SciPinion

SciPi #1052

October 19, 2026

Critical Review and Scoring of Animal Cancer Bioassays

PREPARED FOR:

Vinyl Acetate Council

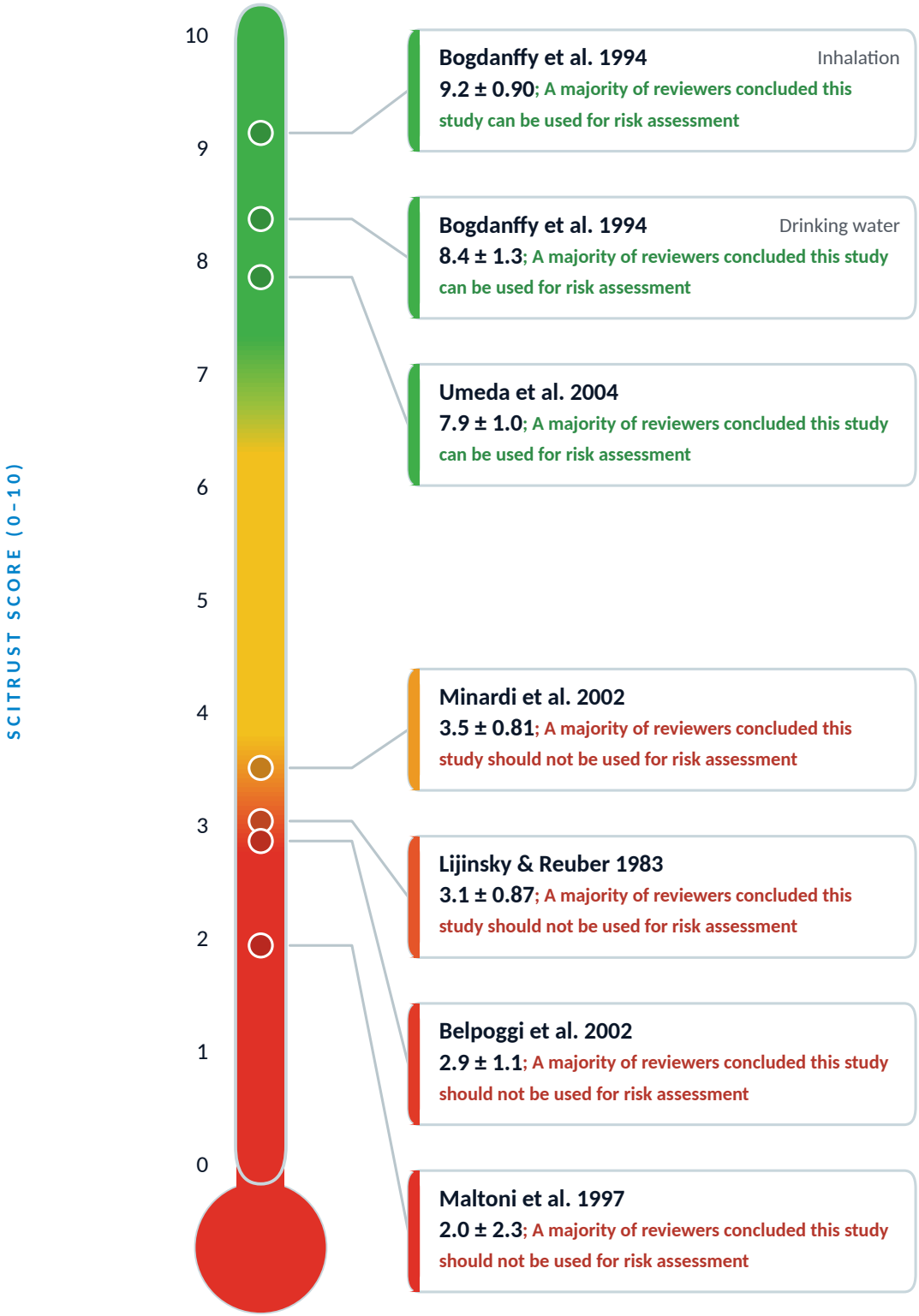
PREPARED BY:

SciPinion

SciPinion L.L.C.
777 E. Main St., Suite 202
Bozeman, MT 59715

shays@scipinion.com
ckirman@scipinion.com
(406) 404-6145

Figure 1. Summary of SciTrust Scores for the Cancer Bioassays Conducted for Vinyl Acetate Monomer



SciPinion

SciPi #1052

October 19, 2026

Round 1: Belpoggi et al. 2002

SciPinion L.L.C.
777 E. Main St., Suite 202
Bozeman, MT 59715

shays@scipinion.com
ckirman@scipinion.com
(406) 404-6145

What are the strengths of Belpoggi et al. (2002) (if any)?

Results | 5 answers

Expert 1

> Longterm exposure of Wistar rats to VAM for 104 weeks (appropriate animal model), test item characterized, fresh preparations of water with test item generated daily, examination of relevant organs by histopatho.; app. controls included.
>
> All studies from Ramazinni show local tumors at point of entry. For this endpoint, data are consistent. Systemic tumors reported in some of the studies, relevance unclear.

2 comments

Expert 2 | 10/10/2026 6:49 SCORE: **1**

Agree with comment on systemic tumors.

Expert 5 | 10/10/2026 8:00 SCORE: **1**

Agree about the unclear relevance of systemic tumors

Expert 2

> [Note: Although several strengths are listed below, it's important to consider the many weaknesses/limitations noted in Q1.2 which substantially impact confidence in the reported findings.]
> Strengths:
> • Two dose levels (1000 and 5000 ppm) including concurrent control were evaluated,
> • Wistar rats (both sexes) were exposed from*in utero* to adulthood to evaluate lifespan effects,
> • Multiple tissues were examined, and tumor data tables are provided,
> • Body weight data and survival data are provided, and based on the plots, these parameters would not have compromised the study,
> • Authors indicated that test material purity >99%; impurities identified-Benzene: 30-40 ppm, Methyl and ethyl acetate: 50 ppm, Crotonaldehyde: 6-16 ppm, Acetaldehyde: 2-11 ppm, Acetone: 330-500 ppm,
> • Authors indicated that vinyl acetate (VA) drinking water solutions were replaced each morning or daily.

1 comment

Expert 1 | 10/10/2026 5:42 SCORE: **0**

Agree that these are the positive aspects. If conclusions are restricted to the tumors in the oral cavity, the Ramazinni-studies are ok. Problem is keeping the animals until natural death with a potential delay between time of death and tissue sampling. These points have little relevance for the local tumors, but confound incidences of systemic tumors, specifically in tissue where there is a steep increase in tumor incidence with age

Expert 4

> This a complimentary study from the same Ramazzini group to the Maltoni et al (1997) study in this case examining rats instead of mice. Because of the fundamental experimental design flaws the only valuable data are: 1) the body weight and survival determinations (although no statistics were performed on the results of these parameters) and 2) the fact that long term treatment of a VAM mixture produced tumors in rats.

2 comments

Expert 1 | 10/10/2026 5:43 SCORE: **0**

Agree with conclusions, body weight gain and survivel does not compromise results regarding tumors in the upper Gi tract

Expert 5 | 10/10/2026 8:02 SCORE: **1**

Agree with these two points. But having the survival data, why not use it to improve the dose response?

Expert 5

> Belpoggi et al. exhibits few strengths: only two doses, no MTD consideration, no verification of drinking water conc. or attempt at estimating intake/dose, no consideration of historical controls, confusing tables, etc.

1 comment

Expert 2 | 10/10/2026 6:48 SCORE: **1**

I'm not necessarily in disagreement, if one considers the weaknesses and limitations (responses to 1.2).

Overall sentiment: Experts broadly agreed that the study's strengths are limited and must be weighed against substantial design weaknesses, but most identified at least a few redeeming features.

Points of agreement:

• The lifetime exposure design in Wistar rats (both sexes, in utero to adulthood, ~104 weeks) was considered an appropriate model, with concurrent controls and two dose levels.
• Test material was reasonably well characterized (purity >99%, impurities identified) and dosing solutions were freshly prepared daily.
• Body weight and survival data were reported and, in the view of several experts, did not appear to compromise the tumor findings — though it was noted that no statistics were applied to these parameters.
• Multiple tissues were examined histopathologically and tumor data tables were provided.
• Strong consensus that findings for local (point-of-entry, upper GI/oral cavity) tumors are consistent across Ramazzini studies and are the most defensible result, whereas the relevance of reported systemic tumors is unclear and potentially confounded by the "until natural death" protocol and delays between death and tissue sampling.

Points of divergence and critique:

• Expert 5 was the most skeptical, arguing the study "exhibits few strengths," citing only two doses, no MTD consideration, no verification of drinking water concentrations or intake/dose estimates, no use of historical controls, and confusing tables; Expert 2 did not disagree once weaknesses were factored in.
• Expert 4 took a similarly restrictive view, holding that fundamental design flaws leave only the body weight/survival data and the qualitative demonstration that long-term VAM treatment produced tumors in rats as valuable.
• Expert 5 additionally questioned why available survival data were not used to refine the dose-response analysis.
• Experts 1 and 2 were somewhat more generous, listing several methodological positives, while explicitly conditioning confidence in the findings on the study's limitations.

AI generated summary content

What are the weaknesses of Belpoggi et al. (2002) (if any)?

Results | 5 answers

Expert 1

Non-guideline study design keeping animals until natural death, no info on drinking water consumption, number of animals/group different from OECD/EPA guidance , little information on hygiene/infectious diseases in colony, no info on source of animals (have been used at institution for 30 years ?), source of test item not specified, degradation of test item in drinking water not assessed.

1 comment

Expert 2 | 10/10/2026 7:31 SCORE: **0**

Agree with all the points

Expert 2

[Note: Belpoggi et al and Minardi et al have many of the same issues which is not surprising given the studies were conducted at same Institute.]
• The actual administered dose of VA is not known. There is a lack of information on the stability and verification of VA in drinking water solutions, and assessment of actual intake. Decomposition to acetaldehyde (AA) and acetic acid can occur and could contribute to the cytotoxicity/inflammation and site-of-contact responses.
• Insufficient number of F0 males (13-14) which is too small to be reliable. F0 females are also low (37).
• Lack of transparency regarding the pathology. It is not known whether the diagnosis was blinded. Moreover, no detail provided in paper on how lesions are classified to support malignant tumor diagnosis. [Note: A common criticism/concern of Ramazzini (RI) lifetime studies has been potential misclassification of malignant lesions.]
• There is no information how breeder animals (F0) and offspring (F1) were allocated across dose groups. Thus, it is unknown whether randomization was done. There was a lack of information on litter of origin for the tumor data. Littermates can respond similarly (intra litter effects/cluster).
• Animals are exposure for 104 weeks and kept under observation until natural death. So, there is an unusually long life span observation which is not necessarily a weakness per se, but makes interpretation of the tumor responses challenging, including pathology, statistics. Also, in the absence of historical data, the issue can be raised whether the background rates are stable in aging rats. There was an independent analysis by Gentry et al (2024) which focused on the RI studies in SD rats and Swiss mice (Critical Reviews in Toxicology, 54(3):153–173). Although Wistar rat data are not available, this independent analysis does raise issues regarding high background rates/variability for certain tumors.
• Survival was "slightly decreased" from the "beginning of treatment to 104 weeks of age" when compared to controls. It is unlikely that slight changes in survival would impact the tumor incidence. Nonetheless, a statistical method was not applied to account for any differences in survival which is good practice,
• It should be noted that bronchopneumonia may have been present during the time this study was conducted. For older RI studies there have been criticisms of chronic infection/inflammation in some colonies which is an issue for certain tumors, and possibly for oral lesions. But we don't know if this was an issue. The authors only vaguely state that the "status and behaviour of the animals were examined 3 times daily."

2 comments

Expert 1 | 10/10/2026 5:46 SCORE: **0**

Agree with the comments, but do not see problems regarding the final verdict that VAM causes tumors in the oral cavity/upper GI tract

2 Replies

Expert 2 | 10/10/2026 7:21 SCORE: **0**

I agree that the oral cavity tumors cannot be entirely dismissed. There is biological plausibility. There was a study by Robinson, Bogdanffy, et al (2002, Toxicol 180:209-220) showing carboxylesterase activity in oral mucosa, thus leading to hydrolysis and local intracellular acidity. However, there are uncertainties with the oral cavity response particularly concerning the pathology (see response to 1.3).

Expert 1 | 10/10/2026 7:27 SCORE: **0**

Ok, pathology descriptions in Ramazinni-studies are apparently not precise, I see this as another restriction

Expert 5 | 10/10/2026 8:06 SCORE: **0**

I did not know about the bronchopneumonia at that time in history. Interesting. Agree that the authors statement is inadequate.

Expert 4

The weaknesses of this study are multiple but fundamentally rest in the experimental design.
1. Treating the rats for 104 weeks but only determining the tumor incidence at spontaneous death is wrought with problems. The paper states *"Upon death, animals were submitted to systematic necropsy"*. The analysis of the tissues from a spontaneous death is not an acceptable means of correctly performing proper histopathology of the tissue and quantitating the tumor response as well as correctly comparing the tumor response between treatment groups. Also, no mention as to when the animals died (time of day) or when (how many times per day) the rats were observed in the post treatment period in order to identify the "spontaneous death". I find it hard to believe that an animal that died overnight produced the same quality histopathology as a rat sampled immediately at death. This sampling issue should have been detailed in the document.

2. Although *"statistical analysis"* was reportedly performed it is unclear as to what treament groups the statistical comparisons were made (Table 11)? How come no statistical analysis was performed on the body weight and survival data?

2 comments

Expert 1 | 10/10/2026 5:48 SCORE: **0**

Agree with the points raised as significant limitations that prevent use of the data for risk assessment

Expert 2 | 10/10/2026 7:30 SCORE: **0**

I also agree with Expert 4. I also had the concern about the lack pathology detail and the potential for misclassification (e.g., were benign lesions upgraded to malignancy).

Expert 5 | 10/10/2026 8:17 SCORE: **1**

Spot on about the poor statistical analysis. The papers by Portier & Bailer on survival adjustment were from the 1980s so there's no excuse for not using these methods.

Expert 5

> Belpoggi et al. exhibits a number of weaknesses--too few animals tested. lack of detail in reporting the statistical methodsand pathology examination, overstating the conclusions about VAM carcinogenicity.

2 comments

Expert 2 | 10/10/2026 7:25 SCORE: **0**

I agree about the problems with the pathology and the potential to over state the statistical significance of the oral cavity tumors. And it is true the number of F0 males were low, but what about the F0 females and F1 males/females? The issue I had was the lack of information of how Breeders and F1s were allocated across dose groups.

Expert 5 | 10/10/2026 8:10 SCORE: **0**

I wrote answers to each of the 21 questions about Methods. The summary in accurate. To obtain scores for Methods, the score I gave for each question ranged from 0 to 0.5. The highest possible score would be 10.5. I summed the scores are rounded to the closest digit. For results, the answers to the 10 questions ranged from 0 to 1, and for Conclusions, the answers to the 7 questions ranged from 0 to 1.4. My full answers and scores are provided in the next section.

1 Reply

Expert 2 | 10/10/2026 10:16 SCORE: **0**

You are correct, I didn't see your detailed response in 1.3 .

Experts were broadly unanimous that Belpoggi et al. (2002) suffers from substantial methodological limitations, though they differed on whether these flaws invalidate the study's central finding.

Points of agreement:

• Test-substance characterisation and dosing:No verification of vinyl acetate stability in drinking water, no data on water consumption or actual intake, unspecified test-item source, and potential degradation to acetaldehyde/acetic acid that could drive site-of-contact effects.
• Inadequate group sizes:Particularly the small number of F0 males (13–14), deviating from OECD/EPA guideline expectations.
• Pathology reporting:Lack of transparency on blinding, lesion classification criteria, and the recognised risk of misclassifying benign lesions as malignant in Ramazzini Institute studies.
• Statistical shortcomings:Unclear comparison groups, no analysis of body weight or survival data, and failure to apply survival-adjusted methods (noted as inexcusable given Portier & Bailer's 1980s methods).
• Non-guideline design:Treatment to 104 weeks with observation until natural death complicates interpretation; sparse reporting on colony hygiene, infection (e.g., bronchopneumonia), and animal source.
• Randomisation/allocation:No information on how F0 breeders and F1 offspring were assigned to dose groups, nor on litter of origin (intra-litter clustering).

Points of divergence: Expert 1 and Expert 4 regarded the limitations—especially tissue collection at spontaneous death and the absent statistical treatment—as severe enough to preclude use of the data in risk assessment, with Expert 4 questioning histopathology quality from animals dying overnight. Expert 1 nonetheless maintained that the conclusion that VAM induces oral cavity/upper GI tumours remains defensible. Expert 2 partly concurred, citing biological plausibility via carboxylesterase activity in oral mucosa (Robinson, Bogdanffy et al., 2002), while stressing residual uncertainty in the pathology. Expert 5 went further, characterising the conclusions on VAM carcinogenicity as overstated. There was also minor disagreement over scope: Expert 2 noted that Expert 5's emphasis on "too few animals" applied mainly to F0 males, with allocation transparency being the larger concern.

AI generated summary content

Please rate your degree of confidence in each of the following study components for Belpoggi et al. (2002) (0-10 scale; 0 = fundamentally flawed; 1=lowest confidence; 5=medium confidence; 10=highest confidence). An overall confidence score will be calculated as the geometric mean of your three ratings. Please provide your rationale for each study component in the box beside it.

Results | 5 answers

	Methods	Results	Discussion/ Conclusions	SciTrust Score
Expert 2	3	3	3	3
Expert 1	4	4	5	4.3
Expert 5	3	6	1	2.6
Expert 4	2	2	1	1.6
Avg	3	3.8	2.5	2.9
Range	2–4	2–6	1–5	1.6–4.3

0 = no confidence · 5 = medium confidence · 10 = highest confidence

SciTrust Score = geometric mean of a reviewer's category scores. A zero in any category zeroes out that reviewer's full score. The **Final SciTrust Score** is the mean of all reviewer SciTrust Scores.

Overall, experts converged on low confidence in this study, with Methods ratings clustering at 2–4, Results at 2–6, and Discussion/Conclusions at 1–5. All four reviewers judged the reporting to be incomplete and the study insufficiently documented for quantitative risk assessment.

Points of agreement:

• No analytical verification of vinyl acetate concentration or stability in drinking water, and no measurement of water consumption, so actual delivered dose is unknown (all experts).
• Absence of historical control data, no mention of blinding or independent pathology review, and no information on randomization or litter of origin for F1 animals.
• Statistical methods were described too vaguely to interpret or reproduce; no survival adjustment (e.g., poly-k), no correction for multiple comparisons, and unclear comparison groups in tables (noted especially by Experts 4 and 5; Expert 5 could not replicate the reported trend results).
• Essentially no discussion section — limitations, dose selection/MTD, and inconsistencies with other VAM drinking-water studies (e.g., Bogdanffy) were not addressed, and the "multi-site carcinogen" conclusion was viewed as overreaching.
• Breeder (F0) group sizes were too small to support conclusions, while offspring numbers were adequate.

Points of divergence:

• Expert 1 gave somewhat higher scores, arguing the portal-of-entry (oral cavity/upper GI) tumor findings are robust because they are consistent with other oral VAM studies; Expert 2 agreed in principle but scored lower due to concerns about diagnostic criteria for malignancy and unknown allocation of pups.
• Expert 5 rated Results relatively highly (6), since tabulated data permitted independent dose–response exploration, but rated Discussion/Conclusions lowest (1); Expert 4 was uniformly most critical, also flagging unaddressed impurities (benzene, acetone, esters, aldehydes) in the test material.
• Interpretation of the elevated F1 female uterine/oral squamous carcinoma response differed: Expert 5 considered it an outlier given weaker responses in other groups, while Expert 2 raised possible clustering by litter, misclassification, lack of precursor lesions, and absent biological plausibility — the two ultimately converged on skepticism.
• Experts disagreed on whether the high dose exceeded the MTD (Expert 1 suspected so in males; Expert 5 saw minimal body-weight effects and comparable survival).

AI generated summary content

ANSWER EXPLANATIONS

ANSWER EXPLANATIONS, CONTINUED

Expert 1 | **SciTrust Score:** 4.3

	Confidence Rating
Methods	4
Results	4
Discussion/Conclusions	5 (medium confidence)

Methods

Will integrate Scipinion guidance in scoring considerations and provide mean of my scoring in confidence rating

Appropriate species (score of 5). Wistar rats can be used for carcinogenicity studies, no justification for selection besides "have been used at institutions for > 30 years"), original source unknown, no information on incidences in historic controls.
No. of animals/group (score of 5).
Controls included and exposed to tap (?) water only (score of 6), low number limits possible conclusions in males. Source of water not stated.
Study duration (score of 2). Exposure for 104 weeks as in guidance, but animals were examined for tumors onyl after natural death (issues with rapidly increasing incidences of tumors in aging rats, histopathology analysis of lysed tissues
Number and spacing of dose levels (score of 5), controls and two dose levels, no justification for dose levels, high dose in males > MTD ? dose received not determined
Rote of exposure may be relevant to humans (score of 7) (migration of VAM from food contact materials into food ? how stable is VAM in food contect materials, food ?)
Test item preparation, source, verification (score of 6). Fresh preparations daily, little information on source of test item, no analytical verification of target conc and stability, presence of actaldehyde as decomposition product
Determination of dose received (score of 0). No information given
Necroscopy, histopatho (score of 4), Organs selected for histopath similar to guidance, apparently all animals including early deaths subjected to histopatho, little information on tissue processing given, issues with histopatho analyses of lysed tissues when dead animals remain in cages for a longer time. No info on blinding of evaluators
Tumor incidences, stats (score of 2). Limited information given, study results may only be used for descriptive purposes. Little information for quantitative assessment. Dose-response for total malignant tumors not evident.
Guideline (score of 2), duration of exposure according to guidance, no terminal sacrifice at defined time point, no info on clin chem and urine analysis, limited information on preneoplastic lesions.

Results

General conduct: no info on infectious disease outbreak or lab hygiene, question of GLP status. Main finding are local tumors, unlikely to be compromised by shortcomings (score 6).

Survival and effects of toxicity on survival/effects: high dose in males may be > MTD, females ok (score 5). Dose-response not evident ?

Reporting of tumor incidences: Mainly in tables giving number and % of animals with tumors (benign and malignant) and total malignant tumors, focus on tumor incidence in oral cavity and upper GI tract, information on precursor lesions in oral cavity, but not on possible precursors for the systemic tumors reported (score 5)

Issues with incidences of haemolymphoreticular neoplasias and infectious disease/age-dependent increases in uterine tumors significantly reduce confidence in the relevance of these tumors (score 2)

Discussion/Conclusions

Conclusions on tumor incidences in oral cavity/upper GI tract well supported (score 8) by observations

Limitations of study design not addressed, impact of design on study outcome not discussed. Actually, the discussion is quite short. However, the main finding (oral and upper GI tract tumors due to VAM) is consistent with the local tumor effects seen with VAM in other oral studies (score 7).

Study results difficult to use in quantitative risk assessment due to deficiencies in reporting and study design (score 2)

Discussion should have addressed differences in study outcome of the drinking water studies (no lesions in oral cavity in Bogdanffy vs. tumors in this study despite identical top concentrations. Differences in kinetics of decomposition of VAM due to different composition of water used for administration ?

Conclusions: Due to issues with study design and reporting, the study receives low quality rating and should not be used for quantitative assessments. In addition, systemic tumors reported should not be used for hazard assessment of VAM due to study design issues. However, study shows that VAM induces tumors at portal of entry after adminsitration of high conc. in drinking water.

ANSWER EXPLANATIONS, CONTINUED

Expert 2 | **SciTrust Score:** 3

	Confidence Rating
Methods	3
Results	3
Discussion/Conclusions	3

Methods

See the many weaknesses in response to Q1.2

There is no information on measurements of VA (as well as AA and acetic acid) in the drinking water solutions at the end of 24 hours. If the drinking water solutions were replaced daily, however, the decomposition is less but not eliminated. Based on the Bogdanffy et al paper, regarding stability of VA in drinking water, up to 14% of VA was lost in the first 24 hr and 40 to 60% over a 7-day period.

There were no statements about whether widely international standardized histopathology criteria were used or whether the diagnosis was blinded. The paper only indicated that slides were examined using same criteria/classification by the same pathologists with a senior pathologist reviewing all tumors and any other cancer lesions of interest.

There is no information on whether randomization was done to allocate animals across dose groups, including at the litter/pup level.

Results

The actual intake of VA is not known. Water intake is not available and VA, as well as AA and acetic acid, wasn't measured in the drinking water solutions. It's unknown to what extent (if at all) potential decomposition products contributed to the effects following chronic exposure.

Port-of-Entry Tumors (Upper GI Tract)

There was a statistically significant increase in carcinomas of the oral cavity in F1 (offspring) males and females at the highest dose tested (5000 ppm). VA does cause cytotoxicity (irritation/inflammation) at the site of contact. Thus, these responses could be due to local rapid hydrolysis of VA in the oral mucosal. Although these tumors should not be dismissed, there are several uncertainties that warrant caution about whether the carcinogenic responses are over interpreted or over represented.

• The possibility of misclassification of carcinomas cannot be ruled out. There is no detailed pathology description (or supporting microphotographs). Thus, it is not clear whether invasion of the basement membrane occurred for the carcinoma diagnosis. Were some non-neoplastic inflammatory/proliferative responses misclassified as carcinomas? There is no grading of the dysplasia (mild, moderate, severe?) provided. [Note; RI drinking water studies on other chemicals have reported increased incidences of upper GI tract tumors which are uncommon, and been criticized for the pathology (e.g., distinction between irritation lesions and true malignancy). Certain studies were especially considered suspect given the chemicals evaluated were not strong irritants.]

• Also, we do not know how the animals were allocated across dose groups. Are the responses independent of litter? A stronger analysis would have provided information on the litter of origin. Only group level tumor percentages are provided.

Systemic Tumors

Uterine tumors-The authors indicated that "The number of malignant tumours (sarcomas and carcinomas) of the uterus was found to have statistically increased in F1 offspring treated with the higher dose. Table 17 lumps the adenocarcinomas and

squamous carcinomas which are epithelial tumors. The adenocarcinomas were only increased at the highest dose in offspring, and the increase in squamous carcinomas is not statistically significant. The fibrosarcoma response is not statistically significant. Pooled data provided in this article should be interpreted with caution.

Furthermore, there is the potential for misclassification (non-neoplastic proliferative responses vs true carcinoma). Aged rats can develop uterine proliferative lesions. There is no detailed pathology description provided, particularly of clear invasive growth. There is a lack of historical data (Gentry et al didn't have information on uterus sarcomas or carcinomas). Thus, background stability or variability over time for these uterine tumors in aging rats is not understood. Also, were all the uterine tumors diagnosed at natural death? Are these tumors independent of litter ?

Hemolymphoreticular tumors - A slight increase was reported at the highest dose test in F0 females. The concurrent control incidence is low. It should be noted that Gentry et al found (SD rats) a high degree of variability across controls for lymphoma/leukemia for SD rats. [Gentry et al,-"lymphoma/leukemia rate of more than 10% has increased from 7% (three out of 43 studies) of RI studies conducted between 1988 and 1989 to upwards of 82% (18 out of 22 studies) of studies conducted from 2002 to 2006".]

Although OEHHA note adrenal pheochromocytomas (F0 females only, not dose related)/pheochromoblastomas (F1 males only, $p<0.5$) and pancreatic exocrine adenomas (F1 males only not statistically significant), Belpoggi et al do not highlight nor discuss these tumors. The same issues apply-lack of historical control data to understand background variability in aging rats, and uncertainty of potential litter effects for F1 responses, diagnosis.

Discussion/Conclusions
There is no discussion section, only a broad conclusion that VA is a multi-site carcinogen which is not supported but the evidence provided. I would not describe a chemical as a multi-site carcinogen just based on site of contact tumors (mostly oral cavity carcinomas) at the highest dose tested. Moreover, the evidence presented is not convincing for the systemic tumors particularly given the uncertainties. The authors specifically note that the upper GI tract tumors are of particular significance, which are biologically plausible and cannot be entirely dismissed. However, caution should be taken given the uncertainties about over interpretation or over represented.

Expert 3

	Confidence Rating
Methods	
Results	
Discussion/Conclusions	

No explanation provided.

ANSWER EXPLANATIONS, CONTINUED

Expert 4 | **SciTrust Score:** 1.6

	Confidence Rating
Methods	2
Results	2
Discussion/Conclusions	1 (lowest confidence)

Methods

The methods are not detailed enough to have another laboratory repeat the study (a foundation of proper science) and anticipate similar results. No analysis of the concentrations of the VAM in the drinking water (during preparation and after administration were reported). The statistics performed were not complete or detailed enough to reach a conclusion on significance. The description of the methods is incomplete.

Results

In several of the tables, data are highlighted in bold but there is no comment on the significance of these bold highlights are mentioned in the document. The analysis of the results is tainted by the experimental design and ill-defined statistics. In Table 11 where the values for tumor bearing animals and number of tumors per 100 animals is presented, it is unclear to what groups the statistical evaluations were being compared. Was the statistical significance between dose groups? Between the treated and control? No indication in the table to what the comparisons are being made.

Discussion/Conclusions

I could not find a discussion section? Only results section and conclusions section were provided. As such no mention of how the results of this study compare with other VAM results from the literature. No mention in the discussion/ conclusions is made in the manuscript on the composition of the VAM used in this study, particularly the contamination of benzene (30-45 ppm), methyl and ethyl acetate (50 ppm), Acetone (330 – 500 ppm) and select aldehydes. Concentrations of benzene at 45 ppm have been linked to blood disorders in rats.

ANSWER EXPLANATIONS, CONTINUED

Expert 5 | **SciTrust Score:** 2.6

	Confidence Rating
Methods	3
Results	6
Discussion/Conclusions	1 (lowest confidence)

Methods

1) Score=0.1. No justification was provided for the use of Wistar rats. That these experiments were conducted 20 years prior to publication seems odd. Perhaps the choice was dictated by the strains available at the Ramazzini Institute.

2) Score=0.2, The number of animals was insufficient. For male breeders, the numbers were 0ppm – 14, 1000ppm – 13, 5000ppm – 13. For female breeders, the numbers were insufficient with 37 in each dose group. The numbers were adequate for offspring.

3) Score=0.2. The breeder control group was inadequate as it had fewer than 50 animals of each sex. The offspring were adequate.

4) Score=0.5. The study duration of 104 weeks was appropriate for lifetime exposure in rats.

5) Score=0.2. Only two dose levels were used: 1000 and 5000ppm. In the other drinking water experiments detailed in Table 4, doses of 1000 and 2500 mg/L were used in Fischer rats. Doses of 200, 1000, and 5000 mg/L were used win SD rats. Doses of 400, 2000, and 10000 ppm were used in F344 rats. Why did this study include a low dose group of 200 – 400 ppm?

6) Score=0. No discussion of the MTD was provided. Effects of body weight were small, and the authors possibly thought Figs. 1 and 2 could suffice for a discussion of MTD.

7) Score=0.4. The concentration of VAM in wastewater of 50 mg/L to which humans could be potentially exposed is orders of magnitude lower than those used in this study. Hence, while the route of exposure could potentially be relevant to human exposures, the actual dose would likely be much lower.

8) Score=0.1. Table 5 gives the purity of the test substance. No discussion of stability in the test solution was provided nor was any ongoing purity testing mentioned.

9) Score=0. Stock solution of VAM was prepared daily but no mention was made of any analysis of the drinking water. Hence, no verification occurred.

10) Score = 0. The study reported water concentrations but did not mention measuring the remaining water from the previous day. Doing so would have enabled calculation of the average daily dose for each set of 5 animals (5 per cage). No mention was made of tap water analysis to ensure it did not contain any VAM.

11) Score=0. As indicated, water consumption was not measured

12) Score=0.5. Table 8 gives a list of the 28 tissues examined during necropsy and seems complete.

13) Score=0.1. The tissues were processed per SOP at CRC/RF rather than widely accepted standards.

14) Score=0. Whether the pathologists were aware of the dose group was not mentioned. No independent pathology review was not mentioned either.

15) Score=0.1. The study indicated the Cochrane-Armitage test for trend in proportions was used to analyze the dose-response relationship and the Chi-square test was used for tumor incidence. No details are provided. Nine years before Belpoggi et al. was published, Bieler and Williams (1993) provided a summary of changes to the Cochrane-Armitage test to incorporate adjustments for survival time and variance in tumor proportions. Presumably, the Chi-square test was used to compare tumor proportions between dose groups. The paper, however, gives neither explanations nor a worked example.

16) Score=0. Figs. 3-6 show the survival curves for breeder and offspring rats of both sexes. Clearly, some animals died before 104 weeks. No adjustment such as poly-k procedure was mentioned.

17) Score=0. Historical controls were not mentioned in the paper.

18) Score=0.1. Benign and malignant tumors were combined in Tables 9 and 10 without any explanation. Table 11 showed malignant tumors only. 28 different tissues were examined by pathologists (Table 8). Tables 9 and 10 (male and female respectively) provide the proportion of animals with benign and malignant tumors, and Table 11 gives the proportions of animals with malignant tumors by sex. The tables seem designed to be confusing and obtuse.

19) Score=0. No discussion of adjustment for multiple comparisons was presented in the paper.

20) Score=0. No, the paper does not meet these guidelines because of the insufficient number of animals and the opaque statistical methods.

21) Score = 0. No, the paper does not meet these guidelines because of the insufficient number of animals and the opaque statistical methods. The deviations from guidelines are not acknowledged.

Results

1) Score=0.7. No infectious disease issues were mentioned and presumably, none occurred.

2) Score = 0.2. Narrative at the bottom of p. 286, Table 9 and Table 10 indicate benign tumor occurrence was unaffected by treatment. The plots in Figs. 3-4 indicate about 60% of both M and F breeders survived to 104 days. Figs. 5-6 indicate 40% of M offspring and 60% of F offspring survived to 104 days. Given the low numbers of breeders, survival adjustment was needed, especially for males with 14, 13, 13 individuals in the 0, 1000 and 5000 ppm groups respectively.

3) Score = 0.8. The survival plots (Figs. 3-5) show that the survival of the control groups was similar to that of treated groups. Hence, I doubt whether excessive toxicity could have affected the results. No discussion was present in the paper.

4) Score=0.5. Figs. 1 and 2 show changes in body weight for breeders and offspring. The treated and control groups' body weights track closely over time. Hence, I doubt whether reduced water intake could have affected the animals. Presenting the daily water consumption figures, which were not obtained, would have helped demonstrate the lack of effect of VAM.

5) Score=0.1. I absolutely despise the tabie format showing decreasing dose. In any case, the data are available. I was able to examine the actual dose-response data after transfer to a spreadsheet.No individual animal data was provided—the paper is not in PubMed and the European Journal of Oncology archive will not open volume 7. In sum, the data are poorly presented.

6) Score = 0.5. The data were reported for the oral cavity, tongue, esophagus, forestomach, upper GI tract with squamous cell carcinomas and squamous cell displasias combined, the uterus as carcinomas, sarcomas and total malignancies, and hemolymphoreticular tumors, lymphomas and leukemias. I was able to copy the values in the tables to a spreadsheet. I

examined the results for those reporting a statistically significant trend using the Cochran Armitage test implemented in th R package 'DescTools'.

· M offspring tumor-bearing animals (Tab. 11) p = 0.4084

· F offsprint oral squamous cell ca (Tab. 12) p = 0.0009

· F breeders uterine total malignant tumors (Tab. 15) p = 0.8585

The point is that I could not replicate the statistical results, but the tables allowed one to explore dose-response.

7) Score=0.7. The findings were reported as # affected and total # at risk. Sometimes various endopoints were combined such as dysplasias and carcinomas as in Table 16. I'm assuming that those of the 28 tissues in Table 8 not reported did not have tumors or or other lesions.

8) Score=1. Body weight effects of the treatment were minimal as shown in Figs. 1 and 2 with comment on p.282.

9) Score = 1. No disease or infection was reported and I assume none were present.

10) Score=0.8. Tables 12, 13,14, and 15 reported squamous cell dysplasias and squamous cell carcinomas separately. Table 17 combined these and reported them as No. per 100 animals.

Discussion/Conclusions

1) Score=0.1. As noted in the Conclusions, VAM exposure was associated with tumors in the oral region and GI tract. The number of breeder animals tested was insufficient to support the causal statement in the first sentence on p. 290.

2) Score=0.1. The concluding sentence of the second paragraph regarding worker protection and the possibility of VAM exposur efrom plastic cannot be supported due to the extremely high concentrations tested on the rats compared with the environmental levels in Table 3.

3) Score = 0.2. Whilst the numbers of offspring were sufficient (>50) for drawing conclusions, any conclusions based on the breeders are suspect, especially in males that had less than 15 animals per dose group. In fact, the male breeders had no significant results save for tumors per 100 animals in Table 11. For the breeders, both male and female, the study lacked sufficient power to draw conclusions.

4) Score =0.2. The narrative of p. 290 fails to discuss the limitations due to the small number of animals.

5) No. Score=0

6) No. Score=0

7) No. Score=0

Debate | 2 comments

Expert 1 | 10/10/2026 5:51 SCORE: **0**

My higher score is based on the outcome of local tumors that actually is observed in several VAM studies. Share all other aspects of critisim made by the colleagues

1 Reply

Expert 2 | 10/10/2026 7:37 SCORE: **0**

Expert 1 in reading your responses, I agree with your points. I gave the study a slightly lower score given concerns about the pathological diagnosis of malignant tumors and the absence on how animals and pups were allocated across dose groups (uncertain if randomization was done). There is no information on the litter of origin for the tumor responses (were the F1 responses concentrated from a litter or 2?).

Expert 2 | 10/10/2026 7:56 SCORE: **0**

Expert 5, thank you for your detailed analysis. I agree with your points. Can you elaborate more about your thinking on the Female F1 squamous carcinoma response (24/95, 25.3%; controls 5/69, 7.2%).

3 Replies

Expert 5 | 10/10/2026 8:44 SCORE: **0**

I did wonder about that in contrast to the breeder F that had lesser response. In addition, the breeder M and offspring M both had lower responses. Hence, I considered the offspring F response an outlier.

Expert 2 | 10/10/2026 13:05 SCORE: **0**

Yes, It could be an artifact or outlier. It is an unusually high incidence in the F1s (mostly adenocarcinomas). Is this a cluster effect (only group level numbers/percentages provided)?, is there some misclassification of malignancy (no invasion criteria provided)? no historical data for context, no precursor lesions provided (eg endometrial hyperplasia), no information on when tumors diagnosed (at natural death, earlier), no dose response, no biological plausibility given VA PK.

Expert 5 | 10/14/2026 8:12 SCORE: **0**

#2, Agree with all those nits you picked.

SciPinion

SciPi #1052

October 19, 2026

Round 1: Bogdanffy et al. 1994 (inhalation)

SciPinion L.L.C.
777 E. Main St., Suite 202
Bozeman, MT 59715

shays@scipinion.com
ckirman@scipinion.com
(406) 404-6145

What are the strengths of Bogdanffy et al. (1994; inhalation) (if any)?

Results | 5 answers

Expert 1

> Guideline inhalation study, detailed assessment of tumor incidences, 3 concentrations and controls, additional animals for interim sacs and recovery, detailed reporting of methods, study design, procedures

Expert 2

> • Two species evaluated (Sprague-Dawley rats; CD-1 mice) in both sexes,
> • Adequate group size (58-60), 3 exposure levels (50, 200, 600 ppm) and exposure duration of 104 weeks; including satellite group for recovery-70 week of exposure followed by 15 weeks of no exposure,
> • Thorough exposure characterization of VA; e.g., purity >99% and was stable under storage conditions; concentration of VA vapor monitored, controlled and measured frequently; levels of acetaldehyde were determined in test material and in high concentration chamber,
> • Range of endpoints and indicators of toxicity evaluated (e.g, clinical observations and chemistry, hematology, urinalyses, organ and body weights)
> • Comprehensive pathology - histopathology was done on respiratory tract tissues from all animals, on all tissues from control and high-dose animals and on tissues (from any exposure group) found to be abnormal during gross necropsy.
> • Multiple cross sections of the nasal cavity,
> • Appropriate statistical methods used including to account for differences in survival,
> • Independent pathological review of respiratory lesions (primary site).
> • Tables of body weight growth curves, non-neoplastic and neoplastic changes in respiratory tract,
> • The highest dose tested was adequate given lesions observed in upper respiratory tract, and did not compromise the study due to excessive toxicity; absolute BW reduction (10% rats; 15% mice) at 104 weeks.

2 comments

Expert 1 | 10/10/2026 6:58 SCORE: **0**

Agree with assessment

Expert 5 | 10/10/2026 8:47 SCORE: **0**

Also agree with these summaries

Expert 4

> This is a well-planned and well executed study. The manuscript is clearly written. The methods section is detailed and well defined. The results were properly presented, and the discussion and conclusion reflected the results is a logical manner.

1 comment

Expert 5 | 10/10/2026 9:07 SCORE: **0**

Agree with this general comment.

Expert 5

This is a well conducted study with sufficient numbers of animals. An unavoidable flaw is the presence of acetaldehyde albeit in low concentrations, that may account for some of the observed effects. Nontheless, a very good study.

3 comments

Expert 2 | 10/10/2026 8:22 SCORE: **0**

Expert 5. I agree with your comment about AA. The authors indicated that the mean concentration was 49 ppm. Couldn't find SD data but based on Wistar results, although we cannot rule out with confidence, I would think the impact would be negligible.

2 Replies

Expert 5 | 10/10/2026 9:07 SCORE: **0**

In rats, vinyl acetate is metabolized to acetaldehyde and acetic acid by carboxylesterases. The acetaldehyde and resulting acidification are clastogenic and may result in tumor formation (PMIDs: 9070362, 10048126). Hence, it is more likely that the endogenously formed AA and acidification was the driver for tumors rather than the impurities in the VA test material.

Expert 2 | 10/10/2026 10:19 SCORE: **0**

in agreement

Expert 4 | 10/11/2026 12:01 SCORE: **0**

Expert 5 ; very good comment and explanation That also has application to a possible mode of action

Expert 1 | 10/12/2026 2:39 SCORE: **0**

agree with the conclusion

Experts were unanimous in judging Bogdanffy et al. (1994) a well-designed, well-executed, and clearly reported inhalation study, with no substantive disagreement emerging in the discussion.

Commonly cited strengths included:

• Robust design:guideline-compliant chronic study in two species (Sprague-Dawley rats, CD-1 mice) and both sexes, adequate group sizes (58–60), three exposure concentrations (50, 200, 600 ppm) plus controls, 104-week duration, and additional animals for interim sacrifices and recovery (70 weeks exposure + 15 weeks recovery).
• Exposure characterization:test material purity >99%, demonstrated stability, and frequent monitoring/control of vapor concentrations, including measurement of acetaldehyde.
• Comprehensive endpoints and pathology:clinical observations, chemistry, hematology, urinalysis, organ/body weights; histopathology of respiratory tract in all animals, multiple nasal cross-sections, and independent pathology review of respiratory lesions.
• Analysis and reporting:appropriate statistics accounting for survival differences, detailed methods, and clear tabulation of body weights and neoplastic/non-neoplastic findings.
• Appropriate top dose:produced upper respiratory lesions without excessive toxicity (body weight reductions of ~10% in rats, ~15% in mice).

The only limitation raised was by Expert 5, who noted acetaldehyde contamination as an unavoidable flaw potentially contributing to observed effects. Expert 2 agreed the issue was worth noting but argued the impact was likely negligible given the low mean concentration (~49 ppm). Expert 5 further clarified—with agreement from Experts 1, 2, and 4—that endogenously formed acetaldehyde and acidification from vinyl acetate metabolism, rather than test-material impurities, more plausibly drive tumor formation, a point flagged as relevant to mode-of-action considerations.

AI generated summary content

What are the weaknesses of Bogdanffy et al. (1994; inhalation) (if any)?

Results | 5 answers

Expert 1

blinding of evaluating histopathologists not stated

2 comments

Expert 1 | 10/10/2026 5:54 SCORE: **0**

Actually, this as a GLP study and I assume that blinding was applied, just did not find the statment in the paper

2 Replies

Expert 2 | 10/10/2026 8:23 SCORE: **0**

Agree

Expert 5 | 10/10/2026 9:09 SCORE: **0**

I gave a low score due to the lack of mention of blinding.

Expert 4 | 10/11/2026 12:02 SCORE: **0**

agree

Expert 2

There are no substantial issues that would distract from the conclusions of this inhalation bioassay.

Expert 4

No apparent weaknesses were detected.

Expert 5

No survival adjustments were made, but such was not widely recognized in 1994.

2 comments

Expert 2 | 10/10/2026 8:25 SCORE: **0**

Authors stated that survival probability functions were estimated via Kaplan-Meirer and compared by log-rank procedure. Wouldn't that be sufficient?

Expert 5 | 10/10/2026 9:14 SCORE: **0**

Kaplan-Meier finds a function that models the number remaining alive. The poly-k procedure accounts for the reduced # of animals in the dose-response (PMIDs: 2744275, 18781526).

Experts broadly agreed that the Bogdanffy et al. (1994) inhalation bioassay is methodologically sound, with no substantive flaws that would undermine its conclusions. Two experts (Experts 2 and 4) explicitly identified no apparent weaknesses.

Where concerns were raised, they were minor and procedural rather than fundamental:

• Blinding of histopathologists:Expert 1 noted that blinding was not stated in the publication, though they presumed it was applied given the study's GLP status. Experts 2, 4, and 5 agreed this was an omission in reporting; Expert 5 indicated this lack of explicit mention lowered their quality score.
• Survival adjustment:Expert 5 observed that no survival adjustment was made to the tumor incidence analysis, while acknowledging such methods were not widely adopted in 1994. Expert 2 questioned this, pointing to the authors' use of Kaplan-Meier estimates and log-rank comparisons as potentially sufficient. Expert 5 clarified the distinction, explaining that Kaplan-Meier models survivorship but does not, unlike the poly-k procedure, account for reduced animals at risk within the dose-response analysis itself.

Overall, the discussion reflects consensus on the study's reliability, with the only genuine point of disagreement being the technical adequacy of the statistical treatment of survival.

AI generated summary content

Please rate your degree of confidence in each of the following study components for Bogdanffy et al. (1994; inhalation) (0-10 scale; 0 = fundamentally flawed; 1=lowest confidence; 5=medium confidence; 10=highest confidence). An overall confidence score will be calculated as the geometric mean of your three ratings. Please provide your rationale for each study component in the box beside it.

Results | 5 answers

	Methods	Results	Discussion/ Conclusions	SciTrust Score
Expert 2	9	10	10	9.7
Expert 1	9	9	9	9
Expert 5	7	9	8	8
Expert 4	10	10	10	10
Avg	8.8	9.5	9.3	9.2
Range	7–10	9–10	8–10	8–10

0 = no confidence · 5 = medium confidence · 10 = highest confidence

SciTrust Score = geometric mean of a reviewer's category scores. A zero in any category zeroes out that reviewer's full score. The **Final SciTrust Score** is the mean of all reviewer SciTrust Scores.

Experts expressed strong and consistent confidence in Bogdanffy et al. (1994), with all four raters scoring each study component between 7 and 10 and explicitly acknowledging their convergence in the debate ("looks like we all are on the same page").

• Methods (7–10):Reviewers agreed the study used appropriate species, adequate group sizes (60/group), air-only controls, a 104-week exposure duration, and well-justified concentrations (0, 50, 200, 600 ppm) based on prior work. Expert 1 and Expert 2 highlighted GLP conduct, analytical verification of chamber atmospheres, monitoring of acetaldehyde as a potential confounder, and two independent pathology reviews. Expert 4 simply judged the methods to follow standard chronic inhalation procedures.
• Points of minor divergence:Expert 5 gave the lowest methods score (7), citing unreported pathologist blinding, limited statistical detail, no survival adjustment, and absence of historical control comparisons. Expert 1 similarly noted blinding was not stated but presumed it based on GLP, illustrating differing tolerance for unreported procedural detail. Expert 2 explicitly rejected OEHHA's concern that non-respiratory tissues in low- and mid-dose animals were inadequately examined, considering the histopathology approach reasonable, and attributed any remaining gaps to publication format rather than study quality.
• Results (9–10):Unanimous confidence in the reporting of dose-related olfactory non-neoplastic lesions and nasal tumors in male rats, reduced body weight gain, absence of exposure-related mortality, and the lack of systemic toxicity or carcinogenicity. Experts agreed mouse findings showed non-neoplastic nasal changes without meaningful malignancy.
• Discussion/Conclusions (8–10):All agreed the conclusions are supported by the data, with praise for the treatment of VAM versus acetaldehyde, the metabolic rationale for olfactory versus respiratory epithelial sensitivity, and consideration of species differences. Expert 1 cautioned that some species-difference discussion is relevant to risk characterization but not directly derived from the presented data; Expert 5 identified the role of acetaldehyde as the principal residual uncertainty.

Overall, Experts 1, 2, and 4 regarded the study as suitable for quantitative risk assessment, and no reviewer identified any flaw substantial enough to undermine the study's utility.

AI generated summary content

ANSWER EXPLANATIONS

Expert 1 | **SciTrust Score:** 9

	Confidence Rating
Methods	9
Results	9
Discussion/Conclusions	9

Methods

Appropriate species (score of 10). SD rats and ICR mice were widely used for carcinogenicity studies, obtained from established breeder. Source of diet given
No. of animals/group > as in guidance (score 10). Additional satelite groups included for clinical chemistry and recovery groups. Health status assessed, clinical. chem and urine analysis included at interim sacs
Controls included and exposed in system to air only (score of 10), sex distribution as in guidance.
Study duration (score of 10). Exposure for 104 weeks as in guidance, daily exposure schedules as in guidance, study performed in well established labs under GLP.
Number and spacing of dose levels (score of 10), controls and three target concentrations, justification for conc. based on results from 3-month inhalation exposures to VAM.
Rote of exposure relevant to humans (score of 10). Inhalation is the most likely occupational exposure scenario.
Test item preparation, source, verification (score of 10). Detailed description of exposure system and generation of exposure atmossphere, analytical verification of conc. and homogeniety of VAM distribution in exposure system. Additional measurement of actaldehyde in chamber air as a potential exhaled metabolite of VAM that may confound results.
Necroscopy, histopatho (score of 9). Organ weights assessed. Organs selected for histopath similar to guidance, apparently all animals including early death subjected to histopatho, two independent histopath evaluations, specific focus on histopatho of upper respiratory system. Blinding of evaluating pathologists not stated, but expected based on GLP.
Tumor incidences, stats (score of 10). Methods appropriate, study results usefull for quantitative risk asssessment
Guideline (score of 10), adequate randomisation at start, duration according to guidance, no animals/dose group according to guidance, terminal sacrifice

Results

General conduct: (score 9). Targeted air conc. of VAM very close to intended conc. Detailed description of results, detailed histpatho analysis of respiratory tract in both species including time-dependent development of lesions

Survival and effects of toxicity on survival/effects (score 9): High dose in the range of maximum tolerate conc in both mice and rats, local effects in respiratory tract as only adverse effects, unlikely to influence validity of study,

Reporting of tumor incidences: (score 10). Detailed reporting of incidences of tumors and precursor lesions in upper respiratory tract in tables. Body weight gains in figs. Recovery groups as verbal description only. Detailed verbal description of lesions in respiratory tract in text.

Discussion/Conclusions

Conclusions supported by data in manuscript and clearly justified, discussions address dose-response, potential modes of action based on literature data. Some points regarding species differences may be relevant for risk characterization, but are are not directly related to the data presented.

Conclusions: High quality inhalation study that may be used for quantitative assessment of tumor risk after inhalation

ANSWER EXPLANATIONS, CONTINUED

Expert 2 | **SciTrust Score:** 9.7

	Confidence Rating
Methods	9
Results	10 (highest confidence)
Discussion/Conclusions	10 (highest confidence)

Methods

As detailed in Q 2.1, this is a high quality study in design, conduct and reporting. Moreover, the Bogdanffy et al article provided sufficient detail and reporting to make an independent judgement regarding the confidence of whether the results could be reproduced, which is an important consideration for a regulatory evaluation.

I would have given a 10, but this is a publication which would not have the detail as a laboratory report. But, it is likely that historical control data, raw data for control and treatment groups, and more methodological details could be obtain by requesting the Hazleton UK report.

Although OEHHA stated that "it is reasonable to conclude that the tissues (other than the respiratory tract) from low- and mid-dose animals were not adequately examined via histology." I do not find this to be a deficiency of note. The histopathology approach was reasonable.

Results

There was sufficient detail provided to have confidence in the findings. The data presented support the authors conclusions. The most significant finding was non-neoplastic and neoplastic lesions (statistically significant) in the olfactory epithelium in male rats. A pathological description of the lesions was provided, including morphologic detail and microphotographs. Olfactory non-neoplastic lesions occurred in females but there was not a significant increase in tumors. The presence of the non-neoplastic lesions in the olfactory epithelium together with the related neoplasms (correlated dose response) supports the conclusion of a carcinogenic port-of-entry effect via the inhalation route.

Although non-neoplastic changes were present in the mouse nasal cavity (similar to the rat with some key differences), no significant malignancies were found in the mouse study.

There was no evidence of systemic toxicity or carcinogenicity.

Discussion/Conclusions

The authors adequately summarized the main findings and provided sufficient information to have confidence in their conclusions. They related their findings to previous available research (given the time period) on VA, acetaldehyde, and acetic acid. They address the differences in bioassay findings of VA vs acetealdehyde. They provided a reasoned argument for the increase sensitivity of olfactory tissue vs respiratory epithelium (olfactory tissue metabolizes VA with greater efficiency leading to cytotoxicity). They explore alternative hypotheses. They discussed dose response implications and species differences (i.e., PK and architectural differences in the nasal cavity, and the fact that rodents are obligate breathers vs humans).

ANSWER EXPLANATIONS, CONTINUED

Expert 3

	Confidence Rating
Methods	
Results	
Discussion/Conclusions	

No explanation provided.

Expert 4 | SciTrust Score: 10

	Confidence Rating
Methods	10 (highest confidence)
Results	10 (highest confidence)
Discussion/Conclusions	10 (highest confidence)

Methods
Methds were well presented in the text. The methods followed standard chronic toxicity evaluation procedures for an inhalation study.

Results
The results were clearly presented in the text.

Discussion/Conclusions
The Discussion of the results was properly presented.

ANSWER EXPLANATIONS, CONTINUED

Expert 5 | **SciTrust Score:** 8

	Confidence Rating
Methods	7
Results	9
Discussion/Conclusions	8

Methods

1) Score = 0.45. The SD rats and Swiss mice are standard fare. The previous work by Owen

2) Score = 0.45. 60 animals were used in each dose group.

3) Score=0.45. Yes, the 0 ppm control group also had 60 animals

4) Score=0.45. The study lasted 104 weeks, equivalent to a rodent lifespan.

5) Score=0.45. The air concentrations of 0, 50, 200, and 600 ppm were based on prior work by Owen that showed the 50 and 200 ppm were no-effect levels in mice and rats (p. 216)

6) Score=0.45. Previous work by Owen and Gage (cited) showed that the MTD was 1000-2000 ppm.

7) Score=0.35. The discussion provides a brief discussion of the differences between rodent and human breathing for risk assessment considerations and cites Heffner (1990) regarding progression to malignancy in humans. Hence, the relevance to human exposure was acknowledged but not stated explicitly.

8) Score=0.45. The test material was >99% pure.

9) Score=0.45. The chamber concentrations were measured at 15 minute intervals using a gas chromatograph. Acetaldehyde is produced from VA and expected in the chamber and test material. The concentrations of acetaldehyde were < 50ppm.

10) Score=0.45. As noted, the chamber concentrations were measured frequently.

11) NA – this is an inhalation study.

12) Score = 0.45. The tissues examined are listed in the 2nd column of p. 217 and goes on for 12 lines of text.

13) Score =0.45. The first pathology review was conducted by the CRO that performed the study, Hazelton. The second review by a Dutch organization was independent and more detailed (p. 217).

14) Score =0. No mention is made of the pathologists' knowledge of dose groups.

15) Score = 0.35. The Cochran-Armitage test and Fisher's exact test were used for trend and group comparisons respectively. No other details were provided.

16) Score=0. No mention was made of survival analysis or adjustment.

17) Score = 0.15. Whilst the mechanism of VA tumor formation is discussed (pp. 225, 227), no historical control comparison was done.

18) Score = 0.45. Table 2 give the occurrence of both benign and malignant tumors in the nose and lungs of rats. A single squamous cell carcinoma was observed in a male mouse in the 600ppm group and a single adenocarcinoma in the control group (p. 223).

19) Score = 0. Not mentioned in the paper.

20) Score = 0.4. The study seems to comport with established guidelines even though these aren't mentioned.

21) Score=0.4. No deviations from guidelines seemed to occur.

Results

1) Score = 1. No disease outbreaks

2) Score= 0.8. No increase in mortality occurred in rats or mice associated with VA exposure. Survival analysis was conducted but few details were provided.

3) Score = 1. No The high dose of 600ppm was selected based on prior work.

4) NA

5) Score= 0.9. Tables were made for rats but mice had too few tumors to warrant such.

6) Score=1. Table 2 provides a good summary of the tumor incidence

7) Score=1. Yes, the finding in both rats and mice, benign and malignant changes were reported (Tables 1,2,3)

8) Score = 1. Significant reduction in weight gain occurred in both mice and rats (p. 218, Figs. 1, 2)

9) Score =0.8. No diseases were mentioned.

10) Score=1. Table 1 indicates olfactory squamous metaplasia, and basal cell hyperplasia in rats and nasal squamous metaplasia in mice.

Discussion/Conclusions

1) Score=1.2. The conclusions are well supported by the data.

2) Score=1.2. The discussion on p. 227 about VA and acetaldehyde is very good in helping the reader to understand both non-neoplastic and neoplastic effects observed.

3) Score=1.2. The study had sufficient power to detect effects.

4) Score=1.2. This well-conducted study had few if any limitations

5) Score=1.2. The discussion on mechanisms in the discussion (pp.223-7) were helpful.

6) Score=1.2. The primary uncertainty is the role of acetaldehyde and this is considered in the discussion.

7) Score=1.2. Yes

Debate | 1 comment

Expert 1 | 10/10/2026 5:56 SCORE: **0**

looks like we all are on the same page

2 Replies

Expert 2 | 10/10/2026 8:42 SCORE: **0**

I agree, and the scoring (although slightly different) reflects this.

Expert 5 | 10/10/2026 9:18 SCORE: **0**

Agree

SciPinion

SciPi #1052

October 19, 2026

Round 1: Bogdanffy et al. 1994 (drinking water)

SciPinion L.L.C.
777 E. Main St., Suite 202
Bozeman, MT 59715

shays@scipinion.com
ckirman@scipinion.com
(406) 404-6145

What are the strengths of Bogdanffy et al. (1994; drinking water) (if any)?

Results | 5 answers

Expert 1

well described study, high dose as MTD based on decrease in water and food consumption, reduced bw weight gains

3 comments

Expert 2 | 10/10/2026 8:46 SCORE: **0**

We all are in agreement.

Expert 5 | 10/10/2026 9:19 SCORE: **0**

Agree

Expert 4 | 10/11/2026 12:03 SCORE: **0**

agree

Expert 2

- Detailed description of methodology including pathology,
- Animals assessed for health, vaccinated against Sendai virus to prevent infection compromising study,
- Adequate group sizes,
- Multiple dose levels (200, 1000, 5000 ppm), and adequate exposure duration,
- Sprague-Dawley rats used, and F0 and F1 exposure (from conception and throughout prenatal development),
- F0 generation rats were selected by a randomization procedure,
- Each lot of the test material was evaluated for purity,
- Drinking water formulations frequency replaced (prepared daily and overformulated by 5% for each dose level to account for decomposition),
- Stability of VA in drinking water determined during 2nd week and every 4 weeks thereafter,
- Body weight, food and water consumption measured and data provided,
- Hematological, clinical observations and chemistry, urinalysis parameters and organ weights measured,
- Appropriate statistical methods including to account for differences in survival
- Tables of growth curves, food and water consumption, organ weight changes, and tumor data provided.

Expert 3

I cannot remember seeing a cancer bioassay designed and executed better than this one, and I have reviewed many.

4 comments

Expert 5 | 10/10/2026 9:19 SCORE: **0**

Agree

Expert 5 | 10/11/2026 8:16 SCORE: **0**

Agree, a very good study

Expert 4 | 10/11/2026 12:03 SCORE: **0**

agree

Expert 2 | 10/13/2026 9:18 SCORE: **0**

agree

Expert 4

> The research question was well defined. The experimental approach address this question. The methods employed are clearly presented in the manuscript. Particular attention was paid to the stability of the VAM in the drinking water. Analytical evaluation of the composition and stability of the VAM in the drinking water was a strength of the protocol. The results and discussion sections were well written and the statistical analysis appears appropriate.

Expert 5

> This was a well-conducted study that turned out negative. The tumor rates of the highest dose group were in the range of historical controls.

Experts were unanimous in judging Bogdanffy et al. (1994) a high-quality, well-conducted drinking water bioassay, with no dissenting views expressed. One expert characterized it as among the best-designed and executed cancer bioassays they had reviewed, a view explicitly endorsed by all other respondents.

Commonly cited strengths included:

• Study design:adequate group sizes, multiple dose levels (200, 1000, 5000 ppm), adequate exposure duration, randomization of F0 animals, and exposure spanning F0 and F1 generations (from conception through prenatal development).
• Dose selection:the high dose was considered an appropriate MTD based on reduced water and food consumption and decreased body weight gain.
• Test material and dosing rigor:purity testing of each lot, daily preparation with 5% overformulation to offset decomposition, and periodic verification of vinyl acetate stability in drinking water — repeatedly singled out as a particular strength.
• Endpoints and reporting:comprehensive hematology, clinical chemistry, urinalysis, organ weights, pathology, and clearly presented data tables; appropriate statistical methods, including adjustment for survival differences.
• Clarity:a well-defined research question, clearly described methods, and well-written results and discussion.

One expert additionally noted that the study was negative and that tumor rates in the highest dose group fell within the historical control range — an interpretive point that was not contested by other respondents.

AI generated summary content

What are the weaknesses of Bogdanffy et al. (1994; drinking water) (if any)?

Results | 5 answers

Expert 1

absence of irritation in oral cavity ? 2 tumors in oral cavity ? test item related?

2 comments

Expert 2 | 10/10/2026 8:58 SCORE: **1**

Agree, It is questionable whether these oral tumors are treatment related. They are observed at different times and the earlier tumor originated from lower jaw/palate/gingiva (so issue raised whether related to tooth lesions). The concurrent control was 0, but the authors state they were within historical control. Too bad, they didn't provide the historical range for the laboratory. My tendency is to conclude not treatment related.

1 Reply

Expert 1 | 10/10/2026 9:27 SCORE: **0**

reasonable conclusion

Expert 5 | 10/11/2026 8:22 SCORE: **0**

Agree, not treatment-related.

Expert 2

There are no weaknesses that would impact confidence or discount the conclusions of the study.

ATSDR concluded that the Bogdanffy et al. rat drinking-water study “may not have achieved the maximum tolerated dose necessary to evaluate potential carcinogenic effects of vinyl acetate,”. It should be noted that selecting an MTD dose can be challenging based on shorter term studies (typically 90 days), and to avoid overt toxicity and overwhelming repair/detoxification pathways. The 5000 ppm is an acceptable MTD given decreases in absolute BW. The highest dose was associated with a reduced body weight gain (11% first year and 17% second year), decrease food consumption (not statistically significant) and significant dose-related decreases in water consumption. The BW effect is likely attributed to reduced intake/palatability. There were decreases in organ weights but these were considered to be related to the changes in body weight.

1 comment

Expert 5 | 10/11/2026 8:21 SCORE: **0**

NCI uses the term "unacceptable side effects" to define the MTD. IMO, squishy. Do decreases in BW mean the rats don't feel well and thus don't eat as much. I think one could argue either way whether the decreases in BW were "unacceptable.

1 Reply

Expert 2 | 10/12/2026 8:14 SCORE: **0**

Agree

Expert 3

It was not immediately apparent that histopathology was performed blindly nor subjected to an independent pathology review.

3 comments

Expert 2 | 10/10/2026 9:25 SCORE: **0**

True, there were no statements in the publication. But Hazleton UK (which we do not have but likely available) conducted the study and very likely followed GLP and internationally accepted procedures for pathology.

Expert 5 | 10/10/2026 9:26 SCORE: **1**

The paper states on p. 217: "Two histopathological evaluations were conducted. The first was conducted at Hazleton Laboratories, UK. Following the initial review by Hazleton UK, a second, independent, and more detailed review of the respiratory tract tissues was conducted at TNO-CIVO laboratories in the Netherlands." This is buried in the tiny print the journal used for Methods.

Expert 4 | 10/11/2026 12:06 SCORE: **1**

I was less concerned about the proper histopathology procedures, given two different groups of pathologists examined the histopath.

Expert 4

No apparent weaknesses

Expert 5

The one weakness is the absence of an even higher dose. 20000ppm or thereabouts might have produced some results.

The authors did not consider possible strain differences as a reason for the negative result in the bioassay.

1 comment

Expert 2 | 10/10/2026 9:20 SCORE: **0**

A higher dose could plausibly produce results given that the oral cavity mucosa has carboxylesterase activity leading to VA hydrolysis and local intracellular acidity/cytoxicity.

Should the authors have tested higher doses? Maybe? But, they were seeing BW loss and palatability issues which can be dose-limiting. It appears they reached a maximum tolerated exposure. And their HDT is far above human drinking water exposures. And going to higher doses could saturate detoxification/repair systems.

From the expert comments I think there is agreement that 5000 ppm is a NOAEL.

1 Reply

Expert 5 | 10/11/2026 8:24 SCORE: **0**

5000 ppm is a reasonable choice for the NOAEL given the effect on BWW and that the study was negative for tumors.

Experts largely regarded the Bogdanffy et al. (1994) drinking-water study as robust, with no weaknesses substantial enough to undermine confidence in its conclusions. Two experts explicitly stated there were no apparent weaknesses, and the remaining critiques were characterized as minor or resolvable.

• Adequacy of the top dose (main point of discussion):One expert argued the principal weakness was the absence of a higher dose (e.g., ~20,000 ppm), echoing ATSDR's concern that an MTD may not have been reached. Others countered that reduced body weight gain (11–17%), decreased food intake, and dose-related reductions in water consumption indicate a maximum tolerated exposure was effectively achieved, and that higher doses could saturate detoxification/repair pathways. Experts acknowledged the MTD definition ("unacceptable side effects") is inherently subjective and could be argued either way. Consensus emerged that 5000 ppm is a reasonable NOAEL.
• Oral cavity tumors:Broad agreement that the two oral cavity tumors were likelynottreatment-related, given their differing times of occurrence, tissue origin, and the absence of oral irritation—though experts noted the paper's failure to provide the laboratory's historical control range as a shortcoming.
• Histopathology quality:One expert questioned whether histopathology was blinded or independently reviewed; this concern was largely resolved when another expert pointed to text (in small print in the Methods) documenting a second, independent review at TNO-CIVO following the initial Hazleton UK evaluation. Others noted the study was likely GLP-compliant.
• Minor unresolved point:One expert noted the authors did not consider possible strain differences as an explanation for the negative bioassay result—a comment that drew no rebuttal or further discussion.

AI generated summary content

Please rate your degree of confidence in each of the following study components for Bogdanffy et al. (1994; drinking water) (0-10 scale; 0 = fundamentally flawed; 1=lowest confidence; 5=medium confidence; 10=highest confidence). An overall confidence score will be calculated as the geometric mean of your three ratings. Please provide your rationale for each study component in the box beside it.

Results | 5 answers

	Methods	Results	Discussion/ Conclusions	SciTrust Score
Expert 2	9	8	8	8.3
Expert 1	8	7	6	7
Expert 5	5	9	9	7.4
Expert 4	10	10	8	9.3
Expert 3	10	10	10	10
Avg	8.4	8.8	8.2	8.4
Range	5–10	7–10	6–10	7–10

0 = no confidence · 5 = medium confidence · 10 = highest confidence

SciTrust Score = geometric mean of a reviewer's category scores. A zero in any category zeroes out that reviewer's full score. The **Final SciTrust Score** is the mean of all reviewer SciTrust Scores.

Experts expressed broadly high confidence in Bogdanffy et al. (1994), with all five agreeing that the study was well designed, GLP-compliant, and that its negative findings for systemic toxicity and oncogenicity are supported by the data. Particular praise was given to the careful handling of vinyl acetate stability in drinking water (overformulation at 105%, daily fresh preparation, periodic assay), adequate group sizes with satellite/interim sacrifice groups, appropriate dose spacing, in utero and lactational exposure, and 104-week duration.

Ratings nonetheless diverged substantially, particularly for Methods (range 5–10):

• Expert 3andExpert 4rated the study at or near the ceiling (10/10 on Methods and Results), viewing descriptions, necropsy, statistics, and reporting as fully adequate.
• Expert 5gave the lowest Methods score (5) but the highest Results and Discussion scores (9), explaining in debate that the low score arose from mechanically weighting all 21 checklist items equally—penalizing omissions such as no MTD discussion, no statement of human relevance for water ingestion, no cited guidelines, and no mention of pathologist blinding—while still calling it "an excellent paper." Expert 3 pushed back that the items should not be treated as equally important.
• Expert 2scored high (9/8/8) but deducted mainly because only one species was tested, arguing a second species would strengthen the no-carcinogenicity conclusion; Expert 3 countered that the paper should be judged as it stands rather than against a regulatory submission package.
• Expert 1gave the lowest Discussion/Conclusions score (6), citing a brief discussion that failed to address the two oral cavity squamous carcinomas in high-dose F1 males and the shorter observation period relative to the Ramazzini Institute design.

Common criticisms across reviewers included the lack of documented pathologist blinding or independent pathology review, limited statistical detail, and incomplete discussion of comparator studies—Expert 1 and Expert 2 focused on the unaddressed oral cavity tumors (Expert 2 noting these are biologically plausible given carboxylesterase-mediated hydrolysis in oral mucosa, though within historical control range), while Expert 4 noted the absence of any discussion of Maltoni et al. Expert 2 also felt the choice of 5000 ppm as the top dose warranted justification.

In debate, consensus converged on the view that this is a high-quality study whose limitations are largely reporting omissions rather than design flaws; Expert 3 conceded a rating of 10 may have been "a bit exuberant," while Expert 5 and Expert 1 affirmed the study's underlying quality despite their lower component scores.

AI generated summary content

ANSWER EXPLANATIONS

Expert 1 | **SciTrust Score:** 7

	Confidence Rating
Methods	8
Results	7
Discussion/Conclusions	6

Methods

Appropriate species (score of 10). SD rats were widely used for carcinogenicity studies, obtained from established breeder. Source of diet given, adeqaute acclimatisation, health assessment
No. of animals/group > as in guidance (score 10). Additional satellite groups included for clinical chemistry. Health status assessed, clinical chem. and urine analysis included at interim sacs
Controls included and exposed to water only (score of 10), sex distribution as in guidance.
Study duration (score of 9). Exposure of F1 for 104 weeks as in guidance, study performed in well established labs under GLP ? Parental animals of F1 were exposed to VAM for 10 weeks before conception and during lactation ?
Number and spacing of dose levels (score of 10), controls and three target concentrations, justification for conc. based on results from previous study with VAM.
Rote of exposure relevant to humans (score of 7), exposure to VAM from food contact material
Test item preparation, source, verification (score of 10). Detailed description of stability of VAM in water, fresh solutions prepared daily, overformulation of 5 % to compensate for decomposition; vendor named, composition and contaminants given
Necroscopy, histopatho (score of 8). Organ weights assessed, body weights, drinking water consumtion determined to calculate doses. Organs selected for histopath similar to guidance, apparently all animals including early death subjected to histopatho.. Blinding of evaluating pathologists not stated
Tumor, stats (score of 8). Methods appropriate, but no tumors found, no information on cytox in oral cavity and upper GI tract presented. Not present ?
Exposure period according to guideline (score of 9) and including conception and lactation, adequate randomisation at start, no animals/dose group according to guidance, terminal sacrifice after 104 weeks of exposure.

Results

organ weights, body weight gain etc all described in detail, statement that no tumors were found contrast between discussion and results

Discussion/Conclusions

discussion is quite short due to limited effects seen. Discussion of the two oral cavity tumors and their relation to test item admin. should be included.

is the absence of oral tumors in this study due to a shorter time of observation as compared to the study design of Ramazzini institute ?

ANSWER EXPLANATIONS, CONTINUED

Expert 2 | **SciTrust Score:** 8.3

	Confidence Rating
Methods	9
Results	8
Discussion/Conclusions	8

Methods

This published study is well designed and conducted. Importantly the authors were attendant to the issue of VA stability in drinking water which is critical since decomposition of VA can result in significant quantities of breakdown products, including AA and acetic acid complicating the interpretation of findings.

Only effects on F1's were reported. This is a publication which would not have the detail as a laboratory report. It is likely that historical control data, raw data for controls and all animal treatment groups, and more methodological details could be obtain by requesting the Hazleton UK report.

Results

There is sufficient data presented in the article to support the authors conclusions based on the doses evaluated. The authors did not find evidence of carcinogenicity at the highest dose tested.

There was an observation of two squamous carcinomas in the oral cavity (which are uncommon) in high dose F1 males (one of which was observed in an unscheduled death and the other on terminal sacrifice). The authors indicated that these responses were within the labs historical control range. Thus, they were considered these to be chance events and related to tooth lesions not treatment related.

Oral cavity tumors are biologically plausible based on VA's mode of action. Bogdanffy was a co-author on a 2002 paper (Robinson DA, Bogdanffy MS, Reed CJ. Toxicology. 2002 Dec 2;180(3):209-20) that investigated carboxylesterase activity in rat and mouse oral cavity mucosa. The authors concluded that the oral cavity mucosa has the potential to hydrolyse VA to its metabolites, acetic acid and AA.

Nonetheless, 5000 ppm in the Bogdanffy et al study establishes a NOAEL for the absence of toxicity and carcinogenicity.

Discussion/Conclusions

The authors conclusions are supported by the data presented. They adequately summarized the main findings. They relate their findings to previous work in supportive of their conclusions, as well as addressing differences in other chronic bioassay DW findings. The authors imply the that the lack of systemic toxicity could be explained by the rapid hydrolysis of VA.

The authors could have discuss their choice of 5000 ppm as the highest dose tested. Moreover, they could have explained why they chose to evaluate rats only and not a second species (mice). Their inhalation study was 2 species.

ANSWER EXPLANATIONS, CONTINUED

Expert 3 | **SciTrust Score:** 10

	Confidence Rating
Methods	10 (highest confidence)
Results	10 (highest confidence)
Discussion/Conclusions	10 (highest confidence)

Methods

An appropriate species was selected with justification. The number of animals per dose group was more than sufficient, allowing interim sacrifices. The study duration was typical near the lifespan of the animal but not so far as to have complications in reviewing tumors against control animals that live their full lifespan. The three doses and control were appropriate and the length of exposure, including *in utero*, was relevant. The methods were well described with the test substance appropriately characterized for purity and stability; an extra effort was made to account for the expected breakdown of the product in water. The documentation of actual versus administered dose was handled appropriately, and food and water were measured per dose group. Animals were vaccinated against a common laboratory rat pathogen. A complete necropsy was performed on animals including those that died early, and collection of tissues appeared to be processed and examined according to standard protocols. It was not immediately apparent that histopathology was performed blindly nor subjected to an independent pathology review. The statistics were appropriate. The study met good laboratory practices.

Results

There did not appear to be any disease outbreaks in the experimental and control animals. Survival was adequate across those groups to allow a meaningful comparison of tumors. There was no apparent excessive toxicity at high doses or reduced water our feed intake that would affect results scene. The data were appropriately summarized, and tumor incidences were appropriatelly reported by group, sex and site. The findings appeared to be reported completely.

Discussion/Conclusions

The authors conclusions were supported by the data presented. The conclusion regarding the lack of carcinogenicity appear was appropriately justified. The limitations of the study were a few but identified and uncertainty in the data or interpretation were acknowledged. The language of the conclusions accurately reflects the strength of the evidence.

ANSWER EXPLANATIONS, CONTINUED

Expert 4 | **SciTrust Score:** 9.3

	Confidence Rating
Methods	10 (highest confidence)
Results	10 (highest confidence)
Discussion/Conclusions	8

Methods
The methods were clearly described and defined in the manuscript and appear scientifically valid.

Results
The results were properly presented

Discussion/Conclusions
The discussion was adequate and discussed the contrasting results from the authors finding in this manuscript with the Lijinsky and Reuber paper. Curiously there was no mention or discussion of the Maltoni et al paper that was published in 1997.

ANSWER EXPLANATIONS, CONTINUED

Expert 5 | **SciTrust Score:** 7.4

	Confidence Rating
Methods	5 (medium confidence)
Results	9
Discussion/Conclusions	9

Methods

1) Score=0.45. The rats were obtained from Charles River for hardiness, longevity and low background disease rate.

2) Score=0.2. The study duration was nominally 104 weeks with 60 animals per dose group and satellite groups for interim sacrifice and 52 and 72 weeks. The rats were bred to produce F1 rats but no distinction was made between these in the presentation of the results.

3) Score=0.5. The control group was also 60 animals, hence, an adequate size.

4) Score=0.45. The study duration was 104 weeks.

5) Score=0.45. A constant 5-fold difference between drinking water concentrations was appropriate. The levels were 0, 200, 1000 and 5000 ppm.

6) Score=0. No discussion or identification of the MTD was provided.

7) Score=0. No discussion was given as to the relevance of water ingestion for humans.

8) Score=0.35. The VA was tested for purity and an evaluation of the stability of VA in water was conducted. Based on this the solutions were made at 105% of the required VA and prepared fresh daily. The drinking water was assayed at 4 week intervals.

9) Score=0. The dose formulations were all within 20% of the initial concentration.

10) Score=0.45. Water consumption was measured and presented in Table 2. From these, doses of 10, 47 and 202 mg/kg/d were calculated for the 0, 1000 and 5000 ppm groups (p. 208).

11) Score=0.45. Water and food consumption were measuredfor each dose group and the results shown in Tables 1 and 2.

12) Score =0.35. The list of tissue on p. 208 suggest a complete necropsy but no guidelines were cited.

13) Score=0. The tissue processing seemed appropriate but no guidelines were cited.

14) Score=0. No mention of pathologist blinding occurred.

15) Score=0.15. The brief description of the statistics used suggest the tests were appropriate for the time (1994). Details were lacking.

16) Score=0.25. Survival analysis was conducted using the Kaplan-Meier method. Survival percents are shown on p. 209.

17) Score=0.4. P. 210 indicates the carcinoma incidence in qll dose groups was within the range of historical controls.

18) Score=0.35. Table 4 shows both the neoplastic and non-neoplastic effects.

19) Score =0. No mention was made of multiple comparisons in the statistics descrption.

20) Score=0. No guidelines cited.

21) Score=0. As noted, no guidelines cited.

Results

1) Score=1. No infectious disease outbreaks occurs and the rats were vaxed and examined by a vet before the study started.

2) Score=0.8. Survival was adequate for meaningful comparisons.

3) Score=1. The highest dose was 5000ppm, lower than some of the studies andwould not likely cause excessive toxicity. The largest survival percentages were in the 5000ppm group (p. 209.

4) Score=1. Water consumption was lower in both males and females in the high dose group (Table 2) and appeared to be the one of the primary effects (p. 214).

5) Score=0.7. The data are summarized and easy to understand in Table 3.

6) Score=1. Table 4 provides tumor incidence by dose group for both M and F.

7) Score=0.8. All findings appeared to be reported. The study was generally negative as tumor incidence in all dose groups was within the range of historical controls.

8) Score=1. Fig. 1 and Table 3 show the group differences in body weight. The high dose group had non-significant lowering of body weight from about 40 weeks on. These did not seem to influence the results.

9) Score=1. No incurrent disease in the vaxed and hardy rats.

10) Score=1. The occurrence of both non-neoplastic and neoplastic effects was rare in all dose groups.

Discussion/Conclusions

1) Score=1.4. The conclusion that “No systemic toxicity or oncogemcity were caused by vinyl acetate administration over the normal life span of the rats is consistent with data.
2) Score=1.4. The negative conclusions of the study are indeed justified in the discussion.
3) Score=1.4. The study did indeed have sufficient power, exposure and duration to detect an effect.
4) Score=1. The discussion of the “hardiness” of the strain as protective against the VA effects , albeit unlikely, was not mentioned.
5) Score=1.2. The limitations of the study are identified.
6) Score=1.2. The uncertainties in the study are acknowledged and discussed.
7) Score=1.2. The language does indeed indicate the strength of the evidence.

Debate | 3 comments

Expert 2 | 10/10/2026 9:37 SCORE: **0**

Generally agree with the comments. I did considered this to be a GLP compliant study and would put high confidence in its conduct and results. I largely gave a lower score because the standard practice is two species for a cancer bioassay. Having two species would strengthen the conclusion of a lack of carcinogenicity including for the systemic tumors. The authors could have explained their choice of rats only.

1 Reply

Expert 3 | 10/14/2026 9:14 SCORE: **0**

Expert 2 is correct that a standard submission to a regulatory agency would be two species. However, this is a stand alone paper and even though it is likely to be part of a submission package, should be evaluated as it stands.

Expert 5 | 10/10/2026 9:37 SCORE: **0**

I gave the lowest methods score based on the 21 questions in the instructions. The authors did a good job of getting a lot of information in this paper. Maybe their choices to omit discussion of the MTD, human relevance, adherence to guidelines and pathologist blinding are reasonable choices to keep the paper appropriately brief. That said, a set of questions was provided and I used those in a way that considered each question/aspect of the paper equally important.

Even with these omissions, this was an excellent paper.

1 Reply

Expert 3 | 10/14/2026 9:16 SCORE: **0**

Hmm... Not all 21 questions should be treated equally, but I may be argued down from a rating of 10.

Expert 1 | 10/12/2026 2:43 SCORE: **0**

Agree that study is of high quality, discussion is somewhat limited and thus I gave a low score

1 Reply

Expert 3 | 10/14/2026 9:23 SCORE: **0**

Hmm... Well again, I acknowledge that my rating of 10 might be a bit exuberant, but I would need a bit more explanation from expert #1 to lower it.

SciPinion

SciPi #1052

October 19, 2026

Round 1: Lijinsky and Reuber 1983

SciPinion L.L.C.
777 E. Main St., Suite 202
Bozeman, MT 59715

shays@scipinion.com
ckirman@scipinion.com
(406) 404-6145

What are the strengths of Lijinsky and Reuber (1983) (if any)?

Results | 5 answers

Expert 1

> Study performed according to standards in the late 1970s with limited resources, can only be used as descriptive. This is acknowleged by authors.

1 comment

Expert 5 | 10/11/2026 8:26 SCORE: **0**

Agree

Expert 4

> The only strength of this paper is that it appears to be the first attempt to look at the chronic effects of VAM in rodents. The reported design and performance of the bioassay reflects the state of the chronic bioassay approaches during the early 1980s at the NCI (limited number of animals, poor statistics, lack of scheduled sampling times). Also, frequently at this time period the NCI chronic studies shared control animals with other chronic studies which compounds interpretation of the results.

1 comment

Expert 5 | 10/11/2026 8:27 SCORE: **0**

Agree

Expert 5

> Th sole strength is the authors' admission that the study was flawed.

1 comment

Expert 4 | 10/11/2026 12:07 SCORE: **0**

agree

Experts were unanimous that Lijinsky and Reuber (1983) offers little of value beyond historical and descriptive interest. All three respondents converged on the same core point: the study's principal "strength" is the authors' own acknowledgement of its limitations.

• Historical context:Experts agreed the work reflects the standards, resources, and methodological conventions of chronic bioassays conducted at the NCI in the late 1970s and early 1980s.
• Priority:One expert noted it appears to be the first attempt to examine chronic effects of VAM in rodents, giving it some value as a pioneering effort.
• Candour about flaws:Multiple experts highlighted the authors' admission that the study was flawed as its sole genuine merit.
• Acknowledged weaknesses:Limited animal numbers, poor statistics, lack of scheduled sampling times, and the then-common practice of sharing control animals across studies were cited as confounding interpretation.

There was no disagreement among respondents; each explicitly endorsed the others' assessments, and the consensus is that the paper can only be used descriptively rather than as a basis for quantitative or regulatory conclusions.

AI generated summary content

What are the weaknesses of Lijinsky and Reuber (1983) (if any)?

Results | 5 answers

Expert 1

limited number of animals, study design makes conclusions difficult (no defined point of sac), lots of "guessing" in manuscript regarding doses received, animals likely exposed to significant amounts of decomposition products, no information on cytox in oral cavity ?

1 comment

Expert 5 | 10/12/2026 8:14 SCORE: **0**

It seems to me the authors weren't really doing a useable study but rather setting the stage for an adequate study

Expert 4

There are multiple weaknesses of this manuscript
1. The experimental design included only 20 rats per group.
2. No statistical analysis was performed on any of the results
3. The drinking water administered to the test animals was not adequately characterized. The authors state that instead of actually measuring the concentration of VAM in the water they "believed each animal received at least half the administered dose".
3. The dose selection was poorly described and not adequate. The authors stated that because of inadequate resources preliminary dose range finding studies were not performed. The doses of 2500 and 1000 ppm were selected. Treatment was for 5 days on and 2 days off. No explanation for the interruption in treatment was provided. Since this was a drinking water study, it is not clear why treatment was not continuously for 7 days.
4. Instead of defined sampling times (sacrifice times) rats were treated for 2 years and sampled after "natural death ' unless moribund. This complicates the analysis of the histopathology and thus pathology evaluation given the resulting deterioration of the tissues after death.
5. IMPORTANT
There appears to be a disconnect between the reported increase in treatment related tumors reported in this manuscript and the pathology report by EPL (1982) of the tissues from this study. The EPL pathology report stated " *A variety of neoplastic and nonneoplastic lesions were present in male and female control rats and in rats treated with Vinyl Acetate. The incidence of neoplastic nodules of the liver in male and female rats and neoplasms of the uterus in female rats was greater in treated groups when compared with control groups. NO treatment-related differences in the incidence, distribution and severity of other lesions were observed among control and treated groups. These lesions had a similar incidence and distribution of those known to occur spontaneously in untreated Fischer 344 rats of similar age*". Thus, in contrast to that reported in the manuscript, the EPL pathology report found no increase in thyroid tumors with treatment. Uterine tumors reported in the manuscript were confirmed in the VAM treated rats by the EPL pathology evaluation. In contrast to the manuscript the EPL pathology evaluation noted the presence of treatment related neoplastic nodules (probably liver adenomas using today's criteria) in both male and female treated rats. These discrepancies in histopathology evaluations are disconcerting.

4 comments

Expert 1 | 10/10/2026 6:00 SCORE: **0**

Thanks for comments on the issue of thyroid tumors

Expert 5 | 10/11/2026 8:41 SCORE: **0**

I took a quick look at the EPL report and noted the absence of any clear dose response in any tissue. Maybe such wasn't possible given <20 animals/group examined.

Expert 4 | 10/11/2026 12:19 SCORE: **0**

The EPL report also noted the incomplete nature of the slides (tissues) sent to EPL
The conclusions were " Conclusions concerning the incidence of neoplasms in many of the other organ systems could not be made because in many instances only tissues from a few of the treated male and female rats were available Before definitive conclusions concerning the effect of Vinyl Acetate on these organs can be made, be prepared for histopathologic examination
I think the emphasizes the poor process of this study at the NCI. Jerrry Hardisty is an exceptional pathologist and i will default to his evaluation over that of the authors of the paper.

Expert 5 | 10/12/2026 9:51 SCORE: **0**

Thanks for this useful comment. I'm not a pathologist.

Expert 5

| THe small number of animals and decomposition of VA in the feeding bottles.

Experts were broadly aligned in judging the Lijinsky and Reuber (1983) study to be methodologically weak, with no dissenting views expressed.

• Sample size:All respondents flagged the small group sizes (20 rats/group) as a core limitation, with one expert noting this likely precluded detection of any clear dose-response relationship.
• Dosing uncertainty:Multiple experts criticized the lack of analytical verification of vinyl acetate concentrations in drinking water, the authors' reliance on the assumption that animals received "at least half" the intended dose, and likely decomposition of the test material in the feeding bottles — raising the possibility of exposure to decomposition products.
• Study design:Shared concerns included the absence of defined sacrifice times (reliance on "natural death"), which complicates histopathology because of tissue autolysis, as well as inadequate dose selection without range-finding studies, an unexplained 5-days-on/2-days-off regimen, and no statistical analysis.
• Pathology discrepancies:One expert raised, and others endorsed, a significant conflict between the published tumor findings (e.g., thyroid tumors) and the EPL (1982) pathology re-evaluation, which instead reported liver neoplastic nodules and incomplete tissue submission. This expert deferred to the EPL pathologist's assessment over the authors'; another acknowledged the point as useful while noting they are not a pathologist.
• Missing data:One expert additionally noted the absence of information on cytotoxicity in the oral cavity.

One expert offered the summarizing view that the work appears less like a usable study and more like groundwork for an adequate one, a characterization not challenged by others.

AI generated summary content

Please rate your degree of confidence in each of the following study components for Lijinsky and Reuber (1983) (0-10 scale; 0 = fundamentally flawed; 1=lowest confidence; 5=medium confidence; 10=highest confidence). An overall confidence score will be calculated as the geometric mean of your three ratings. Please provide your rationale for each study component in the box beside it.

Results | 5 answers

	Methods	Results	Discussion/ Conclusions	SciTrust Score
Expert 1	4	4	4	4
Expert 5	2	2	6	2.9
Expert 4	3	2	2	2.3
Avg	3	2.7	4	3.1
Range	2–4	2–4	2–6	2.3–4

0 = no confidence · 5 = medium confidence · 10 = highest confidence

SciTrust Score = geometric mean of a reviewer's category scores. A zero in any category zeroes out that reviewer's full score. The **Final SciTrust Score** is the mean of all reviewer SciTrust Scores.

All three experts expressed low confidence in Lijinsky and Reuber (1983), with ratings clustering in the 2–4 range across most components and broad agreement on the study's core deficiencies.

Points of agreement:

• Inadequate sample size— all experts flagged the ~20 animals per dose group as far below guideline expectations.
• Absence of statistical analysis— unanimously cited as a fundamental limitation of both methods and results.
• Poor characterization and dosing of test material— VAM decomposition in drinking water was acknowledged but never quantified; actual delivered doses cannot be determined, and water intake was only estimated.
• Limited discussion— experts noted the authors largely restated results and offered no mechanistic interpretation, though several acknowledged the authors' own recognition of the study's flaws.

Points of divergence:

• TheDiscussion/Conclusionscomponent drew the widest spread: Expert 5 rated it relatively favorably (6), crediting the authors for transparently addressing study size and VAM decomposition, whereas Expert 4 rated it 2, emphasizing the absence of any mechanistic discussion for the uterine and thyroid tumors. Expert 1 took a middle position (4).
• Study durationwas interpreted differently — Expert 1 judged the 100-week exposure as shorter than guidance recommends, while Expert 5 considered the study to have run beyond 104 weeks.
• Human relevance of the exposure routewas scored comparatively well by Expert 1 (food-contact exposure), but Expert 5 found no consideration of human relevance at all.
• Only Expert 4 raised theexternal EPL pathology re-analysisthat failed to confirm the female thyroid tumor findings, casting additional doubt on the reported pathology.

Debate comments were brief and confirmatory, with experts explicitly noting that they were "on the same page," indicating that the differences above reflect emphasis rather than substantive disagreement about the study's overall limited reliability.

AI generated summary content

ANSWER EXPLANATIONS

Expert 1 | **SciTrust Score:** 4

	Confidence Rating
Methods	4
Results	4
Discussion/Conclusions	4

Methods

Appropriate species (score of 6). F344 rats were widely used for carcinogenicity studies, obtained from NCI colony. Source of diet not given, process of acclimatisation, health assessment not stated
No. of animals/group low, << in guidance (score 4).
Controls included and exposed to water only (score of 8).
Study duration. Exposure 100 weeks, < in guidance (score 2). Animals were not sacrificed at specific time point, but kept until natural death
Number and spacing of dose levels (score of 7), controls and two target concentrations.
Rote of exposure relevant to humans (score of 7), exposure to VAM from food contact material
Test item preparation, source, verification (score of 2). Limited description of stability of VAM in water, composition and contaminants not given, replacement of solutions did not account for kinetics of decomposition. Animals likely exposed to substantial amounts of decomposition products. Drinking water intake not determined, dose received cannot be determined. Estimated doses stated in manuscript do not make sense
Necroscopy, histopatho (score of 7). Organs selected for histopath similar to guidance, apparently all animals including early death subjected to histopatho.. Blinding of evaluating pathologists not stated.
Tumors, stats (score of 3). No statistical evaluation performed, information on cytox in oral cavity and upper GI tract not assessed/presented. Not present ?
Problem with study design and steep increase of incidence of tumors observed in aging rats make conclusions difficult (uterus, thyroid).

Results

Description of tumor histology is major point in results, tables show no. of animals with tumors, low number of animals/dose group, no stats. Tumor incidences in uterus and thyroid show some dose-dependence. In general, adequate descpription of the limited information presented. No statistical evaluation.

Discussion/Conclusions

Authors discuss presence of tumors in the exposed animals and relate these to background incidences in controls. Acknowledge limitations due to decomposition of VAM. No comments on potential effects of decomposition products. Conclusions generally limited, authors suggest performing another study with adequate numbers and improved study design

Expert 2

	Confidence Rating
Methods	
Results	
Discussion/Conclusions	

No explanation provided.

ANSWER EXPLANATIONS, CONTINUED

Expert 3

	Confidence Rating
Methods	
Results	
Discussion/Conclusions	

No explanation provided.

Expert 4 | SciTrust Score: 2.3

	Confidence Rating
Methods	3
Results	2
Discussion/Conclusions	2

Methods
As noted above, there are major issues with sample size, lack of statistics, determination of the composition and dosing of the VAM material.

Results
While the manuscript reports a response in the female rat for thyroid tumors, a companion analysis by EPL labs failed to show thyroid tumor leading concerns about the completeness and interpretation of the pathology reported in the manuscript. No statistics was performed or reported.

Discussion/Conclusions
The discussion is limited to repeating the findings of the results. This appears to be one of the earliest chronic studies with VAM and thus the lack of a discussion of VAM carcinogenicity in rats in the literature is not surprising. However there is no discussion of the possible mechanism(s) by which VAM induced the reported uterine and thyroid tumors.

ANSWER EXPLANATIONS, CONTINUED

Expert 5 | **SciTrust Score:** 2.9

	Confidence Rating
Methods	2
Results	2
Discussion/Conclusions	6

Methods

1) Score=0.1. No justification for using F344 rats was given

2) Score=0. 20 per dose group—too small.

3) Score=0.4. Control with 0 mg/L was conducted

4) Score=0.5. Study was conducted for > 104 weeks

5) Score =0.3. Both concentrations of VA were high and the reason for choosing 2500 mg/L is thin (p. 43).

6) Score=0. No, as stated by the authors

7) Score =0. No consideration of human relevance was provided.

8) Score=0.2. The authors noted that the VA decomposed in the feeding bottles but did not actually calculate doses.

9) Score=0. No

10) Score=0. No

11) Score=0.1. Water consumption was presented as an estimate

12) Score=0. Unclear from the brief description of methods for pathology.

13) Score=0. No

14) Score=0. Unclear

15) Score=0. No statistics at all.

16) Score=0. See 15

17) Score=0. No

18) Score=0.4. Table 2 separates the tumor types

19) Score=0. No

20) Score=0. No

21) Score=0. No

Results

1) Score=0. An infectious disease outbreak was not mentioned.

2) Score=0.2. Table 1 shows the survival data and by 100 weeks, some had died.

3) Score=0. No discussion of this issue was presented.

4) Score=0. Ongoing water intake was not measured, only estimated.

5) Score=0.5. Table 2 provides a reasonable summary.

6) Score=0.4. Yes. This is provided in Table 2.

7) Score=0.1. The authors seemed more concerned with the 6 uninterpretable figures

8) Score=0. Body weights of the rats were not measured.

9) Score=0.4. No incurrent disease affected the results.

10) Score=0. Pre-neoplastic lesions weren't mentioned.

No AI was used in these answers.

Discussion/Conclusions

1) Score=0.8. Only solid conclusion is that a better study be conducted (p.52).

2) Score=1. A reasonable summary of this tiny study was provided.

3) NA. The study was concluded to be positive.

4) Score=0.5. Wasn't meant to be a definitive study.

5) Score=0.9. the size of the study and the issue of VA decomp were considered

6) Score=1.2. The major flaws in the study are discussed at the end of the paper.

7) Score=1.2. The reporting of the tumors observed seemed reasonable.

Debate | 2 comments

Expert 1 | 10/10/2026 6:01 SCORE: **0**

seems that we all are on the same page

Expert 5 | 10/11/2026 8:43 SCORE: **0**

Agree with #1

SciPinion

SciPi #1052

October 19, 2026

Round 1: Maltoni et al. 1997

SciPinion L.L.C.
777 E. Main St., Suite 202
Bozeman, MT 59715

shays@scipinion.com
ckirman@scipinion.com
(406) 404-6145

What are the strengths of Maltoni et al. (1997) (if any)?

Results | 5 answers

Expert 1

> Longterm exposure of Swiss mice to VAM for 78 weeks, test item characterized, fresh preparations of water with test item generated daily, examination of relevant organs by histopatho.; app. controls included.
>
> Local tumors at port of entry consistently observed. The studies from Ramazinni may serve as descritive work confirmung local tumors at site of 1st contact

1 comment

Expert 5 | 10/12/2026 9:57 SCORE: **0**

I find it interesting that Maltoni mice showed a positive response for oral sq. ca. but the Bogdanffy CR rats did not. Also, the Belpoggi Wistar rats did show a positive response. Strain and species differences??

2 Replies

Expert 1 | 10/12/2026 10:10 SCORE: **0**

If the MoA assumptions are correct, I would not expect species and strain differences regarding the tumors in the oral cavity, Local effect, toxicokinetics not a confounder. Reasonable explanation could be inadequate histopath, differences in VAM degradation result in different exposures ?

Expert 5 | 10/14/2026 8:16 SCORE: **0**

I think the reasons #1 brought up are more plausible than strain and species differences.

Expert 3

> Full lifetime monitoring, although this is not seen to be helpful by many toxicologist due to problems associated with comparison to aging control animals.

1 comment

Expert 1 | 10/12/2026 2:48 SCORE: **0**

agree, this is the issue with all studies from the Ramazinni-foundation

Expert 4

> This is a poorly planned and presented paper describing a chronic study in Swiss mice treated with VAM. I am not sure there are any redeeming results from this study other than the body weight and survival data presented.

1 comment

Expert 5 | 10/11/2026 8:46 SCORE: **0**

Agree that the BW and survival data were informative.

Expert 5

> The paper has few if any strengths and the outcome that VAM causes cancer was pre-determined.

1 comment

Expert 4 | 10/11/2026 12:20 SCORE: **0**

agree

Experts were sharply divided on whether the Maltoni et al. (1997) study has meaningful strengths, with the majority view being largely critical.

• Most positive assessment:One expert identified several methodological strengths, including long-term (78-week) exposure of Swiss mice, characterisation of the test item, daily preparation of fresh dosing solutions, histopathological examination of relevant organs, and inclusion of appropriate controls. This expert considered the consistent observation of local tumours at the port of entry to be a useful descriptive confirmation of tumours at the site of first contact.
• Strongly negative views:Two experts saw few or no redeeming features, describing the paper as poorly planned and presented, with one asserting that the conclusion that VAM causes cancer appeared pre-determined. Both nonetheless agreed that the body weight and survival data were informative.
• Qualified strength:Another expert cited full lifetime monitoring as a strength, but immediately noted the widely recognised difficulty of interpreting results against aging control animals — a caveat endorsed by a colleague as a general limitation of all Ramazzini Foundation studies.

Point of discussion: Experts debated why Maltoni's mice showed oral squamous cell carcinomas while Bogdanffy's CR rats did not (though Belpoggi's Wistar rats did). There was convergence on the view that inadequate histopathology or differences in VAM degradation and resulting exposure are more plausible explanations than genuine species or strain differences, given that the mode of action implies a local effect with toxicokinetics unlikely to be a confounder.

AI generated summary content

What are the weaknesses of Maltoni et al. (1997) (if any)?

Results | 5 answers

Expert 1

Unusual study design keeping animals until natural death, no info on drinking water consumption, number of animals/group different from OECD/EPA guidance, little information on hygiene/infectious diseases in colony, no info on source of animals, source of test item not specified,
data presented in a large number of tables in unusual reporting format.

Studies from Ramazinni should not be used to conclude on the potential of VAM to induce systemic tumors.

2 comments

Expert 5 | 10/11/2026 8:48 SCORE: **0**

Agree about the general lack of utility of Ramazzini-conducted studies.

Expert 4 | 10/11/2026 12:21 SCORE: **0**

Agree with the concerns of using Ramazzini performed studies for concluding VAM carcinogenicity.

Expert 3

This bioassay has a number of significant flaws that preclude making definitive conclusions:
- Only partial utero exposure.
- Contaminants monitored but not discussed.
- No vaccination of experimental animals.
- No interim sacrifices.
- No statistical analysis
- Not GLP compliant (insufficient number of animals; no housing information)

1 comment

Expert 4 | 10/11/2026 12:22 SCORE: **0**

Agree, very good summation of a number of the flaws.

Expert 4

The weaknesses of this paper are multiple
1. The Methods section is incomplete and poorly presented
2. The material administered is contaminated with benzene 30-45 ppm, methyl and ethyl acetate 50 ppm, Acetone 330 –

> 500 ppm and select aldehydes. It is unclear if the administered drinking water was analyzed for the presence and these contaminants throughout the study. The concentration of the VAM in the drinking water was not apparently analyzed either. The drinking water source was not described. Was it tap water? distilled water?
> 3. No statistics were apparently performed.
> 4. The experimental design is not scientifically valid for a chronic/carcinogenicity study. No scheduled sampling of treated animals was performed. How can a comparison of the treated and control groups be made without scheduled sampling.
> 5. There are multiple additional problems with this study, most of which are related to the design and the performance. This is not a valid study to use for carcinogenic evaluation of VAM

1 comment

Expert 5 | 10/12/2026 10:00 SCORE: **0**

Agree

Expert 5

> Plenty. Too few animals, no statistical evaluation, unsupported claims related to the data, unsupported claims related to risk to humans.

1 comment

Expert 4 | 10/11/2026 12:21 SCORE: **0**

Agree

Strong consensus: All four responding experts judged Maltoni et al. (1997) to be seriously flawed and unsuitable as a basis for concluding that vinyl acetate monomer (VAM) causes systemic tumours. Several explicitly endorsed one another's critiques, and two extended their skepticism to Ramazzini Institute studies generally.

Recurring criticisms cited across responses included:

• No statistical analysisof the reported tumour data (Experts 5, 4, 3).
• Inadequate/non-guideline study design— too few animals per group relative to OECD/EPA expectations, no interim or scheduled sacrifices, and the unusual practice of holding animals until natural death (Experts 1, 5, 4, 3).
• Poor reporting of methods and husbandry— missing information on animal source, housing, hygiene/infectious disease status or vaccination, and an unconventional tabular presentation (Experts 1, 4, 3).
• Test material and exposure concerns— contamination of the test substance (benzene, acetates, acetone, aldehydes), lack of verification of VAM concentration in drinking water, undescribed water source, and no data on water consumption (Experts 1, 4, 3).
• Unsupported interpretive claimsabout the data and about human risk (Expert 5); partial in-utero exposure and lack of GLP compliance (Expert 3).

Points of divergence: Disagreement was minimal and largely a matter of emphasis rather than substance. Experts 1 and 5 framed the problem broadly, questioning the utility of Ramazzini-conducted bioassays as a class, whereas Experts 4 and 3 focused on study-specific deficiencies (contaminant analysis, exposure characterization, in-utero design) while still agreeing with the broader concern. Expert 4 placed particular weight on analytical verification of the dosing solutions, a point not raised by others.

AI generated summary content

Please rate your degree of confidence in each of the following study components for Maltoni et al. (1997) (0-10 scale; 0 = fundamentally flawed; 1=lowest confidence; 5=medium confidence; 10=highest confidence). An overall confidence score will be calculated as the geometric mean of your three ratings. Please provide your rationale for each study component in the box beside it.

Results | 5 answers

	Methods	Results	Discussion/ Conclusions	SciTrust Score
Expert 1	3	5	4	3.9
Expert 5	2	2	0	0
Expert 4	0	0	0	0
Expert 3	3	5	4	3.9
Avg	2	3	2	2
Range	0–3	0–5	0–4	0–3.9

0 = no confidence · 5 = medium confidence · 10 = highest confidence

SciTrust Score = geometric mean of a reviewer's category scores. A zero in any category zeroes out that reviewer's full score. The **Final SciTrust Score** is the mean of all reviewer SciTrust Scores.

Experts converged on the view that Maltoni et al. (1997) has substantial methodological weaknesses, but diverged sharply on whether those weaknesses render the study wholly unusable.

Areas of agreement:

• All four experts flagged the absence of any statistical analysis of tumor incidence or dose–response as a core deficiency.
• All noted that the concentration of VAM in drinking water was never verified analytically and that water consumption was not measured, so the delivered dose is unknown.
• Group sizes were widely judged insufficient, no rationale was provided for species or dose selection, and blinding/independent pathology review was not documented.
• Several experts observed that the study does not conform to established guidelines and that deviations were never discussed.

Areas of disagreement:

• Scores ranged from complete rejection (Expert 4: 0/0/0, arguing that a flawed design makes all downstream results "fruit from a poisoned tree") to moderate confidence in the results (Experts 1 and 3: 3/5/4).
• Expert 5 (2/2/0) sided closer to rejection, emphasizing that causal language in the discussion goes far beyond the data and that human-relevance claims were made without any exposure measurement.
• Experts 1 and 3 — who, as Expert 3 noted, reached identical ratings for slightly different reasons — held that the portal-of-entry tumors (oral cavity, upper GI tract) are a credible finding, with Expert 1 arguing this outcome is consistent with other VAM oral studies and unlikely to be undermined by the design flaws, even while rejecting the study for quantitative risk assessment.
• Survival/lifespan design was read differently: Expert 5 and Expert 1 treated the lifetime-to-natural-death approach as a confound (aging tumor background, tissue autolysis), whereas Expert 3 viewed survival as adequate for meaningful comparison.
• Experts differed on contaminants: Expert 3 highlighted benzene as a known carcinogen present in the test material, while Expert 1 considered purity reporting reasonable and raised acetaldehyde instead; Expert 4 said contaminants were not analyzed at all.

As Expert 5 observed in debate, each reviewer picked a somewhat different set of flaws, which itself underscores that the methods were deficient across many dimensions. The practical split is over use: all agree the study cannot support quantitative dose–response assessment, but a subset would retain it as qualitative hazard evidence for site-of-contact tumors.

AI generated summary content

ANSWER EXPLANATIONS

ANSWER EXPLANATIONS, CONTINUED

Expert 1 | **SciTrust Score:** 3.9

	Confidence Rating
Methods	3
Results	5 (medium confidence)
Discussion/Conclusions	4

Methods

Appropriate species (score of 5). Swiss mice can be used for carcinogenicity studies, no justification for selection , original source of animals unknown, no information on tumor incidences in historic controls.
No. of animals/group (score of 3). unequal number of males and females combined/dose group, number of males in some dose group low.
Controls included and exposed to tap water only (score of 6)
Study duration (score of 2). Exposure for a period of 78 weeks, but animals were examined for tumors onyl after natural death (issues with rapidly increasing incidences of tumors in aging rats, histopathology analysis of lysed tissues
Number and spacing of dose levels (score of 5), controls and two dose levels, no justification for dose levels, dose received not determined
Rote of exposure may be relevant to humans (score of 7) (migration of VAM from food contact materials into food ? how stable is VAM in food contact materials, food ?)
Test item preparation, source, verification (score of 6). Fresh preparations daily, little information on source of test item, no analytical verification of target conc and stability, presence of actaldehyde as decomposition product ?
Determination of dose received (score of 0). No information given
Necroscopy, histopatho (score of 4), Organs selected for histopath similar to guidance, apparently all animals including early deaths subjected to histopatho, some information on tissue processing given, issues with histopatho analyses of lysed tissues when dead animals remain in cages for a longer time. No info on blinding of evaluators
Tumor incidences, stats (score of 2). Study results may only be used for descriptive purposes. Little information for quantitative assessment. Dose-response for total malignant tumors not evident.
Guideline (score of 2), duration of exposure not according to guidance, no. animals/dose group not according to guidance, no terminal sacrifice at defined time point, no info on clin chem and urine analysis, limited information on preneoplastic lesions, no information on general toxicity/cytotox in oral cavity

Results

Descriptive overview of tumor incidences, no statistical evaluation of tumor incidences, local tumors at port of entry and upper GI tract evident, relevance of other reported increases in incidence difficult to evaluate due to issues with study design/reporting. Local tumors at site of contact consistent with results and kinetics of decomposition of VAM and tumor location in some other studies.
General conduct: no info on infectious disease outbreak or lab hygiene, question of GLP status. Local tumors as main outcome, unlikely to be compromised by shortcomings in design.

Discussion/Conclusions

very limited discussion of results (score 3),
Conclusions on tumor incidences in oral cavity/upper GI tract well supported (score 8) by observations

Impact of design on study outcome not discussed. Actually, the discussion is quite short. However, the main finding (oral and upper GI tract tumors due to VAM) is consistent with the local tumor effects seen with VAM in other oral studies (score 7) by this lab..

Study results difficult to use in quantitative risk assessment due to deficiencies in reporting and study design (score 2)

Discussion should have addressed differences in study outcome of the drinking water studies (no lesions in oral cavity in Bogdanffy vs. tumors in this study despite identical top concentrations (score 1)

Conclusions: Due to issues with study design and reporting, the study receives low quality rating and should not be used for quantitative assessments. In addition, systemic tumors reported should not be used for hazard assessment of VAM due to study design issues. However, study shows that VAM induces tumors at portal of entry after administration of high conc. in drinking water and thus is consistent with a number of other studies.

Expert 2

	Confidence Rating
Methods	
Results	
Discussion/Conclusions	

No explanation provided.

ANSWER EXPLANATIONS, CONTINUED

Expert 3 | **SciTrust Score:** 3.9

	Confidence Rating
Methods	3
Results	5 (medium confidence)
Discussion/Conclusions	4

Methods

An appropriate species was selected. The number of animals per dose group, however, was less than optimal. Although the study duration exposure was typical, animals were allowed to live out their lifespan, which allowed potential recovery from chemical exposure and possible complications in reviewing tumors against control animals that also lived their full lifespan. The two doses and control were not inappropriate, although a lower dose would have been preferred as well. The length of exposure, including partial *in utero* exposure, was relevant. The test substance was appropriately characterized for purity and stability. However, several contaminants were described, at least one of which is a known carcinogen (i.e., benzene at 3-5% of the low concentration). The documentation of actual versus administered dose was not provided. A complete necropsy was performed on animals including those that died early, and collection of tissues appeared to be processed and examined according to standard protocols. It was not immediately apparent that histopathology was performed blindly nor subjected to an independent pathology review. A statistical analysis was not apparent.

Results

There did not appear to be any disease outbreaks in the experimental and control animals. Survival was adequate across those groups to allow a meaningful comparison of tumors. There was no apparent excessive toxicity at high doses or reduced water our feed intake that would affect results seen. The data were appropriately summarized, and tumor incidences were appropriately reported by group, sex and site. The findings appeared to be reported completely.

Discussion/Conclusions

The authors conclusion regarding of carcinogenicity is generally limited to the high concentration of 5000 ppm (0.5% of the water concentration). In part because of this, the authors overstate the evidence that this chemical causes cancer in the lower and more expected concentrations to which humans are exposed. This overstatement is also due, in part, to the numerous limitations of the study and uncertainties in the data or interpretations that the authors did not acknowledge. For example, the study did not meet several GLP guidelines (number of animals, housing conditions), the monitoring of chemical concentrations was not described nor discussed in relation to tumor findings, and non-oncologic effects were said to be absent despite the presence of tumors at high concentration. This later finding is unusual.

ANSWER EXPLANATIONS, CONTINUED

Expert 4 | **SciTrust Score:** 0

	Confidence Rating
Methods	0 (fundamentally flawed)
Results	0 (fundamentally flawed)
Discussion/Conclusions	0 (fundamentally flawed)

Methods

See above
- no statistics
- no scheduled sampling
- no analysis of compound contaminants in the administered drinking water.
- No analysis of the concentration of VAM in the drinking water

Results

The results are dependent on the experimental design, and the design is flawed therefore all fruit from this poisoned tree are flawed.

Discussion/Conclusions

See above

ANSWER EXPLANATIONS, CONTINUED

Expert 5 | **SciTrust Score:** 0

	Confidence Rating
Methods	2
Results	2
Discussion/Conclusions	0 (fundamentally flawed)

Methods

1) Score =0.2. Swiss mice were used, common in many long term bioassays. No reason for this choice was offered.

2) Score=0. The number of animals was insufficient, both for the breeders and the offspring (Table 1).

3) Score=0.3. The control group was adequate because it was approximately the same size as the treated groups, although both were too small.

4) Score=0.5. The animals were maintained until ~150 weeks, well over the lifespan of 104 weeks.

5) Score=0. The exposures were 0 ppm (control), 1000ppm and 5000ppm. These are very high concentrations. No rationale was given for the choice of doses.

6) Score=0. P. 211 cites an earlier study and notes that 2500 ppm is the maximum tolerated concentration in rats. The use of 5000 ppm seems odd.

7) Score=0.1. P. 210 indicates polyvinyl acetate is used in chewing gum. VAM is used to produce polyvinyl acetate., so presumably human oral exposure could occur. Although polymers made with VAM are also used in packaging, no data on actual exposure to VAM is offered.

8) Score=0.5. The VAM was testing and the impurities are listed on p. 212. The drinking water solutions were changed daily.

9) Score=0. The paper contained no information suggesting the drinking water concentrations were ever tested.

10) Score=0. No mention was made of measurement of water consumption by the mice, so no dose information

11) Score=0. Measurements were not conducted to measure water consumption.

12) Score=0.25. P. 214 describes the pathology procedures, which seem complete, but not guidelines were mentioned.

13) Score=0. No authoritative guidelines on pathology were mentioned.

14) Score=0. Whether the pathologists knew the dose groups of the tissues examined was not discussed in the paper.

15) Score=0. No statistics was performed to analyze either differences or dose response.

16) Score=0. No survival adjustments were made as no dose-response or trend determination was conducted.

17) Score=0. Historical controls were not considered in this work.

18) Score =0.2. Tables 4 and 5 show dose response data on total benign and malignant tumors respectively. In the tissue -specific tables, the tumor types included carcinomas, that I considered to be malignant. No explanation was given in the paper.

19) Score=0. No need for multiple comp. adjustment, as no statistics were performed.

20) Score=0. The study does not match established guidelines.

21) Score=0. The deviations from any guidelines was not discussed in the paper.

Results

1) Score=0. Infectious disease wasn't mentioned in the paper, but Figs. 5-8 showing survival show no sudden drops, as might occur with an infection.

2) Score=0.1. 40-60% of the animals were still alive at 104 weeks. This may have affected the conclusions.

3) Score=0.1. Excessive toxicity may have skewed the results. The authors note that 2500 ppm was poorly tolerated by rats. Whether this concentration affects mice similarly isn't known. The rational for 5000ppm as the maximum dose is not clear.

4) Score=0. Since water intake was not measured, this question cannot be answered.

5) Score=0.2. The data are summarized in the table in a confusing and obfuscatory manner. These data are nonetheless available in these tables.

6) Score=0.3. The data are presented by sex and dose group for both breeders and offspring. I copied the numbers from Tables 5-16 to a spreadsheet to examine the possibility of assessing dose-response. What I observed is that the incidence of specific tumors was generally low save for lymphoma/leukemias but these cancers did not show a clear dose-response.

7) Score=0.4. The tables provide the findings in a fairly complete way save having the individual animal data.

8) Score=0.5. Figs 1-4 show the body weight change. The weight difference between dose groups is small. Slight decrease in weight occurred in the offspring.

9) Score=0.3. Although intercurrent disease was not mentioned, an occurrence seems unlikely given the lifespans shown in figs. 5-8

10) Score = 0.3. Some of the tables show both squamous cell dysplasias and squamous cell carcinomas. Presumably, these are mechanistically related although the paper is very short of discussion of biology or mechanism.

Discussion/Conclusions

1) Score= 0.1. Throughout the paper, the word "cause" is used inappropriately, for example, in the summary on p. 237. Without any statistical examination, the occurrence of tumors at any dose cannot be considered dispositive. These causal statement go well beyond the data.

2) Score=0.1. The food contact material and chewing gum that contain VAM-based polymers may indeed leach VAM. However, without any measurement of the human doses involved, the discussion on p. 237 cannot be justified.

3) Score=0. The study did not have sufficient power due to the small number of animals.

4) Score=0. Consideration of human exposures that would be much, much lower than those experienced by Maltoni's mice does not warrant a call for concern. This study is not a risk assessment but a report of an animal bioassay.

5) Score=0. In the summary, the statement is made that Vam causes an increase in tumors of the Zymbal gland, mouth, tongue, esophagus, forestomach, stomach, lung and uterus. This statement was made without any dose response analysis or statistical comparisons.

6) Score=0. The uncertainties in the data are indeed minimized and seem to be ignored.

7) Score=0. No, the language goes way beyond the results.

Debate | 2 comments

Expert 5 | 10/11/2026 8:53 SCORE: **0**

The various experts each chose a somewhat distinct set of nits to pick, suggesting as noted in the summary that the methods were flawed in many areas.

Expert 3 | 10/14/2026 9:10 SCORE: **0**

Experts 1 and 3 are in complete agreement (how often does that happen?) but for very slightly different reasons.

SciPinion

SciPi #1052

October 19, 2026

Round 1: Minardi et al. 2002

SciPinion L.L.C.
777 E. Main St., Suite 202
Bozeman, MT 59715

shays@scipinion.com
ckirman@scipinion.com
(406) 404-6145

What are the strengths of Minardi et al. (2002) (if any)?

Results | 5 answers

Expert 1

Strength and weaknesses as in Bepoggi et al. and in Maltoni et al.

4 comments

Expert 2 | 10/10/2026 10:26 SCORE: **0**

I agree. And I think all of the experts are in agreement that there are more weaknesses than strengths.

Expert 5 | 10/11/2026 8:55 SCORE: **0**

Agree with #1 and #2 -- Ramazzini papers show more weaknesses than strengths.

Expert 4 | 10/11/2026 12:23 SCORE: **0**

Agree

Expert 1 | 10/12/2026 10:11 SCORE: **0**

also agree

Expert 2

[Note: Although several strengths are listed, it is important to consider the many weaknesses/limitations (Q6.2) which impact the confidence in the study's reported findings.]
Strengths:
- Two dose levels (5000 ppm, 1000 ppm) of vinyl acetate (VA) including concurrent controls were evaluated,
- Sprague-Dawley rats, both sexes, were exposed 104 weeks from in utero to adulthood to evaluate lifespan effects,
- Multiple tissues examined, and tables on tumors are provided,
- Authors indicated that test material purity >99%, and reported the impurities including the carcinogen benzene was 30–45 ppm.

Expert 3

Full lifetime monitoring, although this is not seen to be helpful by many toxicologist due to problems associated with comparison to aging control animals.

Expert 4

> This is another study from the Ramazzini group. It demonstrated that squamous cell oral tumors were induced following chronic exposure to VAM. Although questions with regard to the methods are raised, this finding is useful in assess the tumorigenic potential of VAM.

Expert 5

> The study did use sufficient animals in the offspring group and the duration of 104 weeks was adequate.

Experts identified a limited set of strengths in Minardi et al. (2002), while broadly agreeing that these are outweighed by the study's methodological limitations.

• Design features cited as strengths:Several experts pointed to the use of two dose levels (5000 and 1000 ppm) with concurrent controls, both sexes of Sprague-Dawley rats, adequate group sizes in the offspring, exposure from in utero through adulthood, and examination of multiple tissues with tumor data tabulated. Reported test material purity (>99%, with benzene impurity quantified at 30–45 ppm) was also noted favorably.
• Study duration:Agreement that the 104-week/full-lifetime exposure and monitoring period was adequate, though one expert qualified this by noting many toxicologists view lifetime monitoring as problematic because of difficulties comparing results against aging control animals.
• Relevance of findings:One expert considered the induction of squamous cell oral tumors following chronic VAM exposure a useful contribution to assessing tumorigenic potential, despite reservations about the methods.
• Consensus on overall weight:A clear point of agreement — explicitly endorsed by multiple experts — is that this study, like other Ramazzini Institute publications (e.g., Maltoni et al., Bepoggi et al.), presents more weaknesses than strengths, and that confidence in the reported findings is correspondingly reduced.

No substantive disagreement emerged among respondents; differences were mainly in emphasis, with some listing specific design strengths and others framing even those strengths as heavily qualified.

AI generated summary content

What are the weaknesses of Minardi et al. (2002) (if any)?

Results | 5 answers

Expert 1

see comments on Belpoggi et al. and Maltoni et al.

Expert 2

There are a number of concerns and uncertainties, and a lack of detail to have confidence in the findings.

• **Lack of Test Material Characterization;**There is a lack of information on the verification and stability of VA in drinking water (DW) solutions, and assessment of actual DW intake. Thus, actual administered dose of VA is not known. Decomposition of VA to acetaldehyde (AA) and acetic acid can occur and could contribute to the cytotoxicity/inflammation and site-of-contact responses,
• **Insufficient sample size:** F0 male breeders (13-14) which is too small to be reliable. F0 females also low (37).
• **Unknown Randomization across Dose Groups;** There is no information on how breeders and offspring were allocated across treatment groups. Also, the litter of origin for the tumor responses are unknown. There is the potential for littermates to respond similarly (cluster),
• **Lack of Transparency and Detailed Pathology and Confirmation;**It is not known whether the diagnosis was blinded. No detailed description of pathology provided in paper to support diagnosis of malignant tumors,
• **Lack of Historical Data; L**aboratory historical data for this study design are not available to understand background incidence and its variability. In general, carcinomas of the oral cavity, tongue, esophagus, and forestomach in SD rats are uncommon. The number or percent tumors in concurrent controls in Minardi et al was generally zero or low. However, without the Laboratory historical data it is unknown whether the low concurrent-control incidence reflects a true "very" low spontaneous background rate versus an unusually low incidence in controls. Historical data would provide context,
• **Lack of Body Weight Data;**Although the authors reported that there were no "substantial" differences between treated and control in mean body weight (BW), data were not provided to support their conclusion. It is unclear what "substantial" implies,
• **Lack of Food (and Water) Consumption Dat**a,
• **Lack of Survival Data;**No data were provided on survival. The authors only state there is no "substantial" differences in survival. It is unclear what "substantial" implies. A small difference in survival may not necessarily distort the data. Nonetheless, the authors should have presented survival data and consider statistical methods for survival-adjusted tumor incidence as good practice.
• **Uncertain Health Status;**No information or statements was provided on the health of source animals. [Note: Chronic infection (e.g., bronchopneumonia) in the Ramazzini Institute occurred in some colonies during this time period]. Chronic infection/inflammation in the source animals could influence the results.

2 comments

Expert 5 | 10/11/2026 8:57 SCORE: **0**

Agree with all these issues.

Expert 4 | 10/11/2026 12:24 SCORE: **0**

Good summary of the many weaknesses. I think we have general agreement on our evaluations

Expert 1 | 10/12/2026 2:53 SCORE: **0**

agree

Expert 3

This bioassay has a number of significant flaws that preclude making definitive conclusions.
- Only partial utero exposure.
- Contaminants monitored but not discussed.
- No vaccination of experimental animals.
- No interim sacrifices.
- Not GLP compliant (e.g., insufficient number of animals, lack of housing information).

Expert 4

As with the other Ramazzini studies, VAM, had impurities of benzene 30–45 ppm; methyl and ethyl acetate 50 ppm; crotonaldehyde 6 -16 ppm; acetaldehyde 2–11 ppm; acetone 330–500 ppm. The approach of allowing the study to continue until the rats died is problematic and makes it very difficult in identifying and quantitating tumors in target organs.

1 comment

Expert 5 | 10/12/2026 10:02 SCORE: **0**

Agree

Expert 5

This is one of the most poorly presented papers I've encountered in my career. The presentation was incomplete and lazy.

2 comments

Expert 2 | 10/10/2026 10:28 SCORE: **0**

I am in full agreement that this is a poorly presented paper.

2 Replies

Expert 5 | 10/11/2026 9:04 SCORE: **0**

The thing that irked me most was the lack of any statistical detail save the reference to the Soffritti paper.As first author, Minardi presumably drafted the paper and seemed more like a scribe or assistant who wanted to complete the draft as quickly as possible. Today, Minardi would have used AI and doing so might have improved the presentation.

Expert 2 | 10/11/2026 11:45 SCORE: **0**

The lack of statistical detail/transparency is a valid point, which I think was raised by others.

Expert 4 | 10/11/2026 12:25 SCORE: **0**

agree

Experts were unanimous that Minardi et al. (2002) suffers from substantial methodological and reporting deficiencies severe enough to preclude confident or definitive conclusions. No expert defended the study.

Points of broad agreement:

• Poor reporting and transparency— characterized by one expert as "one of the most poorly presented papers I've encountered in my career," a judgment explicitly endorsed by others. The absence of statistical detail (deferred to a prior Soffritti paper) was singled out as particularly troubling.
• Missing core data— body weight, food and water consumption, and survival data were not presented, with vague claims of no "substantial" differences left undefined; no survival-adjusted tumor analysis was performed.
• Test material concerns— inadequate characterization and verification of vinyl acetate in drinking water, unknown actual dose delivered, potential decomposition to acetaldehyde/acetic acid, and impurities (benzene, crotonaldehyde, acetaldehyde, acetone) that were monitored but never discussed.
• Study design deficiencies— insufficient group sizes (notably F0 breeders), unclear randomization and litter allocation, no interim sacrifices, lack of GLP compliance, and incomplete housing information.
• Pathology limitations— unknown blinding, no detailed pathology descriptions supporting malignancy diagnoses, and no laboratory historical control data to contextualize unusually low concurrent-control tumor incidences.

Individually raised points: the Ramazzini practice of maintaining animals until spontaneous death, which complicates tumor identification and quantitation; uncertain animal health status given documented chronic infections in Ramazzini Institute colonies during this period; only partial in utero exposure; and lack of vaccination of experimental animals.

Disagreement: essentially none. One expert simply cross-referenced critiques of the related Belpoggi and Maltoni studies, and participants explicitly noted "general agreement on our evaluations." Differences were in emphasis only — some focused on data gaps and statistical rigor, others on exposure design and regulatory compliance.

AI generated summary content

Please rate your degree of confidence in each of the following study components for Minardi et al. (2002) (0-10 scale; 0 = fundamentally flawed; 1=lowest confidence; 5=medium confidence; 10=highest confidence). An overall confidence score will be calculated as the geometric mean of your three ratings. Please provide your rationale for each study component in the box beside it.

Results | 5 answers

	Methods	Results	Discussion/ Conclusions	SciTrust Score
Expert 2	2	3	3	2.6
Expert 4	3	6	3	3.8
Expert 1	4	4	5	4.3
Expert 5	2	5	2	2.7
Expert 3	3	5	5	4.2
Avg	2.8	4.6	3.6	3.5
Range	2–4	3–6	2–5	2.6–4.3

0 = no confidence · 5 = medium confidence · 10 = highest confidence

SciTrust Score = geometric mean of a reviewer's category scores. A zero in any category zeroes out that reviewer's full score. The **Final SciTrust Score** is the mean of all reviewer SciTrust Scores.

Experts converged on low-to-moderate confidence in Minardi et al. (2002), with methods ratings clustering tightly at the low end (2–4) and somewhat more variable ratings for results (3–6) and discussion/conclusions (2–5).

Points of agreement:

• The methods section is unacceptably sparse, improperly deferring key design details to Soffritti et al. (2002), and lacks information on animal numbers per group, randomization, blinding, and pathology review.
• There was no verification of vinyl acetate concentrations in drinking water, no water consumption data, and therefore no knowledge of actual administered dose — compounded by concerns over test substance stability and impurities (acetone, benzene).
• Only two dose levels plus control were used, with no MTD rationale and no lower dose to inform the low-exposure range; historical controls and multiple-comparison adjustments were absent.
• There is no true discussion section, study limitations and uncertainties are never acknowledged, and the "multipotent carcinogen" claim is undefined and overstated given that tumors were essentially squamous lesions of the upper GI tract.
• Despite these flaws, the portal-of-entry tumors (oral cavity, esophagus, forestomach) are biologically plausible, consistent with other studies, and should not be dismissed outright.

Points of divergence and nuance:

• Expert 4 and Expert 5 gave relatively higher results ratings (6 and 5), crediting the detailed tabulation and histologic classification of squamous lesions as useful for mode-of-action evaluation; Expert 2 rated results lowest (3), questioning whether inflammatory/dysplastic lesions may have been misclassified as carcinomas and criticizing the "composite analysis" of dysplasias and carcinomas as inflating apparent evidence.
• Expert 3 viewed test substance characterization and necropsy/tissue processing more favorably than Experts 2 and 5, who judged dose verification and pathology documentation essentially absent.
• Experts differed on the lifespan design: Expert 1 saw the long observation period as a possible explanation for tumors not seen by Bogdanffy et al., whereas Experts 3 and 4 regarded death-as-endpoint sampling as a limitation on interpretation.
• Expert 2 uniquely emphasized litter-independence concerns for F1 findings and argued systemic (pancreatic, adrenal) tumor signals were unconvincing.
• Experts 3 and 5 sharply criticized the conclusions' extrapolation to human exposures and call for regulatory action as going beyond the data.

In the debate, all participating experts explicitly endorsed one another's critiques, characterizing the study as straddling low and medium confidence: real and potentially consequential limitations, but a portal-of-entry finding that plausibly holds. Expert 1 concluded the study should not be used for quantitative assessment while remaining qualitatively consistent with other evidence.

AI generated summary content

ANSWER EXPLANATIONS

Expert 1 | **SciTrust Score:** 4.3

	Confidence Rating
Methods	4
Results	4
Discussion/Conclusions	5 (medium confidence)

Methods
general comments on design and context as in Belpoggi et al., and Maltoni et al.,

Results
tumors in oral cavity and upper GI tract, as seen in the other studies from this lab

Discussion/Conclusions
see comments on Belpoggi et al and Maltoni et al,
long duration of study somewhat justified as tumors in oral cavity appeared late in study, not addressed if the longer observation period was the driver for the observation in this study as compared to Bogdanffy et al (no tumors in the oral study)

Conclusions: Due to issues with study design and reporting, the study receives low quality rating and should not be used for quantitative assessments. However, study shows that VAM induces tumors at portal of entry after administration of high conc. in drinking water and thus is consistent with a number of other studies.

ANSWER EXPLANATIONS, CONTINUED

Expert 2 | **SciTrust Score:** 2.6

	Confidence Rating
Methods	2
Results	3
Discussion/Conclusions	3

Methods

The paper lacks methodological detail. See responses to Q 6.2 regarding weaknesses.

With respect to the lack of test material characterization, the authors refer to another paper (Soffritti et al 2002) from the Ramazzini Institute on methyl and ethyl alcohol, which only indicates that drinking water formulations were replaced each morning. Because water consumption data are not reported, the actual administered dose of VA is not known. There is no information in Minardi et al (nor Soffritti et al) whether VA was measured (as well as AA and acetic acid) in the drinking water solutions at the end of 24 hours. If the drinking water solutions were replaced daily, however, the decomposition to AA and acetic acid is less but not eliminated. Based on the Bogdanffy et al paper, regarding stability of VA in drinking water, up to 14% of VA was lost in the first 24 hr and 40 to 60% over a 7-day period.

In neither Soffritti et or Minardi et al, is there information in how animals (breeders and offspring) were assigned across dose groups (i.e., was randomization done?) to have confidence that the tumor responses are independent experimental units and are independent of litter.

There were no statements in Minardi et al or Soffritti et al about whether widely international standardized histopathology criteria were used or whether the diagnosis was blinded. Soffritti et al only indicate that slides were examined using same criteria/classification by the same pathologists with a senior pathologist reviewing all tumors and any other cancer lesions of interest.

Results

Given that VA, as well as AA and acetic acid, wasn't measured in the drinking water solutions, it is not known to what extent (if at all) potential decomposition products contributed to the site of contact effects.

In this study, carcinomas of the oral cavity and forestomach were statistically increased at the highest dose tested in F1's. These upper GI tract tumors cannot be entirely dismissed and are biologically plausible. VA is a site-of-contact toxin given its rapid hydrolysis and accumulation of acetic acid resulting intracellular acidification. The presence of carboxylesterases in the oral cavity (Robinson et al 2002) as well as a longer resident time in the forestomach would be consistent with these responses. These port-of-entry tumors likely reflect chronic irritation/inflammation.

The question could be raised whether some non-neoplastic inflammatory/proliferative responses were misclassified as carcinomas (e.g, squamous dysplasia vs carcinoma). There is no detailed pathology descriptive text or microphotographs. Thus, it is not clear whether invasion of the basement membrane occurred for the carcinoma diagnosis (do all the carcinoma findings reflect a clear carcinogenic signal?). There is no grading of the dysplasia (mild, moderate, severe?) provided. [Note; RI drinking water studies on other chemicals have reported increased incidences of upper GI tract tumors, and been criticized for the pathology (e.g., distinction between irritation lesions and true malignancy). Certain studies were considered suspect given certain chemicals were not strong irritants.]

It is unknown how breeders and offspring were assigned across dose groups. The litter of origin for the tumors is not known,

and thus, it is unclear if the increase in tumors are independent of the litter or driven by a small number of litters.

The authors presented a "composite analysis" of the dysplasia and squamous carcinomas of the upper GI tract (oral cavity, esophagus, forestomach) as supportive information. But such a composite analysis can make the strength of the evidence look greater. Although there is a common broad epithelium lineage, and dysplasia can be a precursor to carcinoma, it is not in itself cancer.

There was no convincing evidence of a systemic tumor response, and in the Results section Minardi only discusses the upper GI tract tumors. Although OEHHA points out the benign pancreatic tumors (islet cell adenomas in male breeders, sample size was 13), pancreatic exocrine adenomas (F1 males) and adrenal pheochromoblastoma (F1 females), these responses were not dose related and could be within normal variation (there are no historical data for context). Also, there is the uncertainty of whether the F1 responses are litter independent.

Discussion/Conclusions
There is no discussion section, only a brief conclusion that VA is a multi-site carcinogen primarily based on the different upper GI tract tumors. The information/evidence presented is not convincing of systemic carcinogenicity. Although the port-of-entry effects in the upper GI tract cannot be dismiss, which are biologically plausible, caution should be taken given the uncertainties about whether the carcinogenic responses are over interpreted or over represented.

ANSWER EXPLANATIONS, CONTINUED

Expert 3 | **SciTrust Score:** 4.2

	Confidence Rating
Methods	3
Results	5 (medium confidence)
Discussion/Conclusions	5 (medium confidence)

Methods

An appropriate species was selected. The number of animals per dose group, however, was less than optimal. Although the study duration exposure was typical, animals were allowed to live out their lifespan, which allowed potential recovery from chemical exposure and possible complications in reviewing tumors against control animals that also lived their full lifespan. The two doses and control were not inappropriate, although a lower dose would have been preferred as well. The length of exposure, including partial *in utero* exposure, was relevant. The test substance was appropriately characterized for purity and stability. However, several contaminants were described, at least one of which is a known carcinogen (i.e., benzene at 3-5% of the low concentration). The documentation of actual versus administered dose was not provided. A complete necropsy was performed on animals including those that died early, and collection of tissues appeared to be processed and examined according to standard protocols. It was not immediately apparent that histopathology was performed blindly nor subjected to an independent pathology review. A statistical analysis was apparent.

Results

There did not appear to be any disease outbreaks in the experimental and control animals. Survival was adequate across those groups to allow a meaningful comparison of tumors. There was no apparent excessive toxicity at high doses or reduced water our feed intake that would affect results seen. The data were appropriately summarized, and tumor incidences were appropriately reported by group, sex and site. The findings appeared to be reported completely.

Discussion/Conclusions

The authors conclusion regarding of carcinogenicity is generally limited to the high concentration of 5000 ppm (0.5% of the water concentration). In part because of this, the authors overstate the evidence that this chemical causes cancer in lower concentrations to which humans are more likely exposed. This overstatement is also due, in part, to the numerous limitations of the study and uncertainties in the data or interpretations that the authors did not acknowledge. For example, the study did not meet several GLP guidelines (number of animals, housing conditions), the monitoring of chemical concentrations was not described nor discussed in relation to tumor findings, and non-oncologic effects were said to be absent despite the presence of tumors at high concentration. This later finding is unusual.

ANSWER EXPLANATIONS, CONTINUED

Expert 4 | **SciTrust Score:** 3.8

	Confidence Rating
Methods	3
Results	6
Discussion/Conclusions	3

Methods

The described methods are very sparse and refer to another Ramazzini group paper for details on the bioassay (Soffritti et al. 2002. Results of long-term experimental studies on the carcinogenicity of methyl alcohol and ethyl alcohol in rats. Ann. N.Y. Acad. Sci. 982: 46–69). This is inappropriate. The general methods should be included in the manuscript. The reliance on the determination of tumor induction using spontaneous death of the rats rather than a defined sampling time limits the interpretation of the results

Results

Statistics were apparently performed (chi Square analysis) but the presentation of the results of the statistical analysis are simply indicated in the tables by an asterisk (no discussion in the results section) without any attribution to what the values are being comparing. The results text should have clarified the results of the statistical analysis. On a positive side, detailed classification of the squamous cell tumors and location of the lesions was described in detail and provides useful data to examine mode of action evaluation.

Discussion/Conclusions

There is no discussion section only a section labelled conclusions. The conclusions reference the other results from the Ramazzini group but fail to discuss the results presented in this manuscript with the previous publications Lijinsky and Reuber (1983) and Bogdanffy et al 1994 a, 1994b). In addition, there is a statement in the conclusions that the results show that "VAM is a multipotent carcinogen". This term is not defined in the manuscript. Usually, the term multipotent infers multiple sites of tumor induction such as seen with many genotoxic agents such as dimethylnitrosamine. With VAM the tumors were in the oral tract with the same histologic presentation (squamous cell).

ANSWER EXPLANATIONS, CONTINUED

Expert 5 | **SciTrust Score:** 2.7

	Confidence Rating
Methods	2
Results	5 (medium confidence)
Discussion/Conclusions	2

Methods

1) Score=0. The methods section indicated that male and female SD rats were used but dose not mention the numbers in each dose group (0, 1000, 5000ppm in water).
2) Score=0.15. The methods do not mention the number of animal. These were not reported until Table 2 on the 10th page of the paper. The number of animals in the offspring was sufficient but not in the breeders.
3) Score=0.3. The only controls were the negative ones, 0ppm in breeders and offspring.
4) Score=0.45. The study duration was adequate at 104 weeks.
5) Score=0.3. Only two exposure levels were used along with a control. A lower level, say 200ppm would have been helpful.
6) Score=0. The MTD was not considered and no rationale for the chosen doses was given.
7) Score=0.4. The narrative in the introduction of leaching from food packaging and chewing gum suggest oral exposure may occur in humans, So, yes, the exposure route was potentially relevant.
8) Score=0.15. The methods section lists the impurities in the VA obtained from an unidentified source with acetone up to 500 ppm.
9) Score=0. No discussion of ongoing validation of VA concentrations was offered.
10) Score=0. No discussion of the actual doses the rats received was provided.
11) Score=0. Measurement of water consumption was not offered.
12) Score=0. The methods reference Soffritti et al. (2002) for the tissues necropsied and other details.
13) Score=0. The details of tissue processing are also in the Soffritti paper.
14) Score=0. No mention is made here or in Soffritti about pathologist blinding to dose groups
15) Score = 0.1. Soffritti et al. indicates that Chi-square tests were used to evaluate differences and the Cochran-Armitage test was used for trend.
16) Score=0. No mention was made of survival adjustment either here or in Soffritti.
17) Score=0. Historical controls were not considered.
18) Score=0.4. Tables 3, 4, 5, and 6 show the tumors of the oral cavity, tongue, esophagus, and forestomach. These include both squamous cell dysplasias and squamous cell carcinomas.
19) Score=0. No accounting for multiple comparisons occurred
20) Score=0. No guidelines were mentioned and the number of breeders was clearly insufficientl.
21) Score=0. Since no guidelines were mentioned, how could deviations be discussed.

Results

1) Score=0.1. Infection disease was not mentioned. The results indicate no difference between dose groups in “survival behavior.”
2) Score=0.2. The lack of differences in “survival behavior” suggests that survival was adequate, but this topic was not explored.
3) Score=0.8. The high dose of 5000ppm was likely not sufficient to produce toxicity that would affected the results
4) Score=0. Food and water intake was mentioned only by reference to other work.
5) Score=0.8. The data were adequately summarized in the tables.
6) Score=0.8. The tables provide data for tumor incidence sufficient for dose-response and the table footnotes mention the Cochran-Armitage test.

7) Score=0.9. The tumor findings are summarized in the tables
8) Score=0. Changes in body weight were mentioned once, in the first sentence of the Results sections. No details were given.
9) Score=0.2. Presumably, no disease that affected survival of animals or the results occurred based on the first sentence in Results.
10) Score=0.9. Tables 3-7 show squamous cell carcinomas and squamous cell dysplasias of the upper GI tract.

Discussion/Conclusions
1) Score=0.05. The first paragraph of the conclusions indicates the increase in tumors in the upper GI tract similar to other rodent studies. The third paragraph that calls for regulatory action goes way beyond the science presented. This is a poorly presented animal bioassay, not a full human health risk assessment.
2) Score=1. The first two paragraphs in Conclusions are justified based on the observed results
3) Score=1.2 The Conclusions mentions the study duration in the second paragraph.
4) Score=0.05. The third paragraph is inappropriate.
5) Score=0.05. The limitations of the study are never discussed. Whichever author wrote this paper seemed either lazy or very poorly trained.
6) Score=0. The uncertainties is the data and not discussed.
7) Score=0. Certainly not in the call for regulation

Debate | 4 comments

Expert 2 | 10/10/2026 10:11 SCORE: **0**

I agree with the comments of all experts. And I can understand how we straddled between Low confidence (Serious problems that could plausibly change the study's conclusion) and Medium confidence (Real limitations; the conclusion probably holds, but with notable uncertainty.) I further understand why the oral cavity tumors are not entirely dismissed. There is biological plausibility and VA is a site of contact toxicant. But there is the real possibility that incidences are overstated in significance. The expert comments clearly point out the flaws/limitations. I raised additional concerns regarding the uncertainty around the diagnosis for malignancies, and the lack of information on randomization of animals/pups across dose groups and the uncertainty of whether results are litter independent.

Expert 5 | 10/11/2026 9:09 SCORE: **0**

Similar to #2, I also agree with the comments of all experts. My sense of the paper is that the message was the association of point-of-contact tumors with VA. The authors seemed to believe this association sufficient to apply these results to humans without any attempt at a risk assessment.

Expert 1 | 10/12/2026 2:56 SCORE: **0**

agree with #2 and #5. In summary, all flaws in the study are clearly identified by the group

Expert 3 | 10/14/2026 9:07 SCORE: **0**

I am in agreement with my colleagues and for the various reasons stated.

SciPinion

SciPi #1052

October 19, 2026

Round 1: Umeda et al. 2004

SciPinion L.L.C.
777 E. Main St., Suite 202
Bozeman, MT 59715

shays@scipinion.com
ckirman@scipinion.com
(406) 404-6145

What are the strengths of Umeda et al. (2004) (if any)?

Results | 5 answers

Expert 1

> Study performed under GLP following testing guideline using drinking water application, detailed reporting of incidences of preneoplastic and neoplastic lesions in oral cavity, upper GI-tract. Some increases in incidence of systemic tumors, but not dose-dependent nor statistically significant

Expert 3

> Followed GLP protocols.

Expert 4

> This is a well-designed chronic study examining the effects of oral (drinking water) administration of VAM in mice (BDF1) and F344 rats. The compound was analyzed to be 98 % pure (although no mention of the other 2% was made). The number of animals per group studied was 50 (appropriate). Statistical analysis of the results was performed. Proper histopathological analysis was performed and reported. The study was performed with a schedule sampling time of two years of treatment.

1 comment

Expert 1 | 10/12/2026 10:13 SCORE: **0**

agreement within the expert group regarding study quality

Expert 5

> The experiments seem well conducted, the statistical aspects were sound, and the comparison to other studies was helpful.

Experts were unanimous in regarding Umeda et al. (2004) as a methodologically sound and well-conducted study, with no dissenting views expressed.

Points of agreement:

• Strong study design and execution, including compliance with GLP protocols and established testing guidelines (Experts 1, 3, 4)
• Appropriate chronic (two-year) exposure via drinking water in both mice (BDF1) and rats (F344) (Experts 1, 4)
• Sound statistical analysis of results (Experts 4, 5)
• Adequate group sizes (50 animals per group) and verified test substance purity (98%) (Expert 4)
• Thorough histopathological evaluation and detailed reporting of preneoplastic and neoplastic lesions, particularly in the oral cavity and upper GI tract (Experts 1, 4)

Additional or differing emphases:

• Expert 5 highlighted the value of the authors' comparison with other studies, a strength not mentioned by others
• Expert 1 noted that increases in systemic tumour incidence were neither dose-dependent nor statistically significant — an observation framed as part of the study's detailed reporting rather than a weakness
• Expert 4 raised a minor caveat that the remaining 2% of the test substance was not characterised
• Expert 3 offered only a brief endorsement (GLP compliance) without further elaboration

AI generated summary content

What are the weaknesses of Umeda et al. (2004) (if any)?

Results | 5 answers

Expert 1

> blinding of patholgists not stated

Expert 3

> - No in utero exposure.
> - No contaminants monitored.
> - No vaccination of experimental animals.
> - No interim sacrifices.
> - High dose appears to be toxic in mice based on BW decrease.

Expert 4

> Several issues
> 1. The chemical analysis of the material was not completely divulged in the manuscript. It was indicated to be 98% pure, but no comment is made regarding the chemicals in the additional 2%. It is unclear what the basis for the selection of the doses used. In particular the 10,000-ppm dose.

1 comment

Expert 1 | 10/13/2026 4:17 SCORE: **0**

Agree that there are limitations

Expert 5

> No justification for the doses used was provided.

Experts broadly agreed that the Umeda et al. (2004) study has methodological limitations, though they emphasized different aspects.

• Dose selection:The most commonly shared concern — two experts (4 and 5) explicitly noted the absence of any justification for the doses used, with Expert 4 singling out the 10,000-ppm dose. Expert 3 added that the high dose appeared toxic in mice based on decreased body weight.
• Test material characterization:Expert 4 criticized the incomplete chemical analysis, noting the material was described as 98% pure with no identification of the remaining 2%. Expert 3 similarly flagged the lack of contaminant monitoring.
• Study design and conduct:Expert 1 raised the absence of any statement about pathologist blinding, while Expert 3 listed several design gaps: no in utero exposure, no vaccination of experimental animals, and no interim sacrifices.

No expert argued that the study was free of weaknesses, and no disagreements were expressed — Expert 1 explicitly endorsed Expert 4's critique ("Agree that there are limitations"). The differences among responses appear to reflect varying focus rather than conflicting judgments, and no debate comments were posted to contest any point.

AI generated summary content

Please rate your degree of confidence in each of the following study components for Umeda et al. (2004) (0-10 scale; 0 = fundamentally flawed; 1=lowest confidence; 5=medium confidence; 10=highest confidence). An overall confidence score will be calculated as the geometric mean of your three ratings. Please provide your rationale for each study component in the box beside it.

Results | 5 answers

	Methods	Results	Discussion/ Conclusions	SciTrust Score
Expert 4	8	8	8	8
Expert 1	9	9	9	9
Expert 5	7	9	8	8
Expert 3	7	8	5	6.5
Avg	7.8	8.5	7.5	7.9
Range	7–9	8–9	5–9	6.5–9

0 = no confidence · 5 = medium confidence · 10 = highest confidence

SciTrust Score = geometric mean of a reviewer's category scores. A zero in any category zeroes out that reviewer's full score. The **Final SciTrust Score** is the mean of all reviewer SciTrust Scores.

Overall confidence was moderate-to-high, with methods ratings ranging from 7 to 9, results from 8 to 9, and discussion/conclusions showing the widest spread (5 to 9).

Points of agreement:

• The study design was broadly judged sound and guideline-compliant (OECD-consistent species, 50/sex/group, 104-week duration, appropriate statistics, GLP conditions).
• Results reporting was consistently praised as clear, complete, and well tabulated, including preneoplastic lesions, site-specific tumor incidences, and historical control comparisons.
• Several experts flagged incomplete characterization of the test substance (purity/stability reporting) and the absence of information on blinded or independent pathology review.
• Near-unanimous agreement that tumor responses were essentially a high-dose (10,000 ppm) phenomenon, likely at or above the MTD, with no clear justification given for selecting that concentration.

Points of disagreement:

• Discussion/conclusions qualitywas the main fault line. Expert 3 rated it 5, arguing the authors failed to acknowledge the apparent threshold and the limited relevance of high-dose findings to human exposures, and questioned ratings of 8-9. Expert 1 defended a high score, noting OECD 453 does not require in utero exposure or interim sacrifices and that the lab and design were strong; while conceding non-linear dose-response, Expert 1 argued the high-dose tumors cannot be dismissed. Expert 5 sided with the threshold/human-relevance concern, and Expert 4 similarly criticized the "dose-related increase" framing as misleading and the rejection of Bogdanffy's cytotoxicity mode of action without an alternative.
• Acidity/decomposition as a confounder:Expert 2 raised whether acetic acid (263 ppm, pH 3.6) could have contributed to oral mucosal irritation and tumors; Experts 1 and 5 agreed this cannot be ruled out and that measurement of acetaldehyde and acetic acid would have helped, though Expert 2 still viewed the response as a high-dose VAM effect.
• BMD modeling:Expert 4 questioned pooling rat and mouse data for the BMDL10 given species differences in response, and Expert 5 noted no survival adjustment was described; Expert 1 by contrast considered the data suitable for quantitative assessment.
• Species/strain differences(weak response in male rats vs. stronger in mice) were noted by Expert 4, with Expert 5 speculating about differential carboxylesterase activity in oral mucosa.

AI generated summary content

ANSWER EXPLANATIONS

Expert 1 | **SciTrust Score:** 9

	Confidence Rating
Methods	9
Results	9
Discussion/Conclusions	9

Methods

Study performed along guideline regarding design, endpoints, duration, animal model. Experienced lab, detailed description of methodology.

Stability of VAM in dosing solutions assessed, some underdosing likely

Results

Detailed description of observations in tables and figures, both for tumor and for preneoplastic lesions. Tables also include incidences in historic controls.

Increased incidences of preneoplastic lesions and tumor at site of 1st contact, dose-dependent.

Changes in systemic tumor incidences unlikely related to test item admin. (no dose-dependence, increases not statistically significant
High conc. may represent MTD

Discussion/Conclusions

Discussions covers histological characteristics of tumors, context of observations made with those of other studies with VAM, role of toxicity induced by hydrolysis products including estimates of doses received by other studies. All conclusions made well supported by data.

Summary: High quality study confirming that oral admin. of VAM in high concentrations induce tumors of the oral cavity. Data can be used for quantitative assessment. Absence of systemic tumors.

Expert 2

	Confidence Rating
Methods	
Results	
Discussion/Conclusions	

No explanation provided.

ANSWER EXPLANATIONS, CONTINUED

Expert 3 | **SciTrust Score:** 6.5

	Confidence Rating
Methods	7
Results	8
Discussion/Conclusions	5 (medium confidence)

Methods

An appropriate species was selected. The number of animals per dose group was sufficient. The study duration was typical near the lifespan of the animal but not so far as to have complications in reviewing tumors against control animals that live their full lifespan. The three concentrations and control were appropriate, although the high concentration was likely beyond the toxic level in mice when comparing body weight gains. The length of exposure was relevant. The methods were well described, but the test substance was not characterized for purity. The documentation of actual versus administered dose was handled appropriately, and food and water were measured per dose group. A complete necropsy was performed on animals including those that died early, and collection of tissues appeared to be processed and examined according to standard protocols. It was not immediately apparent that histopathology was performed blindly nor subjected to an independent pathology review. The statistics were appropriate. The study met good laboratory practices.

Results

There did not appear to be any disease outbreaks in the experimental and control animals. Survival was adequate across those groups to allow a meaningful comparison of tumors. There was no apparent excessive toxicity at high doses, although this may be question in high dose mice, or reduced water our feed intake that would affect results scene. The data were appropriately summarized, and tumor incidences were appropriately reported by group, sex and site. The findings appeared to be reported completely.

Discussion/Conclusions

The authors conclusions were supported by the data presented, and their conclusion regarding this chemical's carcinogenicity at the highest concentration was appropriately justified. The limitations of the study were a few, as discussed above, but, in addition, the authors failed to acknowledge the obvious threshold nature of their results. So, while the chemical is possibly carcinogenic at high dose, its ability to cause such tumors at lower concentrations, those expected in humans, is limited.

ANSWER EXPLANATIONS, CONTINUED

Expert 4 | **SciTrust Score:** 8

	Confidence Rating
Methods	8
Results	8
Discussion/Conclusions	8

Methods

The dose selection made for this study was not clearly defined. The composition of the material was evaluated but not fully reported.

Results

Overall the results are clearly presented, and appropriate statistical analysis was performed. Although not emphasized in the results or discussion sections, only the high dose 10,000 ppm VAM produced squamous carcinomas. Even at this high dose the tumor response, while statistically significant was relatively weak in rats (male rats 5/50) while being more robust 13/50 and 15/50 in male and female mice respectively.

A BMDL10 was determined by apparently combining of the data from all rodents (rats and mice). Is that an appropriate approach given variability in the response of different species (weak response in male rats and a stronger response in mice). The BDL10 analysis should have been performed on each species separately for comparison to the combined results.

Discussion/Conclusions

The discussion appropriately addressees the finding of the current study with the literature. A suggestion for the difference in the responses seen between this study and previous studies are noted. The discussion and the results sections refer to a "dose related increase in squamous cell tumors". While technically correct this is a little misleading given that only the high dose (10,000 ppm) induced squamous cell carcinomas in the rats and mice which points to a high dose only effect. This point should have been further discussed. The authors also reject the Mode of action cytotoxicity hypothesis by Bogdannffy without offering an alternate.

ANSWER EXPLANATIONS, CONTINUED

Expert 5 | **SciTrust Score:** 8

	Confidence Rating
Methods	7
Results	9
Discussion/Conclusions	8

Methods

1) Score=0.45. The study use BDF mice and F344 rats, both from Charles River in Japan. These are commonly used strains for long-term bioassays.

2) Score=0.45. 50 animals of each sex per dose group were used consistent with OECD guidelines.

3) Score=0.45. The control group received drinking water without added VA.

4) Score=0.45. The study duration was 104 weeks, appropriate for a cancer bioassay.

5) Score= 0.3. The doses were appropriate save for the high dose of 10000 ppm.

6) Score=0.2. The high dose of 10000 was above the the range of doses used 10 years earlier by Bogdanffy et al., which was cited. No explanation for the use of 10000 ppm was given and the concept of MTD was not mentioned.

7) Score=0.4. The authors mentioned the unknown relevance to human from ACGIH.

8) Score=0.1. The purity and stability of the test agent was measured “before and at the end of its use. Whether this means twice a week when the water bottles were filled or at week 0 and week 104 is not known.

9) Score=0.45. The dose formulations were measure biweekly.

10) Score = 0.4. Water consumption was measured weekly for 13 weeks and then every 4 weeks.

11) Score=0.4. Body weight, water and food consumption were measured weekly for 13 weeks and then every 4 weeks.

12) Score=0.45. The tissues examined were those specified in OECD guidelines.

13) Score=0. No mention was made of the number of pathologists used or whether an independent pathology exam was conducted.

14) Score = 0. Nothing was written about the pathologists who examined the tissues.

15) Score=0.4. Peto’s test, a version of the Cochran-Armitage test was used for dose response and Fishers exact test was used to compare tumor rates against controls.

16) Score=0. EPA’s BMDS v. 1.3.1 was used for estimating the BMDL. No mention is made of survival adjustment.

17) Score=0.45. The squamous cell carcinomas in the esophagus and larynx were compared with historical controls (p. 90-91,

Tables 2,5).

18) Score = 0.45. Oral cavity squamous cell in mice and rats were distinguished in Tables 2, 4, 5 and 7.

19) Score=0.45. Dunnet's test was used for multiple comparisons with controls (p. 89).

20) Score =0.45. The histopathology and other methods were consistent with OECD guidelines and these guidelines were cited.

21) Score = 0.45. No deviations from the guidelines were apparent.

Results

1) Score=1. No infectious disease outbreaks were apparent ,and the animals were housed individually.

2) Score=1. Survival rates for mice and rats are given in Table 1. Survival rate did not significantly differ between dose groups (p. 89).

3) Score =0.7. For both rats and mice, no consistent VA effects on hematology parameters was observed.

4) Score=1. Water consumption averages for each dose group are show in Table 1 and show a consistent decrease in the high dose group in both rats and mice.

5) Score=1. The data are summarized in the tables, which are easy to understand.

6) Score =1. Tumor incidence is reported fully in the tables.

7) Score=1. Tumor results were reported by site in Tables 2, 3, 4 for mice and Tables 5, 6, 7 for rats. All findings appear to be reported.

8) Score =0.7. Table 1 shows statistically significant body weight differences in the 10000 ppm groups of M and F mice, an M rats at the end of the study. These differences did not appear to affect the bioassay results.

9) Score=1. No disease affected the results and the animals were housed individually.

10) Score = 1. Table 3 shows the pre-neoplastic lesions consistent with the tumor findings.

Discussion/Conclusions

1) Score=1. The conclusions are indeed supported by the data. The comparison made between the results of Lijinsky & Reuber 1983 and Bogdanffy et al. 1994 was careful and provided a good explanation for the difference in the results between those papers and Umeda 2004.

2) Score=1.1. The data presented are consistent with the carcinogenicity of VA and the authors provide a well-reasoned explanation noting both neoplastic and pre-neoplastic lesions.

3) Score=1.2. The study was not deemed negative but did have sufficient power, exposure and duration to detect an effect, which was indeed observed.

4) Score=1.2. The conclusions are appropriately qualified and few study issues are present.

5) Score =0.9. The limitations of the study were few and the impact on conclusions are well reasoned. The authors could have

added additional detail from the Morris study (ref. 28).

6) Score=1.2. The uncertainties in the data are indeed acknowledged.

7) Score=1.2. The language used is consistent with the strength of the evidence.

Debate | 3 comments

Expert 2 | 10/12/2026 9:29 SCORE: **0**

The authors discuss stability of VA in DW solutions and note that at 10,000 ppm the acetic acid concentration was 263 ppm (pH 3.6) at the end of 4 days (acetaldehyde as a decomposition product not measured). Although it is possible that chronic exposure to acetic acid (as a decomposition product) caused some irritation in the oral mucosa, can it confidently be ruled out that acetic acid in the DW solution did not contribute to the oral tumor response?

3 Replies

Expert 1 | 10/12/2026 10:16 SCORE: **0**

I would not rule out a role of acetic acid. If a drinking water study of acetic acid is available, this may clarify

Expert 2 | 10/13/2026 9:15 SCORE: **0**

Only a gavage chronic study on acetic acid. Can't rule out that the pH of 3.6/acetic acid could have mild effects and be a potential confounder in possibly enhancing the local effects of VA but agree this is a high dose effect of VA.

Expert 5 | 10/14/2026 8:21 SCORE: **0**

Would have been good if acetaldehyde and acetic acid were measured in the drinking water when the water bottles were changed out. Perhaps doing so could distinguish effects of endogenous vs. exogenous VA metabolites.

Expert 5 | 10/12/2026 10:07 SCORE: **0**

I agree with #4 about the apparent threshold observed and the lack of human relevance of the high concentrations used.

1 Reply

Expert 5 | 10/12/2026 10:12 SCORE: **0**

Thinking again, Umeda used CR BDF1 mice and F344 rats. I'm wondering if the strain differences I mentioned could result from differences in carboxyesterase activity in the oral mucosa.

Expert 3 | 10/14/2026 8:59 SCORE: **0**

Dear Colleagues, I am having a hard time understanding the ratings of 8 or 9 for the discussion/conclusions of this study. Unlike other studies, there was no in utero exposure nor vaccination of animal nor interim kills, contaminants were not monitored nor discussed, and the high dose appears to be toxic in mice, where all of the statistically significant tumors occurred. This is not really surprising since the Ph of the drinking water was pH 3.6. Based on your comments, however, I might be willing to give this rating a "6."

1 Reply

Expert 1 | 10/14/2026 10:10 SCORE: **0**

The testing guideline (OECD 453) does not require in utero exposure or interim sacrifices. Young adults (age 6 weeks) should be exposed by the pathway most relevant to human exposures and exposure should occur for 2 years. Study was performed along OECD 453 guidance under GLP in a well-established lab and all relevant parameters were assessed. Therefore, my high score.

Regarding the high dose, the tumors may be a consequence of toxicity, but they cannot be dismissed. I agree that dose-response is likely non-linear due to cytotoxicity as an important driver for tumor formation.

SciPinion

SciPi #1052

October 19, 2026

Round 3

SciPinion L.L.C.
777 E. Main St., Suite 202
Bozeman, MT 59715

shays@scipinion.com
ckirman@scipinion.com
(406) 404-6145

How should these studies be used in human health risk assessment? In your explanation below please provide your rationale for each of the studies you considered

Results | 5 answers

	THIS STUDY CAN BE USED FOR QUANTITATIVE RISK ASSESSMENT	THIS STUDY SHOULD ONLY BE USED QUALITATIVELY, FOR HAZARD IDENTIFICATION PURPOSES	THIS STUDY SHOULD NOT BE USED FOR RISK ASSESSMENT	OTHER	TOTAL
Belpoggi et al. (2002)	0.0% 0	0.0% 0	80.0% 4	20.0% 1	5
Bogdanffy et al. (1994; inhalation)	60.0% 3	20.0% 1	0.0% 0	20.0% 1	5
Bogdanffy et al. (1994; drinking water)	80.0% 4	0.0% 0	0.0% 0	20.0% 1	5
Lijinsky and Reuber (1983)	0.0% 0	0.0% 0	80.0% 4	20.0% 1	5
Maltoni et al. (1997)	0.0% 0	0.0% 0	100.0% 5	0.0% 0	5
Minardi et al. (2002)	0.0% 0	20.0% 1	80.0% 4	0.0% 0	5
Umeda et al. (2004)	60.0% 3	20.0% 1	0.0% 0	20.0% 1	5

Strong consensus against the Ramazzini Institute studies. All experts agreed that Maltoni et al. (1997) and Minardi et al. (2002) should be excluded from risk assessment (Expert 4 allowed Minardi only for qualitative hazard identification), and all who reviewed Belpoggi et al. (2002) reached the same conclusion. Recurring criticisms included inadequate pathology reporting and possible misclassification of carcinomas, unknown randomization and litter effects, unverified test-substance stability and actual intake, insufficient group sizes, and absent or poor statistical and survival analyses. Experts emphasized that the similarity of findings across these studies does not constitute independent confirmation, since they share a single laboratory and common methodological flaws, and that biological plausibility alone cannot offset such uncertainties.

Lijinsky and Reuber (1983) was likewise rejected by all reviewers who evaluated it, citing too few animals, no defined sacrifice schedule, undetermined test-article purity and stability (with likely exposure to decomposition products), lack of histopathological detail, and no statistical analysis.

The Bogdanffy et al. (1994) studies were judged highest quality, though with some divergence on application:

• Inhalation: most experts (1, 2, 4) endorsed it for quantitative dose–response and risk assessment; Expert 5 would restrict it to qualitative hazard identification, noting that only total nasal tumors reached statistical significance.
• Drinking water: Experts 1, 3, and 4 supported quantitative use; Expert 2 considered it suitable for hazard assessment and NOAEL derivation; Expert 5 chose "Other," suggesting its main value lies in combination with mechanistic data given the absence of treatment-related tumors.

Umeda et al. (2004) generated the most disagreement. Experts 1, 4, and 5 accepted it for quantitative purposes, citing GLP/OECD consistency, adequate group sizes and duration, measured water consumption, and historical control data—though Expert 1 flagged the need to correct doses for decomposition, and Expert 5 argued the oral cavity tumor response appears threshold-based, favoring a reference dose over the authors' BMDL10 of 476 mg/kg/d. Experts 2 and 3 were more cautious: both highlighted the dose-dependent pH drop (3.6 at the high dose) and unmeasured contaminants as confounders of irritation-driven carcinogenicity, with Expert 2 further arguing the MTD was likely exceeded in mice and that responses represent a high-dose phenomenon of questionable human relevance; Expert 3 would limit the study to hazard identification.

Cross-cutting themes. Several experts stressed that excluding a study from quantitative use does not mean ignoring it—limitations and uncertainties should still be characterized in the assessment. There was also broad concern across studies about test-substance stability, acidity/irritation as a mode of action for oral cavity lesions, and the large margin between experimental doses and realistic human exposures.

AI generated summary content

ANSWER EXPLANATIONS

Expert 1

Study	This study can be used for quantitative risk assessment	This study should only be used qualitatively, for hazard identification purposes	This study should not be used for risk assessment	Oth er
Belpoggi et al. (2002)	0	0	1	0
Bogdanffy et al. (1994; inhalation)	1	0	0	0
Bogdanffy et al. (1994; drinking water)	1	0	0	0
Lijinsky and Reuber (1983)	0	0	1	0
Maltoni et al. (1997)	0	0	1	0
Minardi et al. (2002)	0	0	1	0
Umeda et al. (2004)	1	0	0	0

Studies from the Ramazinni-Lab (Belpoggi et al.,; Maltoni et al., ; Minardi et al.,) should not be used for risk assessment due to issues with study design and reporting of tumor incidences. In addition, the pathological assessment was questioned by some of the experts.

Lijinski and Reuber has a number of limitations such as limited number of animals, study design makes conclusions difficult (no defined point of sac), animals likely exposed to significant amounts of decomposition products,

Both Bogdanffy studies are of high quality, followed respective guidance and have adequate reporting regarding effects relevant for risk assessment. The inhalation study can be used for quantitative assessments

Umeda et al. is also a high quality study with detailed reporting and characterization of lesions/incidences. May also be used for risk assessment regarding local effects of VAM to the upper GI tract, systemic tumors observed do not show dose-response and are not considered as related to VAM by the study authors. When using the study for quantitative assessment, correction of administered doses for decomposition may be required.

ANSWER EXPLANATIONS, CONTINUED

Expert 2

Study	This study can be used for quantitative risk assessment	This study should only be used qualitatively, for hazard identification purposes	This study should not be used for risk assessment	Other
Belpoggi et al. (2002)	0	0	1	0
Bogdanffy et al. (1994; inhalation)	1	0	0	0
Bogdanffy et al. (1994; drinking water)	1	0	0	0
Lijinsky and Reuber (1983)	0	0	1	0
Maltoni et al. (1997)	0	0	1	0
Minardi et al. (2002)	0	0	1	0
Umeda et al. (2004)	0	0	0	1

Context for my rationale:
1. Risk assessment is the integration of hazard, dose-response and exposure.
2. Study quality and reliability is critical when making judgements of human hazard and risk.
3. Although I considered certain published chronic bioassays to be of limited evidentiary value and unsuitable for risk assessment, that does not mean they shouldn't be acknowledged and discussed in the assessment with a thorough characterization of the limitations and uncertainties.

My rationale for each study follows.

Bogdanffy et al, Inhalation. This is a high quality study with sufficient reporting that can serve as the basis for hazard, dose response and risk assessment. There were no deficiencies that would discount the findings/conclusions.
[Note: If the actual laboratory report is needed for raw data, methodological details etc, it might be obtained with the sponsor's permission.]

Bogdanffy et al, Drinking water. This is a high quality study with sufficient reporting that can serve as the basis for hazard assessment and for establishing a NOAEL. There are no deficiencies that would discount the findings/conclusions.
[Note: If the actual laboratory report is needed for raw data, methodological details etc, it might be obtained with the sponsor's permission.]

Belpoggi et al, Drinking water, As noted in my responses to Round 1 (as well as agreement from other reviewers during the debate), there are serious limitations and uncertainties that significantly reduce the evidentiary weight of this study. Although the findings in the oral cavity may be biologically plausible, that is not enough to justify the use in risk assessment (i.e., biological plausibility doesn't resolve the methodological limitations and uncertainties). The issues that raised concern include; lack of pathological detail (don't know if diagnosis was blinded) and potential misclassification of carcinomas (no evidence was presented to support invasion); it is not known if randomization was done to assign animals to dose groups, which is a concern because the effects are in the F1s (offspring) and litter of origin is not provided (i.e., it is not known if the findings are independent of litter); uncertain whether infection in the animal colony is an issue; it is not known when the tumors appeared (at end of life or sooner?); lack of measurements to ensure stability of VA and lack of information on actual drinking water intake; poor statistical analysis, historical data is not available to understand background incidence and its variability in aging rats. Because Belpoggi et al, Minardi et al, and Maltoni et al used different species/strains and observed similar findings, this should not be interpreted as independent confirmation because the studies are from the same

laboratory (Ramazzini Institute) and share common problems.

Minardi et al, Drinking water, As noted in my responses to Round 1 (as well as agreement from other reviewers during the debate), this study has serious limitations and uncertainties. This study shares common problems with Belpoggi et al. (e.g., lack of VA characterization, unknown randomization across dose groups, lack of pathological detail/confirmation). There was also a lack of BW and survival data, and data for DW and food consumption. Simply because Minardi et al used a different strain of ratand observed similar findings to the other Ramazzini studies, this should not be interpreted as independent confirmation because the studies are from the same laboratory and share common problems. Also, as pointed out, biological plausibility is insufficient to justify use of an unreliable study due to significant limitations and uncertainties.

Maltoni et al, Drinking water, This study shares common problems with the other Ramazzini Institute studies, e.g., lack of VA characterization for VA stability in DW solutions, unknown randomization across dose groups, lack of pathological detail/confirmation, as well as, no scheduled sacrifice. Because of the shared problems and being conducted at the same laboratory, it is not considered to be an independent confirmation of findings.

Lijinsky and Reuber, Drinking water, there are too many limitations that render this study unreliable, including insufficient number of animals tested, undetermined VA stability (didn't prepare fresh DW solutions daily and actual administered dose is not known), used a commercial product without rigorous purification, no statistical analysis, lack of histopathological detail.

Umeda et al, Drinking water, This study has some methodological issues. Authors did not provide information on impurities. The DW solutions were not prepared daily (twice a week). There was a dose dependent decrease in pH (as well as an increase in acetic acid; acetaldehyde not measured), which confounds interpretation in that it can not be dismissed that the drop in pH had an irritation effect with chronic exposure. The rationale for the selection of 10,000 ppm as the HDT is not provided. Human exposure via drinking water is very low. At 10,000 ppm, there was a significant drop in body weight at the end of study for both sexes of mice (M-~30% decrease; F ~18% decrease). It appears that the MTD was likely exceeded at highest dose tested (estimated to be ≥989 mg/kg/day) which raises the concern that the mouse oral cavity carcinomas are due to an excessive chronic dose. In the rat study, there were acceptable decreases in BW. There was an increases in oral cavity carcinomas (males 0/50 vs 5/50; females 0/50 vs 3/50) at the 10,000 ppm dose (estimated to be 442 mg/kg/day). There is no historical control range mentioned in the paper but OEHHA indicated the rate was 0.09% in females. Although the pH decrease confounds interpretation, if indeed the laboratory HC is virtually zero, the 3-5 carcinomas cannot be entirely dismissed, but the responses are likely a high dose phenomenon not relevant to human exposures.

ANSWER EXPLANATIONS, CONTINUED

Expert 3

Study	This study can be used for quantitative risk assessment	This study should only be used qualitatively, for hazard identification purposes	This study should not be used for risk assessment	Other
Belpoggi et al. (2002)	0	0	0	1
Bogdanffy et al. (1994; inhalation)	0	0	0	1
Bogdanffy et al. (1994; drinking water)	1	0	0	0
Lijinsky and Reuber (1983)	0	0	0	1
Maltoni et al. (1997)	0	0	1	0
Minardi et al. (2002)	0	0	1	0
Umeda et al. (2004)	0	1	0	0

I did not have a chance to review Bellpoggi et al. (2002), Bogdanffy et al. (1994) or Lijinkkshy and Reuber (1983) and so marked the "Other" column. The Bogdanfffy et all (1994 water) study was bodacious, warranting high marks by all reviewers. The Umeda et al. (2004) was GLP and run well, but the high dose drinking water had a pH of 3.6. Nor were contaminants measured. However, based on my colleagues comments I will raise my score from 5 to 6 for the discussion/conclusion part of my evaluation. The Maltoni et al. (1997) and Minardi et al. (2002) do not make a sufficient contribution to the literature other than to confirm that high doses of acid water will cause damage in the oral cavity and that this damage will sometimes lead to tumors. These studies, while perhaps useful in a theoretical sense, should not be used practically for human health risk assessment.

Expert 4

Study	This study can be used for quantitative risk assessment	This study should only be used qualitatively, for hazard identification purposes	This study should not be used for risk assessment	Other
Belpoggi et al. (2002)	0	0	1	0
Bogdanffy et al. (1994; inhalation)	1	0	0	0
Bogdanffy et al. (1994; drinking water)	1	0	0	0
Lijinsky and Reuber (1983)	0	0	1	0
Maltoni et al. (1997)	0	0	1	0
Minardi et al. (2002)	0	1	0	0
Umeda et al. (2004)	1	0	0	0

No explanation provided.

ANSWER EXPLANATIONS, CONTINUED

Expert 5

Study	This study can be used for quantitative risk assessment	This study should only be used qualitatively, for hazard identification purposes	This study should not be used for risk assessment	Other
Belpoggi et al. (2002)	0	0	1	0
Bogdanffy et al. (1994; inhalation)	0	1	0	0
Bogdanffy et al. (1994; drinking water)	0	0	0	1
Lijinsky and Reuber (1983)	0	0	1	0
Maltoni et al. (1997)	0	0	1	0
Minardi et al. (2002)	0	0	1	0
Umeda et al. (2004)	1	0	0	0

Belpoggi et al (2002) should not be used for risk assessment. The reasons are that the doses used are orders of magnitude greater than human exposures, the lack any clear dose-response, and the lack of adherence to any guidelines for study quality.

Bogdanffy et al. (1994a) was an inhalation study of SD-derived rats and Swiss mice using inhalation concentrations of 0, 50, 200 and 600 ppm. Treatment-related increases in squamous cell carcinomas of the respiratory regions were observed in rats but not in mice. The responses of individual tissues in rats were small and the only statistically significant change was for total nasal tumors.

Bogdanffy et al. (1994b) was a two-gen drinking water study in rats. The doses used were 0, 200, 1000, and 5000 mg/L. Neither systemic toxicity nor treatment-related cancer effects were observed. The study was well-conducted but remains inconsistent with other studies. Whilst the study could be used to rule out any positive hazard ID for VA, such use would be unlikely. The best use would be in conjunction with mechanistic studies regarding the conditions and exposures needed for tumor occurrence. Hence, I marked "Other."

Lijinsky & Reuber (1983) relied on two doses and a control with only 20 animals per dose group. The authors failed to address many of the issues raised in the questions and the overall validity of the study cannot be established. It should not be used for risk assessment.

Maltoni et al. (1997) dosed Swiss mice with 0, 1000 and 5000 mg/L VA in drinking water. The breeder male dose groups had 14,13 and 13 individuals. The breeder female dose groups all had 37 individuals. The F1 male dose groups had 38, 37 and 49 individuals. The F1 female dose groups had 48, 44, and 48 individuals. Hence, none of the dose groups had a sufficient number of individuals.The dosing period was 78 weeks, the lowest recommended duration. The survival of all mice started to drop off between 40-60 weeks. Neither statistical analysis nor survival analysis was conducted. The study did not follow accepted guidelines and should not be used for risk assessment.

Minardi et al. (2002) dosed SD rats withe 0, 1000 and 5000 mg/L VA in drinking water. The breeder male dose groups had 14,13 and 13 individuals. The breeder female dose groups all had 37 individuals. All F1 dose groups had over 50 individuals. The actual water concentrations were not measured. No discussion of statistical analysis or pathology methods was presented save for a reference to another paper; indeed, the methods section consisted of 5 sentences only. The paper should not be used for risk assessment.

Umeda et al. (2004) was a drinking water study in CRL BDF1 mice and CRL F344 rats. Dose groups contained 50 animals each. The water concentrations were 0, 400, 2000 and 10000 ppm. The study duration was sufficient at 104 weeks. Water consumption was measured and used to estimate doses. Survival rates did not significantly differ between dose groups. Historical control data was provided for oral tumors for both species. This paper was consistent with GLP and OECD guidelines and could be used for quantitative risk assessment. Although the authors used EPA's BMDS to predict a BMDL10 of 476 mg/kg/d for oral cavity squamous cell carcinomas, this occurrence appears to be a threshold phenomenon in male rats and in male and female mice, and thus, a threshold assessment, e.g., reference dose, might be a more appropriate outcome than a BMDL.