HHS Public Access

Author manuscript

Regul Toxicol Pharmacol. Author manuscript; available in PMC 2021 August 01.

Published in final edited form as:

Regul Toxicol Pharmacol. 2020 August ; 115: 104682. doi:10.1016/j.yrtph.2020.104682.

Meta-analysis of animal studies applied to short-term inhalation exposure levels of hazardous chemicals

Andrew J. Prussia[*], **Joshua Hill**, **Cheryl R. Cornwell**, **Ram C. Siwakoti**, **Eugene Demchuk**

Agency for Toxic Substances and Disease Registry, Atlanta, GA, USA

Abstract

For short-term chemical inhalation exposures to hazardous chemicals, the incidence of a health effect in biological testing usually conforms to a general linear model with a probit link function dependent on inhalant concentration C and the duration of exposure t. The National Academy's Acute Exposure Guideline Levels (AEGLs) Committee relies on these models when establishing AEGLs. Threshold concentrations at AEGL durations are established by the toxic load equation $C^n \times t = constant$, which toxic load exponent n (TLE or n-value) directly follows from the bivariate probit model. When multiple probit datasets are available, the AEGL Committee routinely pools studies' incidence data. Such meta-analytical models are valid only when the pooled data are homogeneous, with similar sensitivities and equivalent responses to exposure concentrations and durations. In the present study, the homogeneity of datasets meta-analyzed by the AEGL Committee was examined, finding that 70% of datasets pooled by the AEGL Committee are heterogeneous. In these instances, data pooling leads to a statistically invalid model and TLE estimate, potentially resulting in under- or over-estimated inhalation guidance levels. When data pooling is inappropriate, other meta-analysis options include categorical regression, fixed-effect and random-effects models, or even designation of a key study based on scientific judgement. In the present work, options of TLE meta-analysis are summarized in a decision tree contingent on statistical testing.

Keywords

Risk assessment; Short-term inhalation exposure; Ten Berge equation; Toxic load; Toxic load exponent (TLE); n-value; General linear model; Probit; Categorical regression; Meta-analysis; Acute exposure guideline levels (AEGLs)

[*]Corresponding author. Agency for Toxic Substances and Disease Registry, 4770 Buford Hwy NE, Atlanta, GA, USA, aprussia@cdc.gov (A.J. Prussia).

Publisher's Disclaimer: Disclaimer

The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the Agency for Toxic Substances and Disease Registry.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary materials

Supplementary materials to this article can be found online at https://doi.org/10.1016/j.yrtph.2020.104682. MatLab code for performing the probit analyses described here is available through the MATLAB Central File Exchange: https://www.mathworks.com/matlabcentral/fileexchange/73709-probit-meta-analysis-of-short-term-inhalation-studies.

1. Introduction

A key step in determining short-term inhalation exposure levels for hazardous vapors and gases is extrapolation to durations that were not tested experimentally (National Academy of Sciences, 2001). For example, the U.S. Environmental Protection Agency (USEPA) derives Acute Exposure Guideline Levels (AEGLs) for five durations and three health effect severity tiers (example in Table 1). These levels were designed as threshold exposure limits for the general public that could be applicable to once-in-a-lifetime or rare inhalation exposures ranging from 10 min to 8 h. AEGLs are typically derived from short-term inhalation toxicology experiments in animals, where groups of animals are exposed at controlled air concentrations of a chemical for pre-determined durations. Almost no chemical has inhalation experiments tested in mammals for each severity tier and each AEGL duration. Therefore, the vast majority of AEGLs are extrapolated by USEPA from experimental durations according to procedures advised by the National Academy of Sciences (NAS) (National Academy of Sciences, 2001). There are many other government and non-governmental guidelines for acute inhalation exposures, each serving a designated public health safety mission. However, the multiple tier and duration approach pertinent to AEGLs is aligned with duties of the Agency for Toxic Substances and Disease Registry (ATSDR), particularly in emergency response situations.

For deriving AEGLs, NAS recommends using a generalized linear model (GLM) to relate concentration and exposure duration to the health effect incidence monitored in inhalation experiments in animals. GLM with a probit link function is also known as the multivariate probit regression. This is similar to the classical toxicological median lethal dose (LD_{50}) concept, but with the added dimension of time (Finney, 1977):

$$Y = \alpha + \beta \log C + \gamma \log t \tag{1}$$

where Y is the probit of the incidence rate, α is the intercept on the probit axis, β and γ are the slopes of the plane on the log C (concentration) and log t (duration) axes, respectively.

Such an analysis allows calculation of a chemical concentration that has a certain probability of causing an adverse health effect at the specified duration of exposure. For AEGL-3 (i.e. mortality) points of departure (POD), this is often the LC_{01} concentration, or the concentration with a 1% probability of causing death after a specified exposure duration (however, the NAS-preferred method is the lower 95% bound on the 5% probability). Fig. 1a illustrates the relationship expressed as a three-dimensional plane obtained by fitting GLM to the data of a typical mortality incidence experiment.

Experiments confirm that for a given toxic effect, almost all airborne chemicals show a linear relationship on the log-log scale of C and t, but with slopes on the log scale that may be different from 1, thus, deviating from Haber's rule:

$$TL = C^n \cdot t \tag{2}$$

where n is the toxic load exponent (TLE), a chemical-specific constant.

Equation (2), proposed more than a century ago, is commonly referred as the ten Berge equation because of his first systematic survey of these toxicological parameters for 20 chemical inhalants (ten Berge et al., 1986). For a given toxic load, such as necessary to produce an LC_{01} level of effect, Equation (1) can be rearranged to the ten Berge form:

$$Y' = \alpha' + C^n \cdot t \tag{3}$$

where $n = \frac{\beta}{\gamma}$; β and γ are the slopes of the plane in the log C and log t axes, respectively, from Equation (1).

The constant α' adjusts for toxicological sensitivity to the inhalant and the TLE, n, also known as n-value, is the ratio of the parameters for concentration (β) and exposure duration (γ). Fig. 1b illustrates that regardless of algebraic rearrangements this relationship remains linear with a constant slope at fixed TL, because of the constant dihedral angle of the plane.

When incidence data allow, the probit regression is effective at interval estimation of the TLE (and even directly PODs). However, for many chemicals such information is unavailable. For these chemicals, the NAS recommends an alternative method. A point estimation of the TLE can be carried out based on Equation (3) using a simple linear regression at a fixed toxic load (usually the toxic load necessary to produce an LC_{50} level of effect) at multiple durations (National Academy of Sciences, 2001), because Equation (3) can be log-transformed to:

$$\log C = -\frac{1}{n} \cdot \log t + \alpha'' \tag{4}$$

where α'' is the intercept on the concentration axis.

This method may be appealing when LC_{50}s are reported without experimental incidence data; however, this method does not carry over the statistical uncertainties of the LC_{50}s into the interval estimation of the TLE. It is also enticing for the risk assessor to pool LC_{50}s from multiple studies in order to derive a TLE, without regard to the heterogeneity between these studies. A number of TLEs from the AEGL technical support documents (TSDs) are derived this way.

AEGL TSDs represent the largest compendium of peer-reviewed toxicological information pertinent to short-term inhalation exposures. Therefore, it was chosen as the source of information for the present study. In the present study, incidence data for chemicals meta-analyzed by the AEGL Committee were scrutinized using the full range of statistical tests of the probit meta-analysis framework. These data satisfy two conditions: availability of incidence information from *multiple* animal experiments and relevance of the AEGL reference approach to the ATSDR mission (multiple durations and severity tiers). The framework relies on established statistical methods for determining if a chemical's data from multiple studies may be pooled, compliant with a statistical hypothesis of parallel probit planes, or if the planes are non-parallel. When the incidence data could be pooled, the standard probit regression was applied. When the planes were parallel, the categorical probit regression was used. When the planes were non-parallel, fixed and random effects model

analyses were performed on the constituent TLEs to estimate the summary effect TLE. For each chemical, either a statistically sound TLE was derived or an attempt to derivation was made. The outcomes were compared to results of analyses described in AEGL TSDs.

2. Methods

2.1. Assembly of the AEGL incidence data

Upon analyzing the data from 273 AEGL TSDs, it was found that 115 empirically-derived TLEs were identified by the AEGL Committee. Of them, 90 TLEs originated from mortality studies. Relevant experimental details such as study source, species, and the method of TLE derivation were also collected. Of the collected TLEs, 14 were based on pharmacokinetic modeling, 62 were calculated by the AEGL Committee using simple linear regression and 39 stemmed from probit analysis. For 41 chemicals, the AEGL Committee pooled data from multiple studies. These were either binomial incidence data (15 chemicals) or LC_{50}s (26 chemicals). For some chemicals, multiple TLEs have been derived for different severity tiers or species. For both 13 chemicals with pooled LC_{50}s and 15 chemicals for which the probit analysis was used, the binomial incidence data were available. For the remaining 13 chemicals, multiple-duration incidence data were not available so these chemicals were not processed further in this study. Thus, the final dataset contained binomial incidence data for 28 AEGL chemicals either directly from AEGL TSDs or from their cited literature (refer to supporting information spreadsheet: AEGL pooled studies - dataset.xlsx). This database constituted an exhaustive collection of chemicals with incidence data from multiple studies reviewed by the AEGL Committee, thus, excluding author's bias in data selection.

2.2. Calculation of probit regressions

The first step for analysis of the AEGL incidence data was to calculate probit regressions using a procedure similar to the "Concentration × Time" option in the USEPA's Benchmark Dose Software (BMDS) 2.7 (USEPA, Washington, DC). The procedure was encoded in MatLab® 2017b (MathWorks, Inc., Natick, MA) to facilitate batch processing. The results of test calculations were identical to those of BMDS within the machine precision. Using incidence data from the original experimental studies, multivariate probit regressions were calculated for each chemical's study using the natural logarithm-transformed values for concentration and time as in Equation (1). In these calculations, each parameter of Equation (1) was estimated, along with associated uncertainty (i.e. the interval estimation was performed). Using the variances for β and γ, the variance in TLE can be calculated using the error propagation equation for ratios (Ku, 1966).

2.3. Parallel and the goodness-of-fit statistical tests

After calculating the probit regressions for each study, Pearson's χ^2 statistic allowed determination of goodness-of-fit for the probit regression of a single study, χ^2_{study} (Finney, 1977). A *p*-value associated with the χ^2_{study} can be determined from the χ^2-distribution with the model's degrees of freedom; *p*-values < 0.05 were interpreted as the model predicting the experimental incidence poorly. When this is the case, a heterogeneity factor,

$\chi^2_{study}/model\ degrees\ of\ freedom$ was used to correct the variances of model estimates (U.S. EPA., 2008).

To test if the pooled studies for an individual chemical are parallel, a probit regression was calculated for each study, as well as a categorical probit regression that allows each study to have its own intercept on the probit axis (the parallel probit regression; Finney, 1977). Pearson's χ^2 statistic for goodness-of-fit was then calculated for each probit regression. Studies were deemed parallel if the p-value associated with the $\chi^2_{parallel}$ statistic was > 0.05:

$$\chi^2_{parallel} = \chi^2_{catreg} - \sum_{j=1}^{q} \chi^2_{studyj} \tag{5}$$

where q is the number of multiple studies, χ^2_{studyj}, χ^2_{catreg}, and $\chi^2_{parallel}$ are the χ^2 statistics for each study, categorical regression with study intercepts, and the parallelism test, respectively.

Note, the use of the goodness-of-fit p-values in this study is to indicate if the data are incompatible with statistical models, in agreement with principle 1 of the American Statistical Association's statement on p-values (Wasserstein and Lazar, 2016). If the studies for an individual chemical are parallel, random or common-effect meta-analysis is not required to explain the differences in responses. The studies are then measuring either similar response indicating the data can be pooled, or a response that is similar in slope but shifted in sensitivity, which means that the studies can be used together but with a categorical probit regression (parallel).

2.4. Meta-analysis of TLEs

For each chemical that was found to have non-parallel studies, the TLEs from the studies were evaluated as a single-group summary effect, with the TLE variance from the single-study probit regression used as the within-study variance. The common-effect model, also known as the fixed-effect model, and the random-effects model were used (Borenstein et al., 2010).

The uniformity among a chemical's TLEs was evaluated using Cochran's Q statistic and the I^2 statistic (the ratio of excess dispersion to total dispersion, Borenstein et al., 2009). With a Q statistic p-value ≤ 0.05, the hypothesis of uniformity for TLE effect sizes was rejected and the random-effects model was applied.

Additionally, the I^2 values for the TLEs of each chemical were evaluated using a statistical rule of thumb, by which I^2 < 30–40% suggests that heterogeneity might not be important (i.e. supports the common-effect model), while I^2 > 75% suggest considerable heterogeneity (i.e. supports the random-effects model; Borenstein et al., 2009). The I^2 range in between these extremes is guided by scientific judgement. Further in the text, the latter is illustrated by the case of phosgene, whose I^2 was 59%. Even though the underlying single-study data passed the Q test, the weight of evidence suggested that the phosgene data warrant application of the random-effects model.

3. Results

The AEGL Committee identified TLEs from pooled studies for 41 out of 79 unique chemicals with empirically-derived TLEs. For 15 chemicals with pooled data, the AEGL Committee derived meta-analytical TLEs by means of probit analysis. For the remaining chemicals, the Committee carried out simple linear regression analysis using Equation (4), even though incidence data for 13 of these chemicals were available. Thus, 28 groups of incidence data were assembled, all collected in animal mortality studies.

The first step of analysis involved testing for parallelism of probit planes fitted to the individual study data, as given by Equation (5). For each chemical, a categorical probit regression was performed using incidence data from all studies available for the given chemical, a probit regression with all data pooled, and individual probit regressions for each study. In cases when the AEGL Committee pooled a study with a single duration, that study's incidence data were also added to the pooled and categorical regression calculations, but an individual probit regression could not be performed. The resulting TLEs were classified as categorical regression TLEs, pooled-study TLEs, common-effect-model TLEs, designated-study TLEs (Table 2), and the random-effects model TLEs (Table 3). The tables also give confidence intervals (CIs) on the TLEs for all chemicals. The designated study approach was used when multiple studies could not be combined using any other method described above. Auxiliary information on the meta-analytical modeling, such as coefficients and *p*-values is presented in a supporting information spreadsheet (AEGL pooled studies - results.xlsx).

3.1. Categorical regression TLEs

For six of the chemicals, probit three-dimensional planes fitted to their individual study data were parallel and the coefficients for the categorical regression dummy variables were statistically significant, p-values < 0.05 (Fig. 2). The latter implied that each study required a different intercept on the probit Z-axis to explain the incidence data, although the slopes of the plains were statistically identical. Fig. 2b illustrates this case with ammonia data, showing two parallel probit planes. Parallelism of the planes means that the concentration-duration response is similar in each study, but that animal sensitivity varies depending on the test species and experimental conditions. In the case of ammonia, mice appeared to be more sensitive than rats (Fig. 2b). Risk assessors may use the most sensitive probit plane for POD derivation in order to be health protective without losing the response information provided by the other probit planes/studies.

3.2. Pooled-study TLEs

Incidence data for eight chemicals were comprised of parallel-plane studies, but their categorical regression dummy variables were statistically insignificant, p-values > 0.05 (Fig. 3). This implied that data from the individual studies do not require individual intercepts and that they can be pooled directly to fit a single probit plane (Fig. 3b). Data pooling means that with the given size of data, for each study, sensitivity of the cohort demonstrated in one study cannot be distinguished from another study. This is the way the AEGL Committee conducts probit meta-analyses. Similar to categorical regression, study pooling is

advantageous to risk assessors, because it provides greater statistical power and less uncertainty than designating a single key study.

3.3. Random-effects models for TLEs

For the half of studied chemicals, probit planes fitted to the individual study data did not pass the parallel test of Equation (Equation (5)). Studies may not be parallel for a variety of reasons. The studies may legitimately be measuring dissimilar responses, a study may be an outlier, or the studies may have high uncertainties. To investigate these possibilities, toxicological analysis of the literature pertinent to the failed group of chemicals was carried out. It identified four chemicals with incidence data hypothesized to be heterogeneous:

- Hydrogen chloride (HCl). The AEGL Committee relies on an HCl TLE estimate of ten Berge et al. (1986) (National Academy of Sciences, 2010). The authors have derived the TLE by probit regression on pooled mortality data in mice and rats exposed to gaseous HCl. Nonetheless, individual interval TLE estimates for mice and rats that follow from β and γ of Equation (1) published by ten Berge and co-authors are on the opposite sides of the default TLE = 1 (National Academy of Sciences, 2001). A TLE > 1 means that the response to a chemical is more dependent on concentration than exposure duration, while a TLE < 1 means the response is more dependent on duration than concentration. This apparent disparity may stem from discrepancies in the toxicological mode of action (MOA) of the chemical, which suggests that the mice and rat data may be heterogeneous.
- Hydrogen sulfide. The AEGL TSD derives a TLE using simple linear regression on rat LC_{50} data compiled from four independent studies (National Academy of Sciences, 2010). The TSD also suggests that the TLE transitions from a large value at exposure durations shorter than 1 h to smaller values at longer exposure durations. This indicates substantial heterogeneity in the underlying incidence data, possibly reflecting a transition of MOA from concentration-dependent at short durations to more time-dependent at longer durations. This may reflect differences in toxicokinetics, where short durations cause a near-point-of-entry effect, such as respiratory arrest following the paralysis of peripheral nerves, which is similar to hydrogen sulfide's effect on olfactory perception and inhibition of the olfactory nerve (National Academy of Sciences, 2010). Longer durations may cause systematic effects, such as acute metabolic acidosis following the inhibition of cytochrome oxidase.
- Oxygen difluoride. The AEGL Committee presents a TLE derived using probit regression on rat mortality incidence data compiled from two independent studies (National Academy of Sciences, 2014). It appears that the two studies captured different toxicological MOAs, similarly to the studies for HCl.
- Phosgene. The AEGL Committee arrives at a summary TLE based on the Zwart et al. (1990) study (National Academy of Sciences, 2002). In that study, the authors examine mortality in rats and mice using the probit analysis, but the rat data are fitted using an equation with a $\ln(C)*\ln(t)$ interaction term, while the

mice data are fitted without this additional term. In addition, there are general concerns about heterogeneity in phosgene rodent mortality data at short exposure durations due to reflexively induced variations in the animal lung tidal pattern (Li and Pauluhn, 2015).

To corroborate these hypotheses, statistical testing for heterogeneity was carried out. The TLEs were evaluated as a single-group summary effect using the TLE variance, considering both the common-effect and random effects models. The TLE variance was used as the within-study variance to calculate the weighted summary effect. The first three chemicals did not pass the Q test for homogeneity (p-values < 0.05) and had I^2 values greater than 75%. Therefore, they were treated using the random-effects model.

Phosgene represented a special case. Although the phosgene data passed the Q test, their I^2 was estimated to 59% (95% CI: 0–90%). A relatively high I^2, along with a large uncertainty, did not warrant a common-effect treatment, especially considering the variation in lung tidal volumes at short durations. Instead, the random-effects model was applied to the phosgene data.

The random-effects model uses both the within-study TLE variances and between-studies variance to estimate the summary effect TLE. Unlike the common-effect model, which expects that the true TLE is measured in all studies, the random-effects model assumes that each study estimates a different TLE. Together, these TLEs make a distribution. Because the distribution of TLEs is postulated to be normal, the mean of this distribution was calculated as the summary effect, along with CIs that give the range of TLE variation that may be observed for this chemical in similar studies (Fig. 5). It may be counter-intuitive to think that a chemical does not have a single true TLE, but perhaps it is not surprising that different studies, species, and experimental conditions do not yield the same answer. More important, the MOA may also not be the same. This inference directly concerns risk assessment, where the goal is to be protective of populations that may give a range of responses to a chemical exposure.

3.4. Common-effect-size TLEs

Incidence data for five of the ten remaining non-parallel-plane chemicals could be subjected to meta-analysis as well. For these, there was insufficient evidence to reject the hypothesis that the TLEs for each chemical were homogenous. Therefore, an alternative hypothesis was adopted, and these data were processed using the common-effect model (Fig. 4). Like the random-effects chemicals, appropriateness of the common-effect model was verified using the I^2 statistic. For all common-effect-treated chemicals, I^2 was less than 3% (see the supporting information spreadsheet), which suggest the TLEs are not heterogeneous. In a common-effect model, all available data are still utilized, but less certain studies contribute less to the resultant true TLE. A common-effect model is appropriate when neither the categorical regression nor pooling approach approximate the incidence data correctly, but the data are still sufficiently uniform to extract the true summary effect TLE.

3.5. TLEs from designated studies

The last five chemicals had multiple studies, whose data were neither parallel nor could be modeled using the fixed-effect or random-effects concepts. For each of these chemicals, multiple exposure durations have been reported only in one of the underlying studies. Thus, each chemical had only one study with an interval-estimated TLE (Fig. 6). Data from other available studies were non-parallel to the probit plane of the main study (Fig. 6b). For these five chemicals, it was not possible to determine if the single-duration studies provide more information about the response to the chemical, or just additional uncertainty. For these chemicals, the single-study probit analysis was preferred.

4. Discussion

In vivo animal studies are expensive, time-consuming, and raise ethical concerns, yet they remain the gold standard in chemical public health risk assessment. Much of the guidance to acute inhalation exposures to hazardous chemicals is based on such controlled studies, which are few compared to the gamut of chemicals and often limited in terms of experimental methodology. Therefore, it is of prime importance to use the maximum of available data in the most efficient and statistically appropriate way. Combining multiple studies is an apparent path to improving risk assessment. This goal is achieved by means of meta-analysis. However, meta-analytical methods are diverse, and rely on thoughtful application. They require in-depth understanding of how the choice of meta-analytical methodology affects data-driven conclusions, and they should be applied only under the correct statistical framework.

The methodology of short-term inhalation exposure levels relies on a TLE or directly on a probit regression to carry out temporal extrapolation (National Academy of Sciences, 2001). Based on the experience of the AEGL Committee summarized in AEGL TSDs and data analyses of the present report, the following decision tree for the meta-analytical modeling framework is proposed (Fig. 7). The decision logic suggests that multiple studies for one chemical may be treated by one of the five methods. It assumes that the studies are of the typical dichotomous-response design with independent groups of animals exposed at controlled concentrations and fixed exposure durations of the inhalant and that the monitored health effect represents binomial incidence. It also assumes that a bivariate generalized linear model regression (typically with the probit link function – “the probit regression”) is an appropriate statistical model for the response, which might not be the case for non-linear or non-normally distributed data. In particular, durations shorter than 10 min or longer than 8 h may require different assumptions (Verma et al., 2015). Of course, the toxicological quality of each study (appropriate controls, analytical methods, animal health, etc.) must be evaluated by the risk assessor prior to recruiting the study into the modeling database; goodness of model fit may yield insight but is not the ultimate gauge of toxicological quality.

Toxicological quality and compatibility of combined studies defines the meaning of meta-analysis. The studies may be measuring the same true TLE or may be measuring several different true TLEs. Contingent on that, the resultant meta-analytical estimate may express either the *true* TLE or a synthetic mean tendency of a collection of *different* TLEs. This is

similar to the NAS interpretation of the ten Berge et al. (1986) study, by which the limiting bounds of 1 and 3 on 90% of all TLEs known at that time are derived (National Academy of Sciences, 2001). These bounds are known as the default TLEs. They are recommended by the NAS for short-to-long and long-to-short time extrapolation, respectively, when an empirical chemical-specific value is not derived. Similarly, CIs on a synthetic mean TLE may be used to conduct temporal extrapolation in the NAS-recommended manner. There were about 15% of chemicals of this kind in the present study. For half of them the calculated CIs on all random-effects TLEs were more definitive than the default TLEs, and their lower bounds were more conservative. For example, by applying a TLE of 1, the AEGL Committee derives less conservative exposure levels for hydrogen chloride than they would be if the calculated chemical-specific interval estimate of 0.63–1.38 is used (Table 4). For other random-effects-treated chemicals the calculated CIs were less definitive than the default TLEs. For example, for hydrogen sulfide, both the lower and upper confidence bounds exceeded the default TLEs (Table 3). In this case, the default TLEs may be considered as an alternative. However, note that the methodology of a random-effects model entails estimation of a between-study variance (τ^2), which is postulated as an organic property of the studied data. This implies that additional studies may not necessarily result in reduction of CI on the TLE. A large non-reducible variability would imply that the laboratory data for the given chemical are such that the between-study TLE variability exceeds the within-study variance for different chemicals reported by ten Berge et al. (1986). Such a discord clearly defies the initial purpose of a chemical-specific TLE and questions applicability of the standard toxicological framework to the given chemical. For instance, for hydrogen sulfide an adoption of two different (possibly MOA-dependent) TLEs may represent a better public health approach than a single chemical-specific TLE. In any case, such unusual instances deserve a cautious conservative approach, additional research, and thorough contemplation.

For two-thirds of the studied chemicals, however, no statistical justification for denying the hypothesis of a single common TLE was found. For these chemicals, there was no reason to suspect that different studies require different TLEs to explain the data. These data could be explained with just one, an ostensibly *true* TLE. For a majority of the chemicals, the combined studies mutually agreed with each other even regardless of test species. The observed compatibility supports the NAS hypothesis of transferability of the TLE across species and from animal to human, even though the latter has never been tested.

For a few chemicals, no confident conclusion about meta-analytical quality of the TLE could be drawn. For these chemicals, only one multiple concentration-duration study per chemical was available, while single-duration studies disagreed with it. Since the single-duration studies were incompatible with the sole multiple duration study for a chemical, the latter was the designated study for the chemical's TLE. The designated-study and random-effects TLEs were summarized in the same table (Table 3).

If there is a need to combine more than two studies, the decision logic may need to be looped more than once to converge to a single summary-effect TLE. The below case of ammonia illustrates the procedure and provides an example of how the public health guidance may change depending on the method used to examine multiple studies (Table 5).

For the point of departure, the AEGL Committee uses the 1-h LC_{01} concentrations of 3374 and 3317 ppm from two mice studies, Kapeghian et al. (1982) and MacEwen and Vernot (1972), respectively. The other four durations are extrapolated using a TLE = 2 based on the meta-analytical probit regression performed by ten Berge et al. (1986) on incidence data pooled from 1-h study on mice by Kapeghian et al. (1982), 10-min study on mice by Silver and McGrath (1948), and a multi-duration study on rats by Appelman et al. (1982). However, when a categorical probit regression was applied, all individual study data came out, in fact, to be parallel and significant dummy variables appeared to apply. In this regression, dummy variables separating the mice studies were not significant, so the mice studies were pooled, but another dummy variable, differentiating between the rat and mice studies, turned out to be significant. Therefore, the mice studies were pooled, and the decision tree was then followed again, this time showing that the mice pseudo-study and the rat study are indeed parallel with a significant dummy variable, resulting in a TLE estimate of 2.13. Although this TLE estimate is similar to the original TLE, the uncertainty is considerably reduced, with 95% confidence intervals of 1.98–2.27 versus 1.60–2.40. Not only does this indicate the TLE is less uncertain, but PODs (such as LC_{01} concentrations at the 5 AEGL durations) from the categorical probit regression also carry less uncertainty. Directly using a pooled data approach results in unrealistically low LC_{01} values with very wide confidence intervals. This may be why the AEGL committee extrapolated from a 1-h LC_{01} with a TLE instead of using a probit regression, despite the availability of the incidence data.

Using categorical regression on the ammonia studies, LC_{01} concentrations were estimated for all five AEGL durations directly from the fitted probit function using not only the most sensitive species (mice), but also engaging the response information from the rat study, because the rat information improves the accuracy of the slope. Such approach yields a lower 1-h LC_{01} for mice of 2669 ppm and a slightly less steep concentration-duration response (Table 5). For POD-to-AEGL extrapolation, the AEGL Committee uses a total uncertainty factor of 3: 1 for interspecies differences because of high sensitivity of mice (ten Berge et al., 1986) and 3 for individual variability (National Academy of Sciences, 2008). Application of this uncertainty factor to the meta-analytical LC_{01} values for mice calculated at the AEGL-designated durations results in decreased short-term inhalation exposure levels by up to 25% (Table 5). An additional example of the application of this methodology has been recently published for dimethyl sulfide (Demchuk et al., 2018).

The uncertainty in a TLE estimate depends on the fit of the probit model to the incidence data. A large uncertainty in coefficients may originate either from a poor probit fit to the data, a small number of the incidence data points, or both. Usually, uncertainty on the TLE can be reduced by recruiting more data. Of the studied chemicals, chloropicrin represents a good example. An interval estimate of chloropicrin's TLE using only the Yoshida et al. (1987) study was wide (CI: 0.16–4.08, Table 2), exceeding both the lower- and upper-bound default TLE values recommended by the NAS when empirical AEGL values cannot be derived (National Academy of Sciences, 2001). The AEGL Committee derives chloropicrin's TLE from the Yoshida et al. (1987) and Yoshida et al. (1991) studies, with pooled incidence data from the whole-body exposures. The same data were used to calculate the interval TLE shown in Table 2, however, by means of categorical probit regression rather

than pooling because the dummy variable separating these studies was significant, but this resulted in an even wider TLE interval (CI: 0.27—4.52). However, adding the nose-only exposure data from Yoshida et al. (1991) to the categorical probit regression as an additional study results in a valid model and reduces uncertainty on chloropicrin's TLE by almost 50%. The new interval estimate became $n = 2.12$, 95% CI: 1.02–3.22. TLE estimates with large uncertainties should be viewed with suspicion. However, the inclusion of a relatively small additional study under the proper statistical framework may reduce uncertainty considerably.

Nevertheless, most of the chemicals in the present study garnered small uncertainties in their TLE estimates. However, several present the conundrum of having both their TLEs and associated CIs outside of default bounds of 1 and 3. These precise TLEs outside of the default 1 and 3 suggest that the single-sided 95%-CI defaults proposed based on the original 20-chemical study of ten Berge et al. (1986) may be too restrictive. Calculation of TLEs for additional chemicals that expands the number of empirical TLEs available for analysis may suggest more statistically-representative bounds on unknown TLE values, which may be less than 1 and greater than 3.

The AEGL committee often derives TLEs by simple linear regression, i.e. Equation (4). Uncertainty on these point estimates is not reported (because the simple linear regression does not propagate the uncertainty associated with incidence data into the confidence bounds on TLEs). Thirteen chemicals with TLEs derived this way were reexamined in the present study. When these point-estimate TLEs were compared with interval-estimated TLEs of the present study, 6 of these 13 point estimates were found outside of the confidence bounds. This suggests the lack of reliability in TLEs derived by means of simple linear regression or equivalent point estimates derived from multiple studies.

5. Conclusions

Chemical risk assessment often confronts the challenge of selecting a key study from multiple studies that could be used for deriving health guidance. The studies may vary in species, statistical power, and laboratory methods, and it may be elusive which study to choose. By designating only one study as a key study, information from complementary studies that could contribute to the understanding of a chemical's toxicological response is lost. However, correct joint analysis of multiple studies requires a valid meta-analytical framework. Without such a framework, results may be skewed in an indeterminate manner. The resulting inhalation guidance levels may be under-protective or needlessly over-protective. In the present work, uncertainty in temporal extrapolation of short-term inhalation exposures levels was quantified. The quantification exposed commonalities and conflicts among the published studies. Examination of probit meta-analyses performed on 28 chemicals with mortality incidence data has suggested a decision tree for incorporating multiple studies in a statistically appropriate manner for risk assessment of short-term inhalation exposures. The decision tree provides a foundation for evidence-based public health assessment of short-term inhalation exposure scenarios pertinent to emergency response, hazard detection, and preparedness planning. Future work may utilize this framework to update inhalation guidance for these and other chemicals of interest to ATSDR.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

J.H., C.R.C., and R.C.S. were supported by a fellowship from the Oak Ridge Institute for Science and Education.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Adams EM, Spencer HC, Rowe VK, Irish DD, 1950. Vapor toxicity of 1,1,1-trichloroethane (methyl chloroform) determined by experiments on laboratory animals. Arch. Ind. Hyg. Occup. Med 1, 225–236. [PubMed: 15426339]

Appelman LM, ten Berge WF, Reuzel PG, 1982. Acute inhalation toxicity study of ammonia in rats with variable exposure periods. Am. Ind. Hyg. Assoc. J 43 (9), 662–665. 10.1080/15298668291410387. [PubMed: 7148687]

Bonnet P, Francin JM, Gradiski D, Raoult G, Zissu D, 1980. Determination of the median lethal concentration of principle chlorinated aliphatic hydrocarbons in the rat. Arch. Mal. Prof 41, 317–321.

Borenstein M, Hedges LV, Higgins JPT, Rothstein HR, 2009. Introduction to Meta-Analysis, first ed. John Wiley & Sons Ltd, Chichester. 10.1002/9780470743386.

Borenstein M, Hedges LV, Higgins JPT, Rothstein HR, 2010. A basic introduction to fixed-effect and random-effects models for meta-analysis. Res. Synth. Methods 1, 97–111. 10.1002/jrsm.12. [PubMed: 26061376]

Calhoun LL, Lomax LG, Phillips JE, 1988. Aerothene TT: An acute vapor inhalation study in Fischer 344 rats. EPA/OTS Doc # 86–880000173.

Davis HV, 1970. Acute Toxicity of Oxygen Difluoride. AMRL-TR-70-102. Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, O.H. http://www.dtic.mil/dtic/tr/fulltext/u2/727525.pdf.

Demchuk E, Ball SL, Le SL, Prussia AJ, 2018. Concentration-time extrapolation of short-term inhalation exposure levels: dimethyl sulfide, a case study using a chemical-specific toxic load exponent. Inhal. Toxicol 30, 448–462. 10.1080/08958378.2018.1551444. [PubMed: 30600740]

Finney DJ, 1977. Probit Analysis, second ed. Cambridge University Press, London.

Haun CC, MacEwen JD, Vernot EH, Egan GF, 1970. Acute inhalation toxicity of monomethylhydrazine vapor. Am. Ind. Hyg. Assoc 31, 667–677. 10.1080/0002889708506313.

Kapeghian JC, Mincer HH, Jones AB, Verlangieri AJ, Waters IW, 1982. Acute inhalation toxicity of ammonia in mice. Bull. Environ. Contam. Toxicol 23 (3), 371–378. 10.1007/BF01706243.

Kirkpatrick DT, 2008. Acute inhalation toxicity study of allyl alcohol in albino rats (with 1-, 4-, and 8-hour exposure durations). Study Number WIL-14068. WIL Research Laboratories, LLC, Ashland, O.H.

Ku HH, 1966. Notes on the use of propagation of error formulas. J. Res. Natl. Bur. Stand. Sect. C Eng. Instrum 70C (4). 10.6028/jres.070c.025.

Lester D, Adams WR, 1965. The inhalation toxicity of oxygen difluoride. Am. Ind. Hyg. Assoc. J 26, 562–567. 10.1080/00028896509342774. [PubMed: 5879991]

Li WL, Pauluhn J, 2015. Mechanisms involved in the inhalation toxicity of phosgene. In: Salem H, Katz SA (Eds.), Inhalation toxicology, third ed. CRC Press, New York, N.Y., pp. 459–488. 10.1201/b16781.

MacEwen JD, Vernot EH, 1972. Toxic Hazards Research Unit Annual Technical Report: 1972. AMRL-TR-72-62. NTIS AD-755 358. Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, O.H.

McCord CP, 1932. The toxicity of allyl alcohol. J. Am. Med. Assoc 98 (26), 2269–2270. 10.1001/jama.1932.02730520011003.

National Academy of Sciences, 2001. Standing operating procedures for developing acute exposure guideline levels for hazardous chemicals. 10.17226/10122. Washington, D.C.

National Academy of Sciences, 2002. Acute Exposure Guideline Levels for Selected

Airborne Chemicals, Volume 2. National Academies Press, Washington, D.C.. 10.17226/10522.

National Academy of Sciences, 2008. Acute Exposure Guideline Levels for Selected Airborne Chemicals: Volume vol. 6. National Academies Press, Washington, D.C.. 10.17226/12018.

National Academy of Sciences, 2010. Acute Exposure Guideline Levels for Selected Airborne Chemicals, Volume 9. National Academies Press, Washington, D.C.. 10.17226/12978.

National Academy of Sciences, 2014. Acute Exposure Guideline Levels for Selected Airborne Chemicals, Volume 18. National Academies Press, Washington, D.C.. 10.17226/12978.

Silver SD, McGrath FP, 1948. A comparison of acute toxicities of ethylene imine and ammonia to mice. J. Ind. Hyg. Toxicol 30 (1), 7–9. [PubMed: 18895722]

Smyth HF, Carpenter CP, 1948. Further experience with the range finding test in the industrial toxicology laboratory. J. Ind. Hyg. Toxicol 30 (1), 63–68. [PubMed: 18895731]

ten Berge WF, Zwart A, Appelman LM, 1986. Concentration—time mortality response relationship of irritant and systemically acting vapours and gases. J. Hazard Mater 13, 301–309. 10.1016/0304-3894(86)85003-8.

Union Carbide and Carbon Corporation, 1951. Initial submission: Letter from DuPont Chem Regarding a Letter About Toxicity Studies with Allyl Alcohol. Union Carbide and Carbon Corporation, New York January 29, 1951. Submitted by DuPont, Wilmington, DE to EPA with cover letter dated October 27, 1992. EPA Document No. 88–920009857. Microfische No. OTS0571508.

U.S. EPA, 2008. Benchmark Dose Software Development and Maintenance Ten Berge Cxt Models. U.S. Environmental Protection Agency, Washington, D.C. EPA/600/C-08/009.

Verma V, Yu QJ, Connell DW, 2015. A comparison of Reduced Life Expectancy (RLE) model with Haber's Rule to describe effects of exposure time on toxicity. Environ. Pollut 204, 26–31. 10.1016/j.envpol.2015.04.008. [PubMed: 25898234]

Wasserstein RL, Lazar NA, 2016. The ASA's statement on p-values: context, process, and purpose. Am. Statistician 70, 129–133. 10.1080/00031305.2016.1154108.

Yoshida M, Ikeda T, Iwasaki M, Tsuda S, Shirasu Y, 1987. Acute inhalation toxicity of chloropicrin vapor in rats. Nihon Noyaku Gakkaishi (J. Pestic. Sci. Soc. Jpn.) 12, 237–244. 10.1584/jpestics.12.237.

Yoshida M, Murao N, Tsuda S, Shirasu Y, 1991. Effects of mode of exposure on acute inhalation toxicity of chloropicrin vapor in rats. Nihon Noyaku Gakkaishi (J. Pestic. Sci. Soc. Jpn.) 16, 63–69. 10.1584/jpestics.16.63.

Zwart A, Art J, Klokman-Houweling J, Schoen E, 1990. Determination of concentration-time-mortality relationships to replace LC_{50} values. Inhal. Toxicol 2, 105–117.

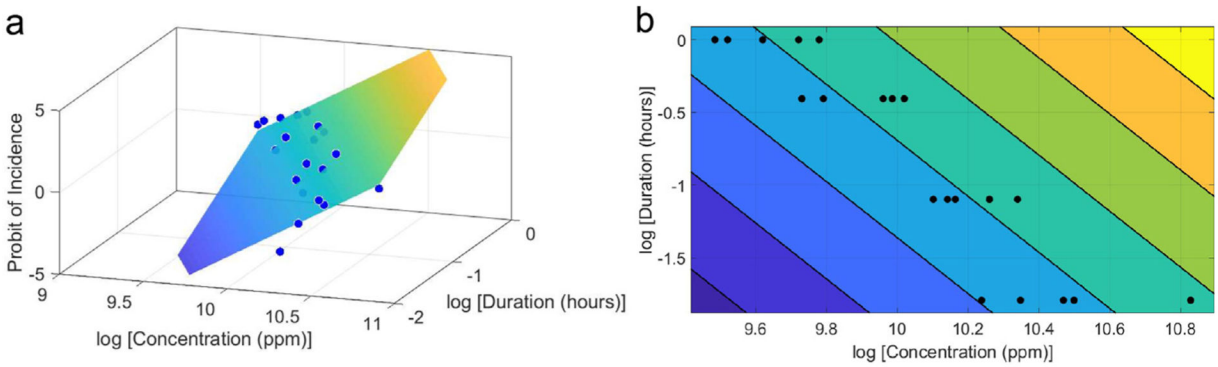

Fig. 1.

Visualization of the relationship between concentration (C) and exposure duration (t) on incidence expressed in probability units (probits) of a rat inhalation mortality study with ammonia (Appelman et al., 1982). Panel (a) relates C and t to the mortality probit. Panel (b) is a 2-dimensional projection of 1a on the probit plane, illustrating that for different incidence probabilities, log-linear relationships between C and t have identical slopes.

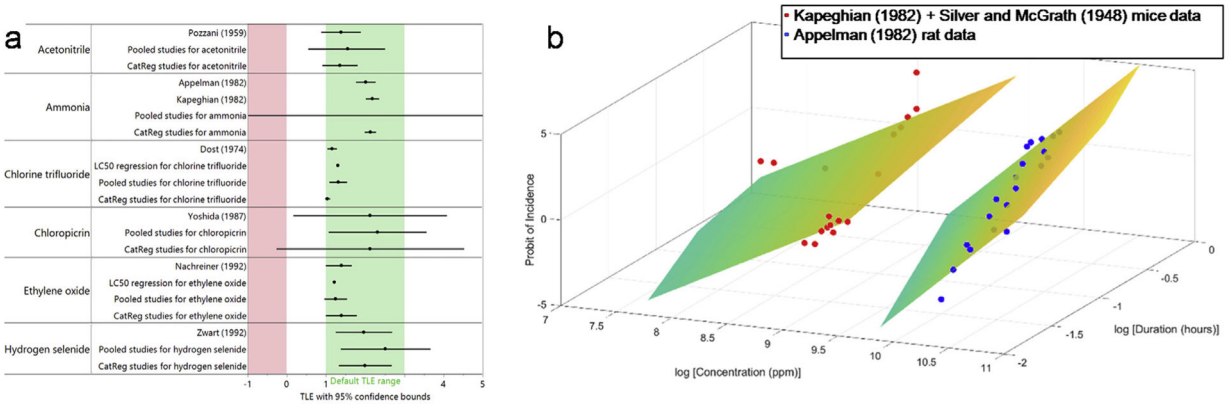

Fig. 2.

Results of categorical regression probit analysis for chemicals with parallel-plane studies. Panel (a) represents a collection of chemicals' TLE interval estimates arranged as a forest plot. The estimates are baed on incidence data from individual studies, pooled studies, and the categorical probit regression. Panel (b) illustrates a typical parallel relationship between probit planes of a chemical. It represents the second chemical on the list and includes the ammonia data from the Appelman et al. (1982) study on rats and the Kapeghian et al. (1982) and Silver and McGrath (1948) studies on mice, in which mice (red dots) appears more sensitive. The planes are parallel within the confidence bounds of the slope coefficient.

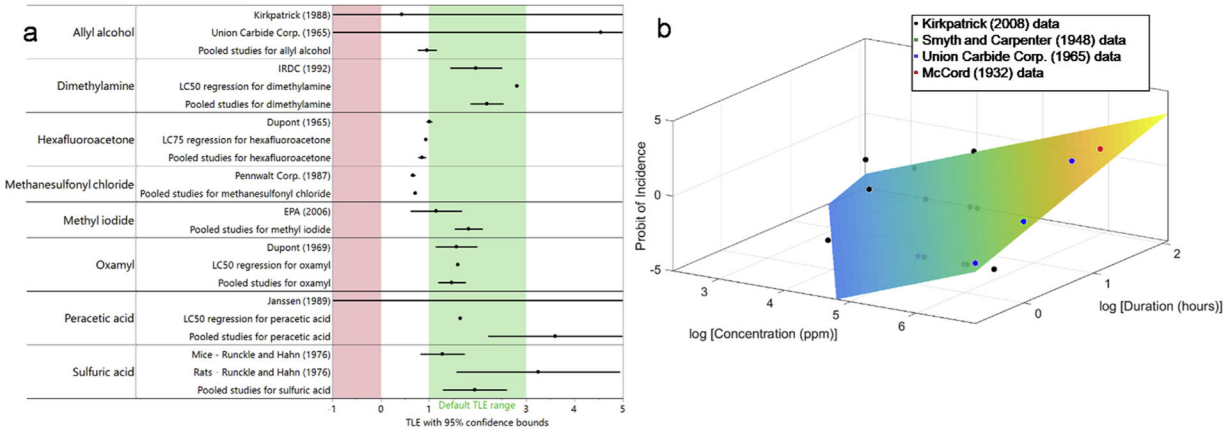

Fig. 3.

Results of probit analysis for chemicals with pooled-study data. Panel (a) represents TLE estimates for each chemical arranged as a forest plot. The TLE estimates were calculated based on individual study data and when all available data were pooled together. Panel (b) illustrates how the pooled data from four studies for allyl alcohol, Kirkpatrick (2008); McCord (1932); Smyth and Carpenter (1948); and Union Carbide and Carbon Corporation (1951) make a single probit plane.

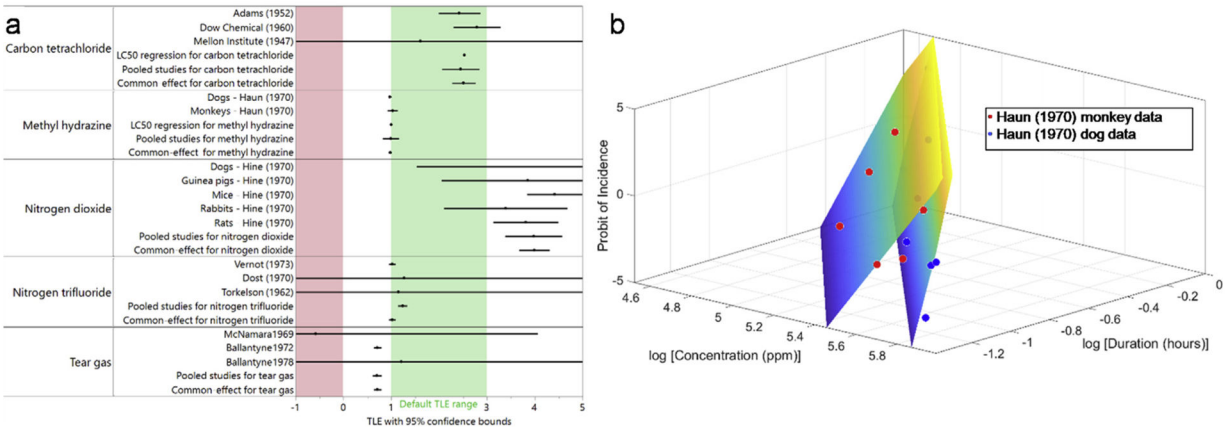

Fig. 4.

Results of probit meta-analysis for chemicals studied using the common-effect model. Panel (a) represents TLE forest plots for each chemical calculated based on individual studies, pooled studies, and as the common-effect model summary effect. Panel (b) illustrates how the data from two studies for methyl hydrazine may not be parallel (Haun et al., 1970), but are appropriate for the common-effect modeling.

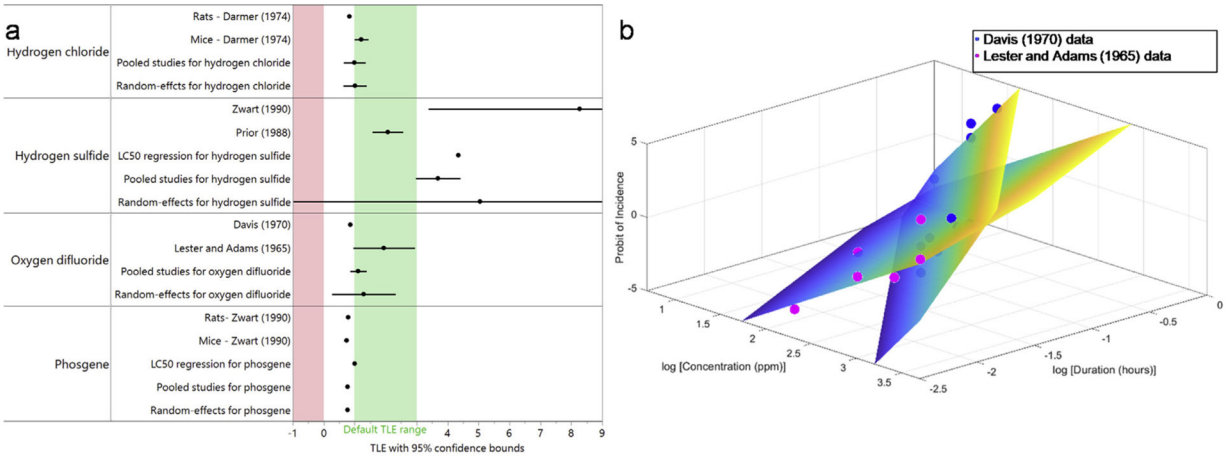

Fig. 5.

Results of probit meta-analysis for chemicals examined using the random-effects model. Panel (a) represents TLE forest plots for each chemical calculated based on individual studies, pooled studies, and as the random-effects model summary effect. Panel (b) illustrates how the data from two studies for oxygen difluoride (Davis, 1970; Lester and Adams, 1965) may not be parallel but still can be combined as a synthetic summary effect using the random-effects model.

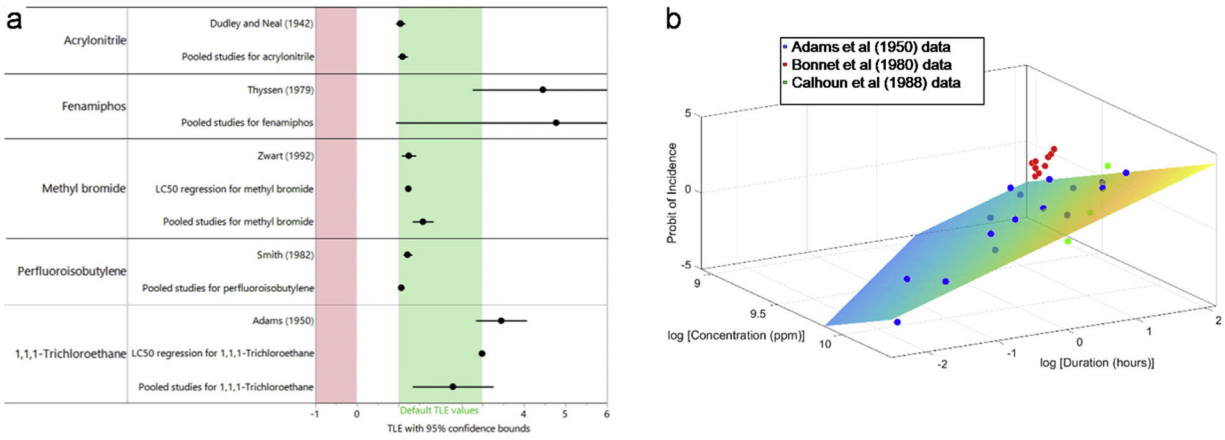

Fig. 6.

Results of probit analysis for chemicals, whose sparse database allows only a single-study TLE estimation. Panel (a) is a forest plot of the TLEs for each chemical calculated based on their main study, and when the main study data were pooled with single-duration studies. Panel (b) illustrates how data from the single-duration studies, Bonnet et al. (1980) and Calhoun et al. (1988), may not be parallel, and lay off the probit plane of the main study, Adams et al. (1950), for 1,1,1-trichloroethane (incidence data from the other single-duration study were unavailable).

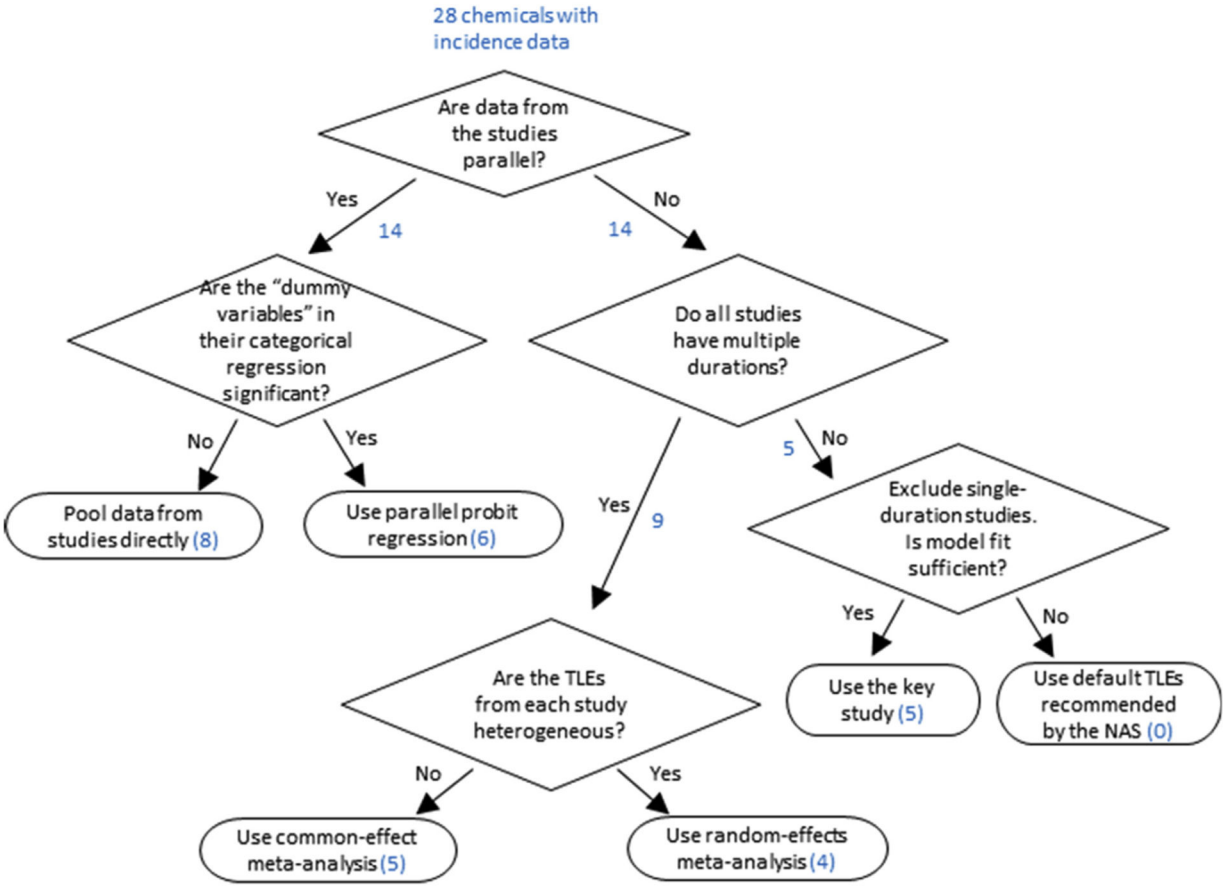

Fig. 7.

Decision tree with logic for deriving a TLE for a chemical with multiple studies. The resulting decisions for the 28 chemicals analyzed in this study are shown in blue text.

Table 1

Acute Exposure Guideline Levels (AEGLs) for ammonia. The concentrations of ammonia in the air are in parts per million (ppm).

Severity Tier	10 min	30 min	1 h	4 h	8 h
AEGL-1 (discomfort)	30	30	30	30	30
AEGL-2 (disabling)	220	220	160	110	110
AEGL-3 (life-threatening)	2700	1600	1100	550	390

Table 2

Meta-analytical TLEs for chemicals with homogeneous multi-study data (true TLEs).[a]

Category	Chemical	Species	Dataset[b]	TLE (95% CIs)
Parallel	Acetonitrile	Rats	Pozzani (1959)	1.38 (0.87, 1.88)
			AEGL probit of pooled studies	1.55 (0.54, 2.5)
			Pozzani (1959)	**1.35 (0.90, 1.80)**
			Union Carbide Corp. (1965)	
			Monsanto (1986)	
	Ammonia	Rats	Appleman (1982)	2.01 (1.76, 2.26)
		Mice	Kapeghian (1982)	2.18 (2.01, 2.35)
		Mice + Rats	AEGL citation of ten Berge TLE	2.0 (1.6, 2.4)
			Probit of pooled studies	−3.33 (−17.64, 10.99)
		Rats	**Appleman (1982)**	**2.13 (1.98, 2.27)**
		Mice	Kapeghian (1982) + Silver and McGath (1948)	
	Chlorine trifluoride	Rats	Dost (1974)	1.15 (1.03, 1.27)
			AEGL LC_{50} regression	1.3
			Probit of pooled studies	1.31 (1.08, 1.53)
			Dost (1974)	**1.03 (0.97, 1.10)**
			Horn and Weir (1955) + MacEwen and Vernot (1970)	
	Chloropicrin	Rats	Yoshida (1987)	2.12 (0.16, 4.08)
			AEGL probit of pooled studies	2.31 (1.06, 3.56)
			Yoshida (1987)	**2.12 (−0.27, 4.52)**
			Yoshida (1991)	
	Ethylene oxide	Rats	Nachreiner (1992)	1.39 (0.99, 1.65)
			AEGL LC_{50} regression	1.21
			Probit of pooled studies	1.24 (0.95, 1.53)
			Nachreiner (1992)	**1.39 (0.99 1.78)**
			Jacobson (1956)	
	Hydrogen selenide	Rats	Zwart (1992)	1.96 (1.24, 2.68)
			AEGL probit of pooled studies	2.51 (1.37, 3.66)
			Zwart (1992)	**1.99 (1.32, 2.67)**

Category	Chemical	Species	Dataset[b]	TLE (95% CIs)
			Zwart and Arts (1989)	
Pooled	Allyl alcohol	Rats	Kirkpatrick (2008)	0.43 (−5e + 4, 5e + 4)
			Union Carbide Corp. (1965)	4.54 (−1e + 6, 1e + 6)
			AEGL probit of pooled studies	**0.95 (0.76, 1.15)**
	Dimethylamine	Rats	IRDC (1992)	1.96 (1.43, 2.50)
			AEGL LC_{50} regression	2.81
			Probit of pooled studies	**2.19 (1.85, 2.53)**
	Hexafluoroacetone	Rats	Dupont (1965)	1.00 (0.94, 1.06)
			AEGL LC_{50} regression	0.93
			Probit of pooled studies	**0.85 (0.77, 0.93)**
	Methanesulfonyl chloride	Rats	Pennwalt Corp. (1987)	0.66 (0.61, 0.71)
			AEGL probit of pooled studies	0.71 (0.30, 1.11)
			Probit of pooled studies (with 45 min data)	**0.71 (0.67, 0.74)**
	Methyl iodide	Rats	EPA (2006)	1.14 (0.61, 1.67)
			AEGL probit of pooled studies	**1.81 (1.52, 2.10)**
	Oxamyl	Rats	Dupont (1969)	1.56 (1.13, 1.99)
			AEGL LC_{50} regression	1.59
			Probit of pooled studies	**1.46 (1.18, 1.75)**
	Peracetic acid	Rats	Janssen (1989)	14.44 (−94.0, 122.8)
			AEGL LC_{50} regression	1.6
			Probit of pooled studies	**3.60 (2.21, 4.99)**
	Sulfuric acid	Mice	Runckle and Hahn (1976) (AEGL used probit of mice only)	1.27 (0.81, 1.73)
		Rats	Runckle and Hahn (1976)	3.25 (1.56, 4.94)
		Mice + Rats	**Pooled probit of studies**	**1.94 (1.28, 2.60)**
Common-effect meta-analysis	Carbon tetrachloride	Rats	Adams (1952)	2.42 (1.99, 2.86)
			Dow Chemical (1960)	2.79 (2.30, 3.28)
			Mellon Institute (1947)	1.61 (−2e + 6, 2e + 6)
			AEGL LC_{50} regression	2.53
			Probit of pooled studies	2.43 (2.09, 2.77)

Category	Chemical	Species	Dataset[b]	TLE (95% CIs)
			Common-effect meta-analysis (p-value for Q = 0.45)	**2.51 (2.27, 2.76)**
	Methyl hydrazine	Dogs	Haun (1970)	0.97 (0.94, 1.01)
		Monkeys	Haun (1970)	1.03 (0.92, 1.13)
		Dogs + Monkeys	AEGL LC_{50} regression	0.99 (dogs) 0.97 (monkeys)
			Probit of pooled studies	0.99 (0.82, 1.16)
			Common-effect meta-analysis (p-value for Q = 0.33)	**0.98 (0.94, 1.01)**
	Nitrogen dioxide	Dogs	Hine (1970)	5.25 (1.53, 8.89)
		Guinea pigs	Hine (1970)	3.85 (2.05, 5.66)
		Mice	Hine (1970)	4.41 (3.84, 4.99)
		Rabbits	Hine (1970)	3.39 (2.10, 4.68)
		Rats	Hine (1970)	3.81 (3.13, 4.49)
		Dogs + Guinea pigs + Mice + Rabbits + Rats	AEGL citation of ten Berge TLE	3.5 (2.7, 4.3)
			Probit of pooled studies	3.98 (3.38, 4.57)
			Common-effect meta-analysis (p-value for Q = 0.39)	**3.99 (3.67, 4.31)**
	Nitrogen trifluoride	Rats	Vernot (1973)	1.02 (0.95, 1.09)
			Dost (1970)	1.27 (−3e + 4, 3e + 4)
			Torkelson (1962)	1.15 (−6e + 4, 6e + 4)
			AEGL probit of pooled studies	1.23 (1.03, 1.43)
			Probit of pooled studies (with 5000 ppm data)	1.24 (1.06, 1.42)
			Common-effect meta-analysis (p-value for Q = 1.00)	**1.02 (0.95, 1.09)**
	Tear gas (CS, 2-Chlorobenzylidenemalononitrile)	Rats	McNamara (1969)	−0.58 (−5.22, 4.06)
			Ballantyne (1972)	0.71 (0.63, 0.79)
			Ballantyne (1978)	1.21 (−9.83, 12.25)
			AEGL probit of pooled studies	0.70 (0.54, 0.87)
			Fixed effects meta-analysis (p-value for Q = 0.86)	**0.71 (0.63, 0.79)**

[a]The chemicals are categorized by the statistically most appropriate method of analysis, identified by the bold rows.

[b]Shows dataset labels, not the entire inventory of studies. Refer to supporting information for complete listing of studies. “+” denotes pooling.

Table 3

Synthetic and designated study TLEs estimated from heterogeneous data.[a]

Method of analysis	Chemical	Species	Dataset[b]	TLE (95% CIs)
Random- effects meta-analysis	Hydrogen chloride	Rats	Darmer (1974)	0.83 (0.77, 0.89)
		Mice	Darmer (1974)	1.21 (0.99, 1.44)
		Mice + Rats	AEGL citation of ten Berge TLE	1.0 (0.7, 1.3)
			Random-effects meta-analysis (p-value for Q = 0.001)	**1.01 (0.63, 1.38)**
	Hydrogen sulfide	Rats	Zwart (1990)	8.27 (3.38, 13.17)
			Prior (1988)	2.07 (1.57, 2.56)
			AEGL LC_{50} regression	4.35
			Probit of pooled studies	3.69 (2.97, 4.41)
			Random-effects meta-analysis (p-value for Q = 8e-7)	**5.05 (−1.03, 11.12)**
	Oxygen difluoride	Rats	Davis (1970)	0.86 (0.81 0.92)
			Lester and Adams (1965)	1.94 (0.95, 2.94)
			AEGL citation of ten Berge TLE	1.1 (1.0, 1.2)
			AEGL probit of pooled studies	1.11 (0.82, 1.40)
			Random-effects meta-analysis (p-value for Q = 0.03)	**1.29 (0.26, 2.32)**
	Phosgene	Rats	Zwart (1990)	0.79 (0.74, 0.84)
		Mice	Zwart (1990)	0.74 (0.65, 0.79)
		Rats + Mice	AEGL LC_{50} regression	1
			Probit of pooled studies	0.77 (0.70, 0.83)
			Random-effects meta-analysis (p-value for Q = 0.12)	**0.77 (0.72, 0.82)**
Designated study analysis	Acrylonitrile	Rats	**Dudley and Neal (1942)**	**1.04 (0.93, 1.15)**
			AEGL citation of ten Berge TLE	1.1 (1.0, 1.2)
			AEGL probit of pooled studies	1.09 (0.98, 1.21)
	Fenamiphos	Rats	Thyssen (1979)	**4.46 (2.77, 6.14)**
			AEGL probit of pooled studies	4.78 (1.42, 8.15)
			Probit of pooled studies	4.78 (0.93, 8.63)
	Methyl bromide	Rats	**Zwart (1992)**	**1.24 (1.07, 1.41)**
			AEGL LC_{50} regression	1.2
			Probit of pooled studies	1.58 (1.33, 1.83)

Method of analysis	Chemical	Species	Dataset[b]	TLE (95% CIs)
	Perfluoroisobutylene	Rats	**Smith (1982)**	**1.21 (1.11, 1.32)**
			AEGL LC_{50} regression	1.04
			Probit of pooled studies	1.06 (0.99, 1.13)
	1,1,1-Trichloroethane	Rats	Adams (1950)	**3.46 (2.85, 4.07)**
			AEGL LC_{50} regression	3
			Probit of pooled studies	2.23 (1.23, 3.23)

[a]The summary effect TLE is bolded for each chemical's random-effects meta-analysis. For chemicals requiring a designated study, the study and its resulting TLE are also bolded.

[b]Shows dataset labels, not the entire inventory of studies. Refer to supporting information for complete listing of studies.

Table 4

Comparison of the published and recalculated AEGLs for gaseous hydrogen chloride. The concentrations are in parts per million (ppm).

Time	10 min	30 min	1 h	4 h	8 h
AEGL-3	620	210	100	26	26
Recalculated	381	172	100	12	12[a]
Difference	40%	20%	0%	2-fold	2-fold

[a]The AEGL TSD ackowledges uncertainty in extrapolation from 1-h to 8-h duration, however, adopts the least health protective approach extrapolating from a shorter 4-h duration to longer 8-h duration, essentially, using an infinitely large TLE. With evidence-based uncertainty in mind, an 8-h exposure level would be 4 ppm, i.e. the difference would increase 7-fold.

Table 5

Comparison of short-term inhalation exposure levels for ammonia derived using three meta-analytical procedures.

Ammonia inhalation guidance (concentrations in ppm)		10 min	30 min	1 h	4 h	8 h
AEGL Committee procedure	1-h LC_{01} from mice studies			3300		
	Divide by total UF of 3			1100		
	Extrapolate to other durations using TLE = 2 from pooled mice and rat studies (AEGL-3 levels)	2700	1600	1100	550	390
Pooled probit regression	LC_{01} for rat & mice studies in a pooled regression (95% CI)	2.5 (0.00009–54)	3.6 (0.0003–57)	4.3 (0.0007–61)	6.4 (0.003–78)	7.9 (0.005–94)
	Divide by total UF of 3 (recalculated AEGL-3 levels)	0.83	1.2	1.4	2.3	2.6
Categorical probit regression approach	LC_{01} for mice studies in categorical regression (95% CI)	6100 (5400–6700)	3700 (3200–4000)	2700 (2300–2900)	1400 (1200–1600)	1000 (850–1200)
	Divide by total UF of 3 (recalculated AEGL-3 levels)	2000	1200	890	470	340