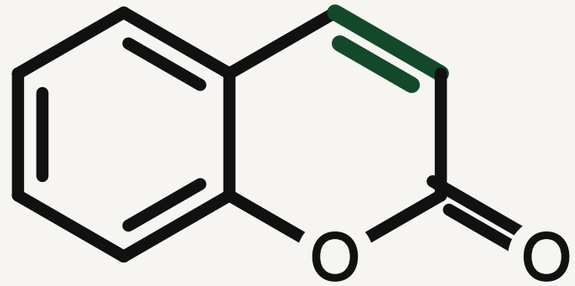
µg/kg bw/day · Log scale. >700,000× end to end.

Two molecules.
One double bond apart.



Dihydrocoumarin

O1C(=O)CCC2=CC=CC=C12



Coumarin

O1C(=O)C=CC2=CC=CC=C12

Both occur in food. Both are in FDA's EDT database.

The EDT puts them 15× apart.

Dihydrocoumarin

Class II

45 µg/kg bw/day

15×

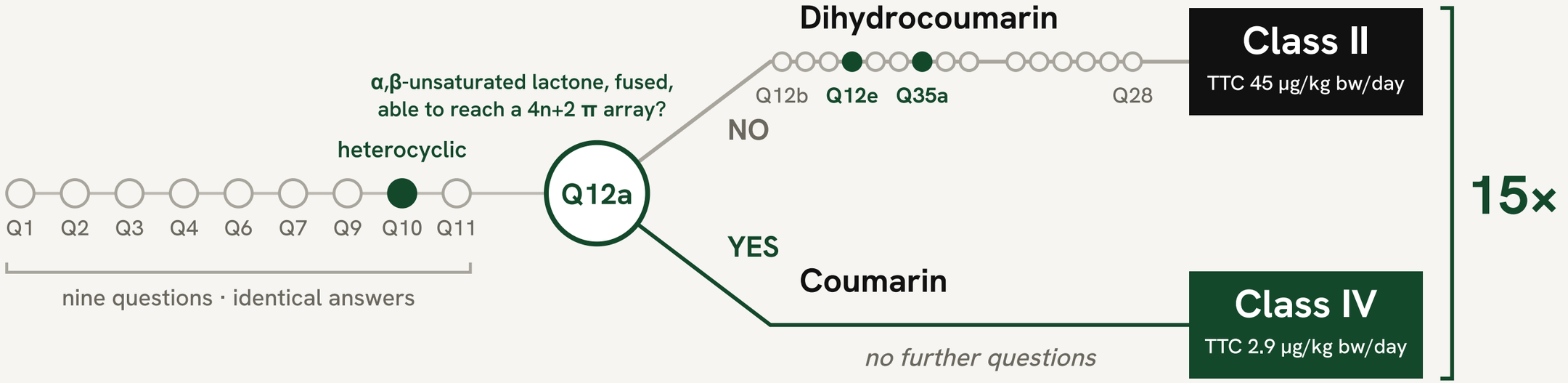
Coumarin

Class IV

2.9 µg/kg bw/day

Why?

Nine questions. Identical answers. Then one fork.



Nine top-level questions, answered identically, then one fork. Counting every sub-question: coumarin 79 checks, 2 answered yes; dihydrocoumarin 109. Both paths match those recorded in FDA's EDT database.

Now I can see why. That is not the same as knowing it's right.

What would it take to believe this?

- A dose level, from a real study, for each.
- Same species. Same lab, if possible.
- A reason the double bond would matter.

This is where agents can materially assist with the verification problem.

Dihydrocoumarin has a no-effect level. Coumarin doesn't.

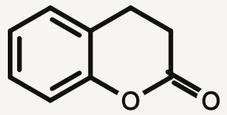
	Dihydrocoumarin	Coumarin
Study	NTP 1993, F344/N rat, gavage	NTP 1993, F344/N rat, gavage
Duration	91 days	2 years
No-effect level	**107 mg/kg bw/day**	**none established**
Lowest-effect level	**214 mg/kg bw/day**	**18 mg/kg bw/day**
Survival, high dose, males	—	**0 of 50**

For coumarin there was no dose that did nothing.

Values as recorded in FDA's post-validation EDT database (Appendix 2, October 2026), citing NTP 1993. Retrieved automatically from federal databases by an agent.

Same bond.

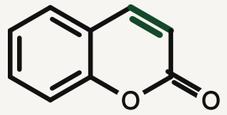
Q12a asks about it. The liver acts on it.



Dihydrocoumarin

ring opens by hydrolysis →

o-hydroxyphenyl-propionic acid → a normal metabolite



Coumarin

Q12a

3,4-epoxidation across the double bond →

reactive 3,4-epoxide → liver injury in rat

Remove the bond and the epoxide route is gone.

In rat. Humans favour a different route — more on that in a moment.

Claim. Reason. Evidence.

15× apart → Fork at Q12a → ✓ NOAEL 107 vs none epoxide route

Minutes, not an afternoon.

Every number above traces to a named study.

And one thing the screen can't settle.

Coumarin's Class IV reflects a rat bioactivation route. Humans mostly take a different one. A structure-based screen does not adjust for species.

Verification is also how you find the edge of what the tool knows.

The dividend is biggest in the middle.

Obvious

You already know. Verification is free, and so is the answer.

Checkable with work

An afternoon of reading. This is where the dividend lives.

Unknown

No amount of checking settles it.

▲

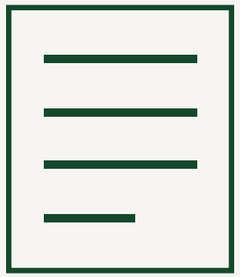
coumarin / dihydrocoumarin

Encode the logic.



Logic learned

The answer is a weight.
“Why” is a reconstruction.



Logic encoded

The answer is a path.
“Why” is a lookup.

This is not an argument against neural networks.
It is an argument about where the decision rule lives.

Then take it to the source.

- **Is there authoritative data on this exact question?**
- **How fast can it be pulled?**
- **Does it trace back to a named study, not a summary?**

A claim is cheap to check when it can be traced to a named study.

Three questions for any AI tool you're asked to trust

1 What is the claim?

2 Why was it made?

3 Is it correct?

How cheaply you can answer all three is what turns an AI tool into a useful one.

Mattson Thieme · mattson@humanchemical.com · humanchemical.com