Regulatory Toxicology and Pharmacology 145 (2023) 105505

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

journal homepage: www.elsevier.com/locate/yrtph

Nitrosamine acceptable intakes should consider variation in molecular weight: The implication of stoichiometric DNA damage

Jonathan Fine [a], Leonardo Allain [a], Joerg Schlingemann [b,*], David J. Ponting [c], Robert Thomas [c], George E. Johnson [d]

[a] Merck & Co., Inc., Rahway, NJ, USA
[b] Merck KGaA, Frankfurter Str. 250, 64293, Darmstadt, Germany
[c] Lhasa Limited, Granary Wharf House, 2 Canal Wharf, Leeds, UK
[d] Institute of Life Science, Swansea University Medical School, Swansea, UK

ARTICLE INFO

Handling Editor: Dr. Martin Van den berg

Keywords:
Nitrosamines
N-Nitrosamines
NDSRIs
Risk assessment
Acceptable intake
Read-across
Carcinogenicity
Percentile
TD_{50}
Toxicology
18 ng AI default
ICH M7 cohort of concern

ABSTRACT

N-nitrosamines (NAs) are a class of compounds of which many, especially of the small dialkyl type, are indirect acting DNA alkylating mutagens. Their presence in pharmaceuticals is subject to very strict acceptable daily intake (AI) limits, which are traditionally expressed on a mass basis. Here we demonstrate that AIs that are not experimentally derived for a specific compound, but via statistical extrapolation or read across to a suitable analog, should be expressed on a molar scale or corrected for the target substance's molecular weight. This would account for the mechanistic aspect that each nitroso group can, at maximum, account for a single DNA mutation and the number of molecules per mass unit is proportional to the molecular weight (MW). In this regard we have re-calculated the EMA 18 ng/day regulatory default AI for unknown nitrosamines on a molar scale and propose a revised default AI of 163 pmol/day. In addition, we provide MW-corrected AIs for those nitrosamine drug substance related impurities (NDSRIs) for which EMA has pre-assigned AIs by read-across. Regulatory acceptance of this fundamental scientific tenet would allow one to derive nitrosamine limits for NDSRIs that both meet the health-protection goals and are technically feasible.

1. Introduction

NAs are a class of chemical compounds with the common structure depicted in Fig. 1. Their presence in certain foods and tobacco products is well-known (Chain, E. P. o. C. i. t. F., 2023), but since 2018 there have also been reports and recalls due to their presence in pharmaceuticals (Nudelman et al., 2023). NAs can be formed when vulnerable amine moieties (especially secondary amines) are present under conditions that promote the formation of nitrosating agents such as the nitrous acidium ion, the nitrosonium ion or dinitrogen trioxide (depending on conditions), especially from inorganic nitrite. The detected NA compounds were initially those formed from small dialkyl amines due to side reactions in or impurities from the synthetic process (Burns et al., 2020; Schlingemann et al., 2022a), but more recently an increasing number of larger nitrosamines derived from amine-containing API or API impurities (NDSRIs) has been reported. This is due in no small part to the presence of nitrite as an impurity in many common excipients (Boetzel et al., 2022), and due to the widespread presence of secondary amines in APIs and their impurities, and the number of drug products that may form a nitrosamine under storage conditions is significant (Schlingemann et al., 2022b).

Some nitrosamines are exceptionally potent genotoxic carcinogens, and hence the threshold of toxicological concern (TTC) concept from ICH guideline M7 that limits exposure to mutagenic compounds to a maximum acceptable intake (AI) of 1500 ng/day cannot be applied. Snodin estimated that this applies to ca. 50% of all nitrosamines (Snodin, 2023). Instead, compound-specific exposure limits must be established, either by read-across to a suitable surrogate compound or a dedicated carcinogenicity study (ICH, 2017). For novel structures, Regulatory Authorities have put in place a default AI of 18 ng/day based on the 5th percentile of N-nitroso compound TD_{50} values from the Lhasa carcinogenicity database (EMA, 2022). TD_{50} is the lifetime dose that

* Corresponding author.
E-mail address: joerg.schlingemann@merckgroup.com (J. Schlingemann).

https://doi.org/10.1016/j.yrtph.2023.105505
Received 10 July 2023; Received in revised form 15 September 2023; Accepted 2 October 2023
Available online 5 October 2023

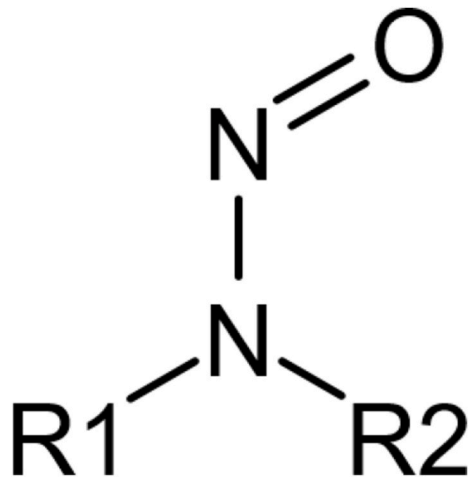


Fig. 1. Common structure of N-nitrosamines. R1 and R2 must be carbons with either single or aromatic bonds only.

induced tumors in 50% of the treated animals, and the AI in humans is calculated as AI = $TD_{50} \times 2 \times 50$ kg/100,000, resulting in an additional theoretical lifetime risk of 0.001% to the natural cumulative lifetime background risk of approximately 40% of cancer in the population (American Cancer Society, 2022). Setting the default AI based on the TD_{50} value at the 5th percentile shall ensure that 95% of NAs with carcinogenicity data have a theoretical compound specific AI higher than the default; by extrapolation, it is assumed that the same distribution applies to the larger chemical space of nitrosamines without known carcinogenic potency. This assumption may be conservative, since the chemical space of unknown nitrosamines consists mainly of larger, more complex molecules – expected to be of lower potency. TD_{50} values are commonly provided on a mass scale [mg/kg/day] for useful analogy with the LD_{50} (Peto et al., 1984), and because, for reasons of simplicity, the compound dosing in the lab is accomplished via weight measurements. Consequently, the derived compound AIs and ultimately the default AI obtained by statistical extrapolation are expressed on a mass scale as well. There is, however, a conceptual flaw in this approach, since the mechanism for DNA damage from nitrosamines requires a stoichiometric conversion to the actual alkylating agent, a diazonium ion. It is worth noting that the same applies to other assays, in which a stoichiometric reaction is primarily responsible for the effect, for example the local lymph node assay (LLNA) for skin sensitization, which is experimentally dosed, and often reported, as a weight/volume unit but more usefully compared by molarity (Lidén et al., 2016; Natsch et al., 2014).

The carcinogenic potential of nitrosamines stems from the so-called metabolic activation by enzymes of the cytochrome P450 complex that creates DNA-alkylating diazonium ions (Hecht, 1998; Li and Hecht, 2022a, 2022b). Each nitrosamine functional group can at most release a single diazonium ion, which can at most result in a single DNA alkylation that can become manifest, at most, in a single DNA mutation. The number of nitroso groups per mass unit and hence the maximum possible number of DNA alkylations per mass unit of nitrosamine is proportional to the compound's molecular weight. Consequently, it can be argued that the default AI should be calculated and expressed on a molar basis, which allows the conversion to a mass scale for specific compounds by means of their molecular weight. Likewise, compound-specific AIs derived by read-across must be scaled to the target compound's molecular weight, as previously suggested (Nudelman et al., 2023; Snodin, 2023). However, both approaches are currently not regulatory practice. In this work we provide a revised default AI that has been calculated on a molar basis as well as MW-corrected AIs for the NDSRIs in the EMA Q&A document (EMA, 2022) to facilitate further discussion between regulatory bodies and pharmaceutical manufacturers. This translation is particularly significant in the context of the current discussion, since there is a discrepancy in the chemical space and molecular weight (Oliveira et al., 2022) between the small molecular weight nitrosamines, which were historically of most interest for investigation into their carcinogenic mechanisms, and structure-activity relationships – as reported extensively by Lijinsky and other investigators (Lijinsky, 1987), and the diverse NDSRIs now being detected. NDSRIs, being derived from drug molecules, have molecular weights many times the reference carcinogens NDEA and NDMA. Additionally, the trend for the average MW of pharmaceuticals is progressively increasing, even since Lipinski codified the 'druggable space' (Agarwal et al., 2022; Lipinski et al., 2001), meaning that many NDSRIs may have molecular weights far above 500 g/mol.

2. Methods

The Carcinogenic Potency Database (CPDB), commonly referred to as the Gold Carcinogenicity Database after its long-term curator Lois Gold from the University of California, Berkeley, is a publicly available retired database that includes the results of 6540 chronic, long-term animal cancer tests on 1547 chemicals (Gold et al., 1991). The data were extracted from the general literature through 2001 and from the National Cancer Institute/National Toxicology Program through 2004, providing standardized information for researchers, regulators, and other stakeholders. The CPDB contains TD_{50} data for 117 N-nitroso compounds (available e.g., in the supplementary material of Thomas et al. but is no longer updated with additional data. The project has been taken on by Lhasa Limited as the Lhasa Carcinogenicity Database (LCDB) (Lhasa Limited, 2023), and data added as well as Lhasa TD_{50}s which have been recalculated with a requirement for at least two dose groups and a re-implemented TD_{50} algorithm. There are 46 Lhasa TD_{50} for N-nitroso compounds (compared with 117 Gold TD_{50}), as stricter criteria were applied to the data sets compared with the original CPDB study quality inclusion criteria (Gold et al., 2005; Thomas et al., 2021), especially the exclusion of data sets with only a single dose group. The Gold TD_{50}s are provided in the supplementary material of a previous publication (Thomas et al., 2022). As further studies affecting the Lhasa TD_{50}s may be added in the future, CAS numbers identifying the 46 compounds have been added as supplementary information to ensure that up-to-date information can be received through LCDB, which is freely available at https://carcdb.lhasalimited.org.

The EMA recommendation of 18 ng/day is derived from the 5th percentile of nitroso compound TD_{50} values from the LCDB, for a 50 kg patient to have a 1 in 100,000 chance of increased cancer risk (EMA, 2022). Although the key method details for this calculation of the 5th percentile were not discussed in regulatory guidelines, several potential derivations of that value exist. Here we present one: first the distribution of TD_{50} values (expressed in mg/kg/day) is assumed to follow a log-normal distribution. Then, the mean and standard deviation of this distribution are fitted using the likelihood method presented in Thomas et al. (2021). The 5th percentile of the generated distribution serves as the TD_{50} reference value for the calculation of the default AI using equation (1).

$$AI = \frac{p5\left\{TD_{50}^{\mathrm{LCDB}}\right\}}{Risk} * BodyWeight * 2 \quad (1)$$

It should be noted that the precise value of the 5th percentile calculated using non-parametric methods (i.e., those other than the method mentioned above) is extremely sensitive to the precise values of the compounds bracketing it. These are, incidentally, small even amongst the reference set of small molecules, being NDEA, nitrosomethyl-2-oxopropylamine and nitrosomethylphenylethylamine (the 8 ng/day summary TD_{50} for which has recently been challenged (Woolley and Cross, 2023) - something that would in itself increase the 18 ng/day by moving this compound upwards in terms of relative potencies).

The methodology described above can be repeated for distributions of TD_{50} values measured in mmol molecule/kg/day and mmol nitrosamine/kg/day as long as the distributions of these TD_{50} values do not significantly differ from a log-normal distribution. The molar values are calculated using the formula weight of the free form of the nitrosamine, ignoring the mass of any counter ions. The difference between mmol molecule and mmol nitrosamine is that the TD_{50} of the latter value is scaled by the number of nitrosamine functional groups present in the molecule. The rationale for scaling the TD_{50} value is that the AI value should be calculated for each nitrosamine group because the mechanism for toxicity consumes this functional group. Note that this argument assumes that the potential contribution to carcinogenic potency of each nitrosamine group in a molecule is equivalent regardless of its local environment and incorporation of the impact for these effects is out of scope for this work. A good example of this is 2,6-dimethyl-N,N′-dinitrosopiperazine, which contains a hindered nitrosamine expected to be of low potency and an unhindered one expected to be of high potency; the potency of the molecule as a whole could conservatively be considered to be double that of the more potent site although in practical terms the contribution of the hindered site to activity may be minimal.

RDkit was used to calculate the number the nitrosamine groups in a molecule using the SMARTs structure "O = [N; X2]-N(-[#6])(-[#6,#7,#8])". This definition of a nitrosamine group allows for a single R-group to be Oxygen or Nitrogen, otherwise compounds with CAS numbers 135-20-6, 3276-41-3, 40548-68-3, and 63885-23-4 in the Gold dataset would not contain any nitrosamine functional groups. Note that this feature definition is wider than the dialkyl N-nitrosamines discussed in the introduction, covering also nitrosamides, nitrosoureas, nitrosocarbamates, nitrosoguanidines and nitrosated heterosubstituted amines to be included in the analysis; this is because the same (or similar) definition – of all N-nitroso compounds - was used in the derivation of the 18 ng/day regulatory limit. The relative potencies of these two subsets of N-nitroso compounds have been discussed (Ponting and Foster, 2023; Thomas et al., 2021, 2022).

Statistical data analysis was performed with the Python programming language (Van Rossum and Drake Jr, 2009) version 3.9.1 in combination with the libraries pandas (McKinney, 2010) version 1.5.3 (McKinney, 2010), NumPy (Harris, 2020) version 1.24.2 and SciPy (Virtanen et al., 2020). Data visualization was done with the R language for statistical computing (R_Core_Team, 2021) version 4.2.1 using the packages ggplot2 version 3.4.1 (Wickham, 2016), dplyr version 1.1.0 (Wickham et al., 2023a) and tidyr version 1.3.0 (Wickham et al., 2023b). The code is provided as online supplementary material along with the "renv.lock" and "requirements.txt" files which can be used to recreate the Python and R environments exactly.

3. Results

In order for a log-normal fit to the TD_{50} values to be valid, it is necessary that the distribution of these values does not significantly differ from a log-normal distribution. Fig. 2 shows the TD_{50} distributions of the Gold and Lhasa datasets measured in mass, molecular-molar, and normal units (equivalent to the number of NNO groups in a molecule). Only di(N-nitroso)-perhydropyrimidine, N,N-dinitrosohomopiperazine, and dinitrosocaffeidine have more than one nitrosamine.

A null-hypothesis that the distributions do not significantly differ from a log-normal distribution was tested using a Shapiro-Wilks test (p-values shown in Fig. 2). All p-values are above a = 0.05, therefore the null-hypothesis cannot be rejected. The re-calculated daily intake values based on the 5th, 33rd and 50th percentile of the log-normal distribution of Lhasa TD_{50} values are shown in Table 1.

Table 1
Calculated Daily Intake using various percentiles of the fitted TD_{50} distributions with the assumption of a 50 kg *patient and doubling of the TD_{50} value and a risk of 1 in 100,000 developing tumors as a result of nitrosamine exposure. Values are given for the 50% confidence intervals. The molar equivalent of the EMA default AI, which is calculated from the* 5th *percentile of the log-normal distribution of Lhasa TD_{50} values, is* 0.163 nmol/*day (163 pmol/day). The molar equivalent of the EMA t-AI, which is calculated from the* 33rd *percentile of the log-normal distribution of Lhasa TD_{50} values, is* 1.224 nmol/*day (1224 pmol/day).*

Database	Percentile of fitted distribution	Calculated Daily Intake (ng/day) with 50% CI	Calculated Daily Intake (nmol/day of molecule) with 50% CI	Calculated Daily Intake (n equ./day of NNO) with 50% CI
LCDB	5	**25.4 (17.8–36.3)**	**0.163 (0.114–0.233)**	0.147 (0.102–0.212)
LCDB	33	186.0 (149.9–230.7)	1.224 (0.984–1.523)	1.152 (0.922–1.439)
LCDB	50	384.6 (326.4–453.3)	2.556 (2.164–3.019)	2.443 (2.061–2.896)

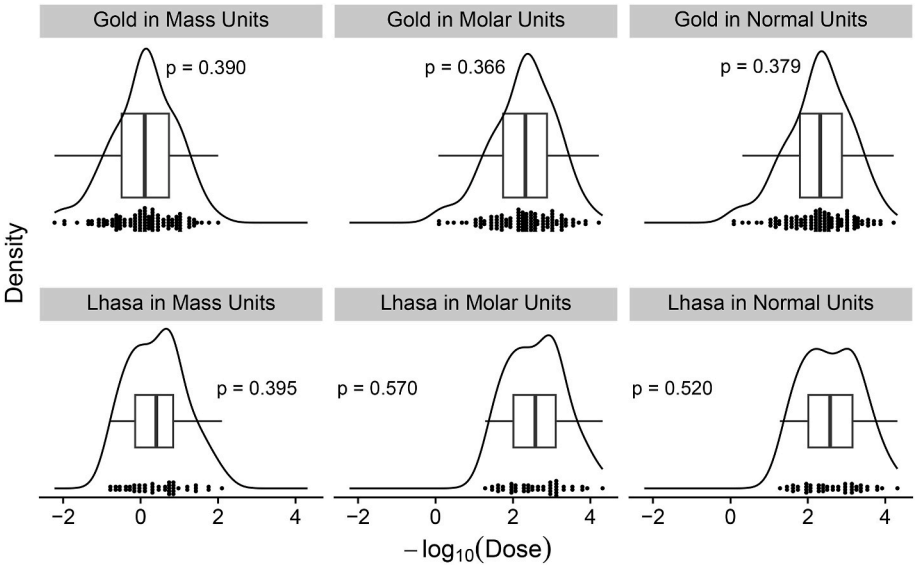


Fig. 2. log-normal distributions for LCDB are shown with mass units (left), molar units (middle), and molar units scaled to the number of nitrosamine groups (right). The normality of each distribution was calculated using a Shapiro-Wilks test where the null hypothesis states that the distributions do not significantly deviate from the normal distributions.

The calculated mass-based AIs for the 5th and 33rd percentile of TD_{50} values from LCDB ("default AI" and "temporary AI") differ slightly from the respective values stipulated by EMA (25.4 ng/day vs 18 ng/day and 186 ng/day vs 178 ng/day). This can be attributed to the fact that different mathematical approaches can be taken to estimate these percentiles (Schoonjans et al., 2011). As details about EMA's approach have not been disclosed, we used the parametric likelihood model (Fig. 3), which is expected to be superior to counting and interpolation (if the percentile has no corresponding data point), especially for small sample sizes (Parrish, 1990), and has been shown to provide robust estimates of the N-nitrosamine TD_{50} distribution (Thomas et al., 2021).

4. Discussion

Nitrosamines are mutagenic when they act as DNA-alkylating agents. The actual alkylating species is not the nitrosamine itself but a diazonium ion that is created upon metabolic activation by cytochrome P450 enzymes. Different DNA-adducts can be formed from N-nitrosamines depending on their respective chemical structure and that of the alkylating diazonium ion created upon metabolic activation. Dimethyl- and Diethyl-groups are considered the more important in terms of mutagenic/carcinogenic potency compared to longer chain and cyclic compounds (EMA/CHMP, 2020). Small adducts are more likely to be misrecognized during replication, leading to O6-methyl-G and O6-ethyl-G lesions that may cause GC > AT mutations and O4-alkyl-T lesions that may cause TA > CG mutations. These small alkyl-adducts at O6-G and O4-T are repaired by human O6-alkylguanine DNA alkyltransferase (MGMT) through direct transfer of the alkyl group to Cys-145 in the protein, leading to irreversible inactivation of MGMT. In contrast, the efficiency of MGMT-mediated repair of O6-alkyl-dG lesions decreases with increasing size of the alkyl groups (Du et al., 2019), and larger and more complex O6-alkyl-dG lesions strongly block replicative polymerases, which often leads to cell death. The latter adducts can be overcome by translesion synthesis (TLS). Depending on the type of DNA lesion and error-prone TLS-polymerase, mismatched nucleotides can be inserted, and if not repaired by mismatch repair (MMR), mutations may arise (Fahrer and Christmann, 2023). Apart from direct lesion reversal and TLS, nucleotide excision repair (NER) is thought to be the predominant pathway for repairing bulky O6-alkyl-dG adducts (Du et al., 2019).

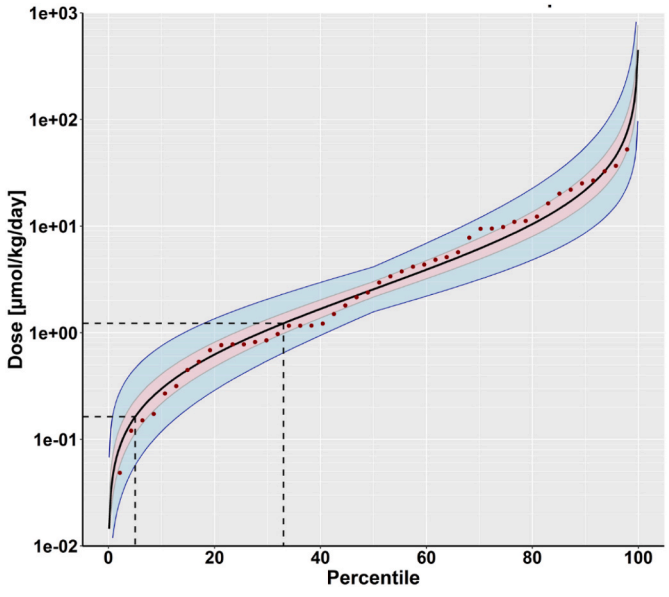


Fig. 3. Percentile estimates for molar-based Lhasa TD50 values for nitroso-compounds. Dashed lines indicate the 5th and 33rd percentile. Red shading represents the 50% confidence interval; blue shading shows the 95% confidence interval.

As a result of these different DNA adducts and DNA adduct spectra (and other factors, such as efficiency of metabolic activation impacted by steric hindrance and electron configuration), there is a wide range of mutagenic and cytotoxic potencies of N-nitrosamines, which can be estimated via SAR (Cross and Ponting, 2021; Thomas et al., 2022). Note, that the larger nitrosamines which induce larger DNA adducts are predicted to be more cytotoxic than smaller nitrosamines which induce smaller DNA adducts. Smaller DNA adducts are widely accepted as having higher mutagenic potency than larger DNA adducts for this reason.

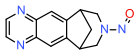
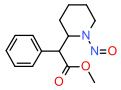
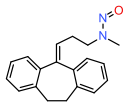
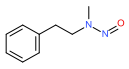
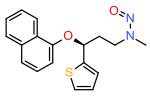
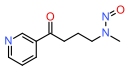
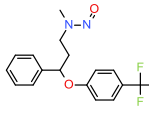
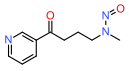
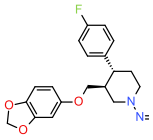
One activated nitroso group can only cause a single DNA adduct at maximum, which may manifest in a single mutation if not properly repaired. Because of this mechanism, AI values for nitrosamines should be scaled to the nitrosamines molecular weight, as the number of nitroso groups per mass unit of nitrosamine molecule is proportional to the nitrosamines molecular weight. Looking from a different angle, the carcinogenic potency of a nitrosamine depends on both the rate at which it can be metabolically activated to form DNA-alkylating diazonium ions and the rate at which the respective DNA lesions can be repaired. The potency is measured as the endpoint of *in vivo* carcinogenicity studies as the effect of a certain quantity (not mass) of nitrosamine molecules being present. In other words, the true potency as equilibrium between activation and repair is only encoded in the molarity at the endpoint, hence if any statistical extrapolation from TD_{50} is to be performed, it must be done based on the molar values.

We have re-calculated EMA's default AI using molar equivalents of TD_{50} values from the LCDB, resulting in a molar default AI of 163 pmol/day. To derive a compound-specific AI on a mass scale this value just needs to be multiplied by the compound's molecular weight. The same rationale can be applied to derive a pmol/kg TD_{50} and associate AIs for different positions in the histogram. For example, the lower 33rd percentile, listed as 178 ng/day (EMA, 2022), would be equivalent to 1224 pmol/day. In addition, the same rationale of expressing AIs in nanomoles/day, can be applied to different databases, or for future revisions of the Lhasa database (LCDB). Following the same logic, AIs derived by read-across to structurally related surrogate compounds should also be scaled to the target compound's molecular weight. We have performed this exercise for the NDSRIs listed in the EMA Q&A document (Table 2), with the outcome that the MW-scaled AIs are 1.6–3.1 times higher. It should be noted that this scaling has lifted the proposed AIs for those compounds with nitrosopiperidine as a read-across analogue above the 1.5 µg/day TTC; should this scaling to values above the TTC not be acceptable in a regulatory context, an alternate approach would be to cap those NDSRIs at the TTC, pending a future conversion of the TTC to molar units and general acceptance of molar scaling. This treatment would essentially consider them as all other non-cohort mutagenic impurities without compound-specific carcinogenicity data.

There are currently two NDSRIs, i.e., N-Nitrosorasagiline and N-Nitrosodabigatran, for which EMA has stipulated the use of the default AI of 18 ng/day (updated Q&A document (EMA, 2022)); considering their molecular weight of 200.3 g/mol and 500.5 g/mol, and using the molar default AI of 163 pmol/day, their AI expressed on a mass scale would be 32.6 ng/day and 81.6 ng/day, respectively, corresponding to 181% and 453% of the original default AI.

Mathematically, if expressed on a mass scale, the default AI provides the intended protection of a maximum of 1 additional cancer case in 100,000 people, in their lifetime, only for a substance with a molecular weight of 155.9 g/mol (25.4 ng/day/0.163 pmol/day). Many nitrosamines and especially NDSRIs have a much higher molecular weight. It was recently demonstrated that a substantial fraction of APIs and API impurities are potential nitrosamine (NDSRI) precursors as they are secondary or tertiary amines (Schlingemann et al., 2022b). The median molecular weight from this group of ca. 8000 potential NDSRIs is 399.4 g/mol. At this point, the protection factor of an 18 ng/day AI is not 10^5,

Table 2
NDSRIs and their stipulated AIs from the EMA Q&A document derived by read-across as well as alternative AIs that are scaled to the compound's molecular weight (Nudelman et al., 2023). This table does not include correction for LTL exposure, as it is currently not accepted for long-term application, although Bercu et al. (2021) have clearly shown that LTL corrected AIs are protective for potential carcinogenic risk to patients.

NDSRI	Structure	MW	Source	EMA AI [ng/day]	Read-across surrogate	Surrogate MW	MW ratio	MW-scaled AI [ng/day]
MeNP		129.2	Rifampicin	**26.5**		73.1	1.8	**46.8**
NNV		240.3	Varenicline	**37**[a]		112.1	2.1	**79.3**
NMPH		262.3	Methyl-phenidate	**1300**		114.2	2.3	**2986**
Nitroso-nortriptyline		292.4	Amitriptyline	**8**[b]		164.2	1.8	**14.2**
Nitroso-duloxetine		326.4	Duloxetine	**100**		207.2	1.6	**157.5**
Nitroso-fluoxetine		338.3	Fluoxetine	**100**		207.2	1.6	**163.3**
Nitroso-paroxetine		358.4	Paroxetine	**1300**		114.2	3.1	**4080**

[a] *Nitroso-hexamethylenimine (NHEX) with a* TD_{50} *of* 313 μg/kg/*day may be a better point of departure for read-across* (Nudelman et al., 2023; Ponting et al., 2022).
[b] *The use of this value for this compound has been challenged based on the use of a multi-organ (upper GI tract) value rather than an organ-specific value for deriving the* TD_{50}*; the corresponding single-organ (esophagus) study would indicate a value of 40.*1 ng/*day (Woolley and Cross, 2023).*

but 3.62 * 10^5, or in other words, default nitrosamine AI limits could be on average almost 4x higher (the median being at 65.2 ng/day) than they currently are and would still be fully protective as intended. Fig. 4 illustrates how an 18 ng/day AI (dotted black line) falls into histograms of MWs of nitroso compounds from the Gold and Lhasa databases and nitrosamines derived from collections of APIs and API impurities (NDSRIs). Clearly, the mass-based default AI of 18 ng/day is too conservative for essentially all these NDSRIs due to their increased MW compared with the low-MW nitroso compounds that were used to derive it. Although not all these NDSRIs will truly exist and of the ones that do not all will be potent mutagens, this still underlines the urgent need for reasonable AIs for structurally complex nitrosamines. In this regard the recent updates of the EMA Q&A on nitrosamines (EMA, 2023) as well as the respective guidances from FDA and Health Canada and the introduction of the so-called carcinogenic potency categorization approach (CPCA) have substantially improved the situation around AIs for NDSRIs. CPCA allows the structure-based classification of nitrosamines into five categories, of which only the most potent category 1 requires adherence to the 18 ng/day default AI, whereas the less potent categories 2; 3; 4 and 5 allow for AIs of 100 ng/day, 400 ng/day and 1500 ng/day (for categories 4 & 5), respectively. However, further improvement could be achieved by application of a molar scale to the AIs connected with the CPCA categories, at least for categories 1–3. For category 1 the applicable limit would be 163 pmol/day as previously described. As the derivation of the 100 ng/day and 400 ng/day AIs hasn't yet been disclosed, we cannot calculate their molar equivalents at this time, but it can be assumed that the back-calculated mass-based AIs for most category 2 and 3 NDSRIs would be considerably higher than 100 ng/day and 400 ng/day, respectively, on average most likely by a factor of 3–4 as it is the case for the 18 ng/day default AI. For CPCA categories 4 and 5, molar scaling requires the conversion of the mass based TTC and would frequently result in back-calculated AIs above the current 1500 ng lifetime daily limit for regular mutagens defined in the ICH M7 guideline. Although the CPCA even in its current format (mass scale) brings relief for many NDSRIs by placing them in categories 4 and 5, there is still a substantial number that will remain in the high-potency categories and would benefit greatly from molecular weight scaling (Burns et al., 2023).

While undeniably protective, the use of over-conservative AI limits has potential unintended negative patient impacts (e.g., unnecessary product recall and drug shortages). From a chemical/manufacturing point of view it may be impossible in many cases to obtain a product that complies with the 18 ng/day criterion or another non-MW-corrected AI, whereas compliance with the 163 pmol/day criterion or a MW-corrected

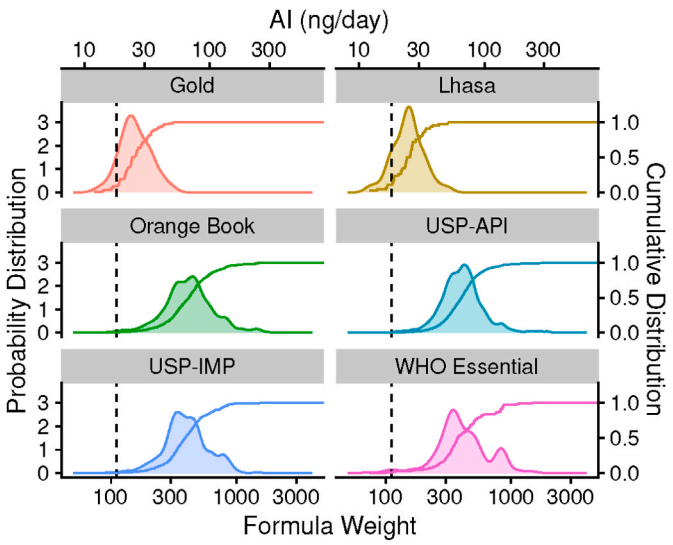


Fig. 4. Acceptable intake [ng/day] based on 117 Gold/CPDB and 46 Lhasa TD_{50} values for N-nitroso compounds and calculated from the molecular weights of potential nitrosamines (taken from the Schlingemann et al. (Schlingemann et al., 2022b)) derived from 2211 APIs (1739 potential nitrosamines from secondary and tertiary amines) listed in the FDA "Orange Book" of Approved Drug Products with Therapeutic Equivalence Evaluations, 563 APIs (287 potential nitrosamines) from the WHO list of essential medicines as well as from 8611 APIs (7895 potential nitrosamines) and 3564 API impurities (2213 potential nitrosamines) from the USP's installation of the Global Substance Registration System (GSRS), using the molar default AI of 163 pmol/day to convert from formula weight to mass-based AI. The dotted black vertical lines represent an AI of 18 ng/day.

AI derived by read-across is more likely to be technically feasible. More importantly, the assessment described within this paper demonstrates that patient safety will be maintained through application of MW-corrected AI. *In vivo* dose response work that is being carried out on N-nitrosamines by many stakeholders will provide further information on the use of mutagenic potency to determine protective and achievable AI limits, and this will lead to improved ability for stratification across the complete range of N-nitrosamine structures and more accurate read across predictions. The results of this *in vivo* mutagenicity work can be compared to analysis of chemical series of compounds with carcinogenicity data to further reinforce this, potentially using benchmark dose analysis as has been done for NDMA and NDEA (Johnson et al., 2021). Concurrently, as there is no indication for increased mutagenic potency in larger nitrosamines, it will be essential to allow for the consideration of molecular weight when nitrosamine AIs are to be derived by statistical extrapolation or read across. In conclusion, there are at least three scenarios (discounting the use of a class-based limit such as has been proposed for the 'olol' drugs, non-carcinogenic nitrosamines, or those positive but of low enough potency that the TTC is sufficiently protective) to consider when determining the AI for a novel NDSRI, and the use of molar scaling will allow for the determination of more relevant AI limits in two of them:

- Compound-specific carcinogenicity data: No scaling needed
- Specific read-across analogue identified: Scale by relative molecular weight of NDSRI and analogue
- No suitable analogue or class identified, default limit needed: 163 pmol * MW of NDSRI, rather than the mass-based 18 ng.

While the focus of this manuscript is on nitrosamines, it should be noted that the arguments here can, as described, be applied to all stoichiometric covalent reactivity-mediated toxicological outcomes, such as the LLNA described, where the mechanism that leads to toxicity consumes the toxicant.

Supplementary material

The R/Python code used for data analysis and visualization is provided as online supplementary material.

Funding & conflict of interest

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. All authors are employed by their respective affiliations. The authors declare no conflict of interest, financial or otherwise. GJ is a consultant who evaluates the risks posed by pharmaceutical impurities. His clients did not influence the content of this manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. All authors are employed by their respective affiliations.

CRediT authorship contribution statement

Jonathan Fine: Writing – original draft, Writing – review & editing, Formal analysis, Visualization. **Leonardo Allain:** Conceptualization, Writing – original draft, Writing – review & editing. **Joerg Schlingemann:** Conceptualization, Writing – original draft, Writing – review & editing. **David J. Ponting:** Writing – review & editing, Investigation. **Robert Thomas:** Writing – review & editing, Formal analysis, Visualization. **George E. Johnson:** Writing – review & editing, All authors reviewed the results and approved the final version of the manuscript.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: GJ is a consultant who evaluates the risks posed by pharmaceutical impurities. His clients did not influence the content of this manuscript.

Data availability

Data will be made available on request.

Acknowledgements

We thank Mark Powley (Merck & Co., Inc., Rahway, NJ, USA) and Andrew Thresher (Lhasa) for their substantial support and valuable comments on the manuscript.

References

Agarwal, P., et al., 2022. Trends in small molecule drug properties: a developability molecule assessment perspective. Drug Discov. Today 27, 103366.
American Cancer Society, 2022. Lifetime Probability of Developing and Dying from Cancer, 2017-2019 (Cancer Facts & Figures 2023 Supplemental Data).
Bercu, J.P., et al., 2021. Use of less-than-lifetime (LTL) durational limits for nitrosamines: case study of N-Nitrosodiethylamine (NDEA). Regul. Toxicol. Pharmacol. 123, 104926.
Boetzel, R., et al., 2022. A nitrite excipient database: a useful tool to support N-nitrosamine risk assessments for drug products. J. Pharmaceut. Sci. 112, 1615–1624.
Burns, M.J., et al., 2023. Revisiting the landscape of potential small and drug substance related nitrosamines in pharmaceuticals. J. Pharmaceut. Sci.
Burns, M.J., et al., 2020. Controlling a cohort: use of mirabilis-based purge calculations to understand nitrosamine-related risk and control strategy options. Org. Process Res. Dev. 24, 1531–1535.
Chain, E. P. o. C. i. t. F., 2023. Risk assessment of N-nitrosamines in food. et al. EFSA J. 21, e07884.
Cross, K.P., Ponting, D.J., 2021. Developing structure-activity relationships for N-nitrosamine activity. Comput. Toxicol. 20, 100186.

Du, H., et al., 2019. Repair and translesion synthesis of O6-alkylguanine DNA lesions in human cells. J. Biol. Chem. 294, 11144–11153.
EMA, 2022. Questions and Answers for Marketing Authorisation Holders/applicants on the CHMP Opinion for the Article 5(3) of Regulation (EC) No 726/2004 Referral on Nitrosamine Impurities in Human Medicinal Products (EMA/409815/2020 Rev.11; 29 July 2022).
EMA, 2023. Questions and Answers on "Information on Nitrosamines for Marketing Authorisation Holders" (EMA/409815/2020 Rev.16).
EMA/CHMP, 2020. Procedure under Article 5(3) of Regulation EC (No) 726/2004; Assessment Report: Nitrosamine Impurities in Human Medicinal Products.
Fahrer, J., Christmann, M., 2023. DNA alkylation damage by nitrosamines and relevant DNA repair pathways. Int. J. Mol. Sci. 24.
Gold, L.S., et al., 2005. Supplement to the carcinogenic potency database (CPDB): results of animal bioassays published in the general literature through 1997 and by the national Toxicology Program in 1997–1998. Toxicol. Sci. 85, 747–808.
Gold, L.S., et al., 1991. The Carcinogenic Potency Database: analyses of 4000 chronic animal cancer experiments published in the general literature and by the U.S. National Cancer Institute/National Toxicology Program. Environ. Health Perspect. 96, 11–15.
Harris, C.R., et al., 2020. Array programming with NumPy. Nature 585, 357–362.
Hecht, S.S., 1998. Biochemistry, biology, and carcinogenicity of tobacco-specific N-nitrosamines. Chem. Res. Toxicol. 11, 559–603.
ICH, 2017. ICH Guideline M7(R1) on Assessment and Control of DNA Reactive (Mutagenic) Impurities in Pharmaceuticals to Limit Potential Carcinogenic Risk.
Johnson, G.E., et al., 2021. Permitted daily exposure limits for noteworthy N-nitrosamines. Environ. Mol. Mutagen. 62, 293–305.
Lhasa Limited, 2023. Lhasa Carcinogenicity Database.
Li, Y., Hecht, S.S., 2022a. Metabolic activation and DNA interactions of carcinogenic N-nitrosamines to which humans are commonly exposed. Int. J. Mol. Sci. 23, 4559.
Li, Y., Hecht, S.S., 2022b. Metabolism and DNA adduct formation of tobacco-specific N-nitrosamines. Int. J. Mol. Sci. 23, 5109.
Lidén, C., et al., 2016. Comparative sensitizing potencies of fragrances, preservatives, and hair dyes. Contact Dermatitis 75, 265–275.
Lijinsky, W., 1987. Structure-activity relations in carcinogenesis by N-nitroso compounds. Cancer Metastasis Rev. 6, 301–356.
Lipinski, C.A., et al., 2001. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings1PII of original article: S0169-409X(96)00423-1. The article was originally published in Advanced Drug Delivery Reviews 23 (1997) 3–25.1. Adv. Drug Deliv. Rev. 46, 3–26.
McKinney, W., 2010. Data Structures for Statistical Computing in Python.
Natsch, A., et al., 2014. Predicting skin sensitizer potency based on in vitro data from KeratinoSens and kinetic peptide binding: global versus domain-based assessment. Toxicol. Sci. 143, 319–332.
Nudelman, R., et al., 2023. The nitrosamine 'saga': lessons learned from five years of scrutiny. Org. Process Res. Dev. in press.
Oliveira, A.A., et al., 2022. Collaborative analysis of complex nitrosamines. In: Society of Toxicology Annual Meeting, p. 5063.
Parrish, R.S., 1990. Comparison of quantile estimators in normal sampling. Biometrics 46, 247.
Peto, R., et al., 1984. The TD50: a proposed general convention for the numerical description of the carcinogenic potency of chemicals in chronic-exposure animal experiments. EHP (Environ. Health Perspect.) 58, 1–8.
Ponting, D.J., et al., 2022. Strategies for assessing acceptable intakes for novel N-nitrosamines derived from active pharmaceutical ingredients. J. Med. Chem. 65, 15584–15607.
Ponting, D.J., Foster, R.S., 2023. Drawing a line: where might the cohort of concern end? Org. Process Res. Dev. https://doi.org/10.1021/acs.oprd.3c00008.
R_Core_Team, R, 2021. A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Schlingemann, J., et al., 2022a. Avoiding N-nitrosodimethylamine formation in metformin pharmaceuticals by limiting dimethylamine and nitrite. Int. J. Pharm. 620, 121740.
Schlingemann, J., et al., 2022b. The landscape of potential small and drug substance related nitrosamines in pharmaceuticals. J. Pharmaceut. Sci. 112, 1287–1304.
Schoonjans, F., et al., 2011. Estimation of population percentiles. Epidemiology 22, 750–751.
Snodin, D.J., 2023. Mutagenic impurities in pharmaceuticals: a critical assessment of the cohort of concern with a focus on N-nitrosamines. Regul. Toxicol. Pharmacol. 141, 105403.
Thomas, R., et al., 2022. What makes a potent nitrosamine? Statistical validation of expert-derived structure–activity relationships. Chem. Res. Toxicol. 35, 1997–2013.
Thomas, R., et al., 2021. Utilisation of parametric methods to improve percentile-based estimates for the carcinogenic potency of nitrosamines. Regul. Toxicol. Pharmacol. 121, 104875.
Van Rossum, G., Drake Jr., F.L., 2009. Python 3 Reference Manual. CreateSpace, Scotts Valley, CA.
Virtanen, P., et al., 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods 17, 261–272.
Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer, New York.
Wickham, H., et al., 2023a. Dplyr: A Grammar of Data Manipulation.
Wickham, H., et al., 2023b. Tidyr. Tidy Messy Data.
Woolley, D., Cross, K., 2023. Risk (Re)assessment of N-methyl-N-Nitrosophenethylamine for use in computing acceptable intake levels of N-nitrosamine drug substance-related impurities. In: SOT 62nd Annual Meeting & ToxExpo, Nashville TN.