Supplementary Material I

Geometric means (ratio) and logarithmic data transformation

The geometric mean (*GM*) for a variable X estimated from a sample of size *n* is given by

$$\left(\prod_{i=1}^{n} X_i\right)^{\frac{1}{n}} \leq \frac{1}{n}\sum_{i=1}^{n} X_i.$$

The value is always less than or equal to the arithmetic mean and can alternatively be obtained by back-transformation of the arithmetic mean of the data logarithms ($\mu_{\log}$) according to

$$GM = \exp\left\{\mu_{\log X}\right\}.$$

In a similar manner, the geometric mean ratio for two variables, *X* and *Y*, is obtained by back-transformation of the difference between the means on the logarithmic scale, as

$$GMR = \exp\left\{\mu_{\log X} - \mu_{\log Y}\right\}.$$

Logarithmic transformation of the data prior to statistical analysis has been extensively used in statistical applications in the biosciences [1-3]. Equally, the use of logarithmically transformed data as input for statistical analyses has been accepted and recommended by regulatory authorities, notably in the bioequivalence studies framework [4, 5].

In the comparative analysis of the effects of 3R4F and 1R6F cigarette smoke, we worked on log-transformed data and followed a standard two independent samples comparison setting. We estimated the differences between the two products and the associated standard errors on the logarithmic scale, following

$$\mu_{\{\log X - \log Y\}} = \mu_{\log X} - \mu_{\log Y} \quad \text{and} \quad \mathrm{var}_{\{\log X - \log Y\}} = \mathrm{var}_{\log X} + \mathrm{var}_{\log Y},$$

and we back-transformed the final estimates to obtain the final estimates.

Confidence intervals and equivalence limits

Regulatory recommendations in bioequivalence studies stipulate that a claim of bioequivalence is possible if the 90% confidence interval of the geometric mean ratio is included between the equivalence limits $\delta_{\mathrm{L}}$and δ_{U}, with these limits set to 0.8 and 1.25 [4, 6, 7]. Note that these limits are symmetric around 1 on the ratio scale, given that $1.25 = 1/0.8$. These limits are sometimes short given the high variability inherent to the endpoint under investigation. By introducing the blue boxes in Figure 1 of the main text, we allowed these limits to extend beyond the standard 0.8 and 1.25 values when the variability of the reference product is high. To assess the variability of the reference product we need to estimate its residual variance (denoted as σ_{R}^2) after removal of any systematic patterns in its long-term variability, e.g. study-to-study variability. Historical data are used for this purpose, and the residual variance is estimated as

(1) $$\sigma_{\mathrm{R}}^2 = \mathrm{var}(\varepsilon)$$

from the linear statistical model:

(2) $$\mathrm{Y}_{mn} = \mu + \mathrm{S}_m + \varepsilon_{mn}.$$

The term S_m represents the study-specific effect and contributes to random study-to-study variation, which is removed from the residual variance estimate σ_R^2 . Any other design factor (e.g., a blocking term if it exists) should also be removed from the error term, so that the latter reflects only the residual or within-subject variability. The subscripts m and n denote study and subject in the study, respectively.

The equivalence ranges are then computed according to

(3) $$[\delta_{\mathrm{L}}, \delta_{\mathrm{U}}] = \exp\left\{\pm \max\left(\log(1.25), \left(t_{\alpha,df} + t_{\beta/2,df}\right) \frac{\sigma_{\mathrm{R}}}{\sqrt{n}}\right)\right\},$$

where the constants t_α and $t_{\beta/2}$ represent the student percentiles at consumer's and producer's risks α and β, respectively, with $df = 2n - 2$ degrees of freedom where $2n$ is the total sample of animals allocated to the two products. Note that $\exp\{\delta\}$ back-transforms the limits (δ) from the logarithmic scale to the original scale in the same manner as for the geometric mean ratio. The equivalence limits $[\delta_{\mathrm{L}}, \delta_{\mathrm{U}}]$ are extended by means of expression (3) beyond the [0.8, 1.25] standard bioequivalence limits to accommodate for extra variability commonly present in *In vivo* testing. Expression (3) was proposed in [8] and was principally motivated in bioequivalence studies for the development of drugs with highly variable

response. As a referee correctly mentioned it would be beneficial to the reader to make the link between expression (3) and the minimum sample size requirements for achieving a $1-\beta$ statistical power at the nominal significance level α; for this, see Chapter 5 in Chow 1992 [9] as well as Hauschke et al. 1992 [10]. Note that we use here the nominal α level, not the $\alpha/2$ level. This results from testing interval hypotheses using the TOST approach as described in the main text of the manuscript. Proposals other than expression (3) for widening the equivalence limits can be found in refs [11-13]. An alternative approach used for widening the equivalence limits, recommended by the European Medicines Agency (EMA) Committee in their revised guidelines [5], is to define the equivalence limits according to

$$(4) \quad [\delta_L^*, \delta_U^*] = \exp\{\pm \max(\log(1.25), k\,\sigma_{\mathrm{R}})\},$$

where *k* is commonly set to 1 or less. This approach dissociates the equivalence limits from the consumer's and producer's risks as well as from the sample size of the study. We report in Table 1 and Table 2 of the main text lower and upper equivalence limits for the two proposals under the two different scenarios in expression (3) and (4), respectively.

Extensions and generalizations

The analyses described above fixed a (1 - 2α)% confidence interval for equivalence assessment, used the geometric mean ratio for comparisons with data analyzed on the logarithmic scale, followed a two-sample problem formulation for the comparative assessment, and used an intuitive approach to widen the equivalence bands. Other statistical methods could be implemented for each step to improve the comparisons framework. A linear statistical model of the form

$$\mathrm{Y}_{ijn} = \mu + \mathrm{B}_i + \mathrm{P}_j + \varepsilon_{ijn}$$

could be used to estimate the difference between the two products via the estimated coefficients for P_j (e.g., P_j for $j = 3\mathrm{R}4\mathrm{F}, 1\mathrm{R}6\mathrm{F}$) while using model parameters (e.g., B_i) to control for other study design parameters and blocking factors such as laboratories or analytical methods. The latter becomes very interesting in interlaboratory studies and proficiency tests where laboratories can share their historical data and allow the variable 'laboratory' entering the statistical model above as a covariate (B_i) reflecting potential differences in analytical methods, conditioning and other factors. The EMA guidelines [5] recommend the use of statistical models in bioequivalence studies. Such models should then be checked

for all model parametric assumptions. Model diagnostics can be used, and if necessary, corrective action taken.

References

1. Crow, E.L. and K. Shimizu: Log-normal Distributions: Theory and Application., Dekker, 1988
2. Kondylis, A.: Median Estimation of Chemical Constituents for Sampling on Two Occasions under a Log-normal Model; Biometrical J. 57 (2015) 914–929. DOI: 10.1002/bimj.201400095
3. Limpert, E., Stahel, W.A. and Abbt, M.: Log-normal Distributions across the Sciences: Keys and Clues; BioScience. 51 (2001) 341–352.
4. Chow, S.C. and J.P. Liu: Recent Statistical Developments in Bioequivalence Trials - A Review of the FDA Guidance; Drug Inf. J. 28 (1994) 851–864.
5. EMA, *Guideline on the Investigation of Bioequivalence*. 2010: London.
6. Berger, R.a.H., J.: Bioequivalence Trials, Intersection-union Tests and Equivalence Confidence Sets; Stat. Sci. 11 (1996) 283–319.
7. United States Food and Drug Administration (FDA), *Statistical Approaches to Establishing Bioequivalence*, in *Guidelines for industry*. 2001, Food and Drug Administration, Center for Drug Evaluation and Research.
8. Boddy, A.W., F.C. Snikeris, R.O. Kringle, G.C. Wei, J.A. Oppermann, and K.K. Midha: An Approach for Widening the Bioequivalence Acceptance Limits in the Case of Highly Variable Drugs; Pharm. Res. 12 (1995) 1865–1868. DOI: 10.1023/a:1016219317744
9. Chow, S.a.L., J.: Design and Analysis of Bioavailability and Bioequivalence Studies; 3rd ed., Dekker, 1992
10. Hauschke, D., V.W. Steinijans, E. Diletti, and M. Burke: Sample Size Determination for Bioequivalence Assessment using a Multiplicative Model; J. Pharmacokinet. Biopharm. 20 (1992) 557–561. DOI: 10.1007/BF01061471
11. Dangi, Y.S., M.L. Soni, and K.P. Namdeo: Highly Variable Drugs: Bioequivalence Requirements and Regulatory Perspectives; Int. J. Curr. Pharm. Res. 3 (2010) 24–28.
12. Endrenyi, L. and L. Tothfalusi: Bioequivalence for Highly Variable Drugs: Regulatory Agreements, Disagreements, and Harmonization; J. Pharmacokinet. Pharmacodyn. 46 (2019) 117–126. DOI: 10.1007/s10928-019-09623-w
13. Tothfalusi, L., L. Endrenyi, K.K. Midha, M.J. Rawson, and J.W. Hubbard: Evaluation of the Bioequivalence of Highly-variable Drugs and Drug Products; Pharm. Res. 18 (2001) 728–733. DOI: 10.1023/a:1011015924429

Supplemental material II

A) Test atmosphere

The following analytical parameters were determined at the breathing zone of the rats:

Parameter	Method	Determination schedule
TPM	Gravimetry after trapping on Cambridge filter pads	≥1 time per 6-h exposure; approximately 0.5 h per sampling for sham chamber and ≥4 times per 6-h exposure; approximately 0.5 h per sampling for cigarettes chambers
Carbon monoxide	Nondispersive infrared photometry of gas/vapor phase	Continuously
Nicotine	Capillary gas chromatography after trapping on sulfuric acid-impregnated diatomaceous earth	≥1 time per week; approximately 0.5 h per sampling for sham chamber and ≥4 times per 6-h exposure; approximately 0.5 h per sampling for cigarettes chambers
Aldehydes	Reverse phase high-performance liquid chromatography of DNPH derivatives after trapping in DNPH solution	≥1 time per week; approximately 0.5 h per sampling
Particle size distribution	Aerodynamic particle sizer (spectrophotometric)	≥1 time per week

B) Cigarette smoke uptake parameters

Aerosol uptake and exposure were monitored during the study through measuring the following parameters:

i) Respiratory physiology which included respiratory minute volume, tidal volume, respiratory frequency and peak inspiratory flow was assessed during a time period in the middle of exposure phase.

ii) Carboxyhemoglobin (COHb) level in blood.

iii) Nicotine and selected nicotine metabolites (trans-3'-hydroxycotinine, norcotinine, cotinine, nicotine-N'-oxide, nornicotine) were evaluated in 24-hour urine samples (including 6-hour exposure period and approximately 18 hours post-exposure period).

Parameter	Method	Frequency	Number of rats (OECD)	Remarks
Respiratory physiology	Head-out plethysmography	Once during 90-d exposure period	At least 8 male and 8 female rats per group	Individually
Blood carboxy-hemoglobin	Hemoglobin derivatives measured spectrophotometrically on the basis of Lambert-Beer's law	Once during 90-d exposure period	At least 8 male and 8 female rats per group	Individually
Nicotine and selected nicotine metabolites in urine	LC-MS/MS (non-GLP)	Once during 90-d exposure period	At least 8 male and 8 female rats per group	24-h urine collection

C) Inflammatory cells from bronchoalveolar lavage analysis

Bronchoalveolar lavage fluid (BALF) was collected from the right lungs of all rats at necropsy. Two lavage media were used for the collection of free lung cells from the first to fifth cycle: cycle 1, PBS without bovine serum albumin (BSA); cycle 2-5, PBS with 0.325% w/v BSA. Cells were pooled from cycles 1-5 and analyzed by flow cytometry.

Parameter	Method	Number of rats	Remarks
Collection of free lung cells (FLC)	Bronchoalveolar lavage of the right lung	All rats	At dissection
FLC count	Flow cytometric counting using Trucount® tubes	All rats	At dissection

Parameter	Method	Number of rats	Remarks
Determination of cellular viability	Flow cytometric analysis after staining with fluorescein diacetate and propidium iodide	All rats	At dissection
FLC differentiation	Flow cytometric analysis after staining with cell-type specific antibodies conjugated to fluorochromes and counterstaining with the nucleic acid dye propidium iodide	All rats	At dissection

D) Clinical chemistry biomarkers

Blood samples were taken from rats under pentobarbital anesthesia during exsanguination via the abdominal aorta and processed to isolate the serum. Analysis of the serum samples was performed on the UniCel® DXC 600 clinical analyzer for the parameters described below.

Parameter	Principle/method
Calcium concentration	Potentiometric determination (ion-selective electrode)
Inorganic phosphate concentration	Photometric determination of ammonium-12-molybdo-phosphate which is formed by reaction of inorganic phosphate with ammonium molybdate
Chloride concentration	Potentiometric determination (ion-selective electrode)
Sodium concentration	Potentiometric determination (ion-selective electrode)
Potassium concentration	Potentiometric determination (ion-selective electrode)
Glucose concentration	Photometric determination of β-nicotinamide adenine dinucleotide (NADH)* after hexokinase-catalyzed phosphorylation of glucose followed by enzymatic oxidation of glucose-6-phosphate
Alanine amino-transferase activity	Photometric determination of NADH* during enzymatic reduction of pyruvate, the product of alanine aminotransferase activity
Aspartate amino-transferase activity	Photometric determination of NADH* during enzymatic reduction of oxalacetate, the product of aspartate aminotransferase activity
Alkaline phosphatase activity	Photometric determination of the hydrolysis of the colorless organic phosphate ester substrate, p-nitrophenylphosphate, to the yellow-colored product, p-nitrophenol, and phosphate
Total bilirubin concentration	Photometric determination of azobilirubin
Protein concentration	Photometric determination of a dye complex formed from peptide bonds and copper ions in alkaline solution (Biuret method)
Albumin concentration	Photometric determination of a dye complex from albumin and bromocresol green
Globulin concentration	Total protein minus albumin
Creatinine concentration	Photometric determination of a creatinine picrate complex

Parameter	Principle/method
Urea concentration	Photometric determination of NADH* after enzymatic reduction of α-ketoglutarate to glutamate in the presence of ammonium ions cleaved from urea by urease
Total cholesterol concentration	Cholesterol esterase hydrolyzes cholesterol esters to free cholesterol and fatty acids. Free cholesterol is oxidized to cholestene-3-one and hydrogen peroxide by cholesterol oxidase. Peroxidase catalyzes the reaction of hydrogen peroxide with 4-aminoantipyrine and phenol to produce a colored quinonimine product which is determined photometrically.
Triglyceride concentration	Photometric determination of the reaction product of 4-aminophenazone, 4-chlorophenol, and hydrogen peroxide; hydrogen peroxide is produced during enzymatic oxidation of glycerol-1-phosphate after enzymatic cleavage of triglycerides and phosphorylation of glycerol.

* NADH, reduced form of nicotinamide adenine dinucleotide