# Noncompartmental Analysis, Statistical Evaluation

# Noncompartmental Analyis

- ## NCA a.k.a. SHAM (Shape, Height, Area, Moments)
  - PK metrics (plasma)
    - Single dose
      - Extent of Absorption (WHO, EEA, …), Total Exposure (USA):
        *AUC* (Area Under the Curve)
        » In most jurisdictions the PK metric for BE is $AUC_{0-t}$,
          where $t$ is the last time point with a quantifiable concentration
        » WHO, EEA: For IR products with a long half life $AUC_{0-72}$ is sufficient
        » USA and EEA (controlled release products only): additionally $AUC_{0-\infty}$
      - Rate of Absorption (WHO, EEA, …), Peak Exposure (USA): $C_{max}$
      - $t_{max}$ (Russia, Eurasian Economic Area, …)
      - Rarely relevant
        » $t_{75\%}$, POT-25 (Plateau time or peak occupancy time; time span where
          $C(t) \geq 75\%$ $C_{max}$: Russia for modified release products)
        » *MRT* (Mean of Residence Times)
        » Therapeutic Occupancy Time
          (time span where $C(t) \geq$ some given limit, *e.g.*, the MIC)

# Noncompartmental Analyis

- Multiple dose
  - Extent of Absorption (WHO, EEA, …), Total Exposure (USA):
    $AUC_{0-\tau}$ (AUC covering the dosing interval $\tau$)
    If chronopharmacological variation and more than o.a.d. regimen:
    $AUC_{0-24}$
    No extrapolation of *AUC* in any case
  - Rate of Absorption (WHO, EEA, …), Peak Exposure (USA):
    $C_{max,ss}$
  - Minimum concentration
    $C_{min,ss}$ (*l*owest observed concentration within the profile; originators)
    $C_{\tau,ss}$ (concentration at the end of the dosing interval; generics)
  - *PTF* (Peak-to-Trough Fluctuation)
    $(C_{max,ss} - C_{min,ss}) / C_{av,ss}$, where $C_{av,ss} = AUC_{0-\tau} / \tau$
  - Mentioned in some GLs but practically obsolete due to its
    extreme variability
    *Swing* = $(C_{max,ss} - C_{min,ss}) / C_{min,ss}$

# Noncompartmental Analyis

- PK metrics obtained by NCA depend much more on the sampling schedule than PK parameters estimated with a PK model
  - Examples
    - It is unlikely that one is able to 'catch' the true $C_{max}/t_{max}$ in every subject
      - Hence, frequent sampling around $t_{max}$ mandatory
    - To obtain a reliable estimate of the apparent elimination $\lambda_z$, *at least* three samples required
  - However, contrary to PK modeling NCA is independent from software
    - Paper, pencil, brain…

# PK model | *AUC*

- *AUC* is the integral of the concentration-time curve
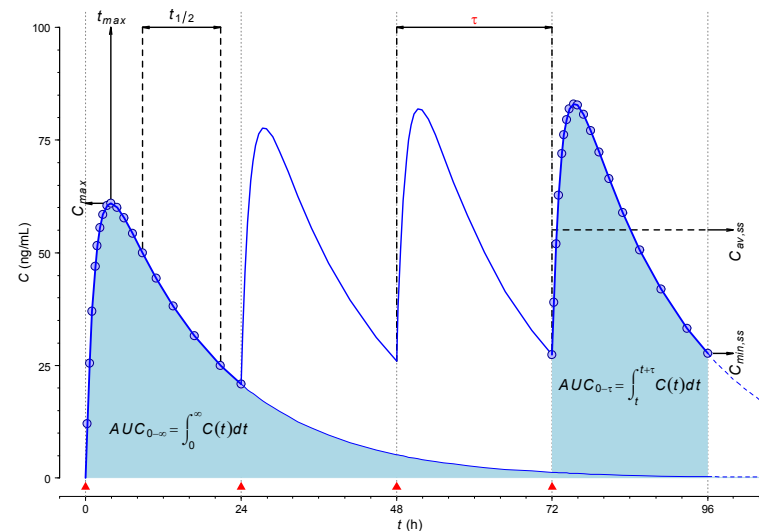  - One compartment, extravascular dose, no lag-time

$$C(t) = \frac{f \cdot D}{V} \frac{k_a}{k_a - k_e} \left( e^{-k_e \cdot t} - e^{-k_a \cdot t} \right)$$

$$AUC_{0-\infty} = \int_0^\infty C(t)\, dt$$

$$= \frac{f \cdot D}{V} \frac{k_a}{k_a - k_e} \left( \frac{1}{k_e} - \frac{1}{k_a} \right)$$

$$= \frac{f \cdot D}{V \cdot k_e}$$

$$= \frac{f \cdot D}{CL}$$



Superposition Principle of linear PK
$$AUC_{0-T} \approx AUC_{0-\infty}$$

# NCA | *AUC*

- In NCA numeric approximation of the integral is required
  – Linear trapezoidal method
  – Linear-up / logarithmic-down trapezoidal method
  – Of academic interest
    - Cubic splines
    - Lagrange polynomials
    - Simpson's rule

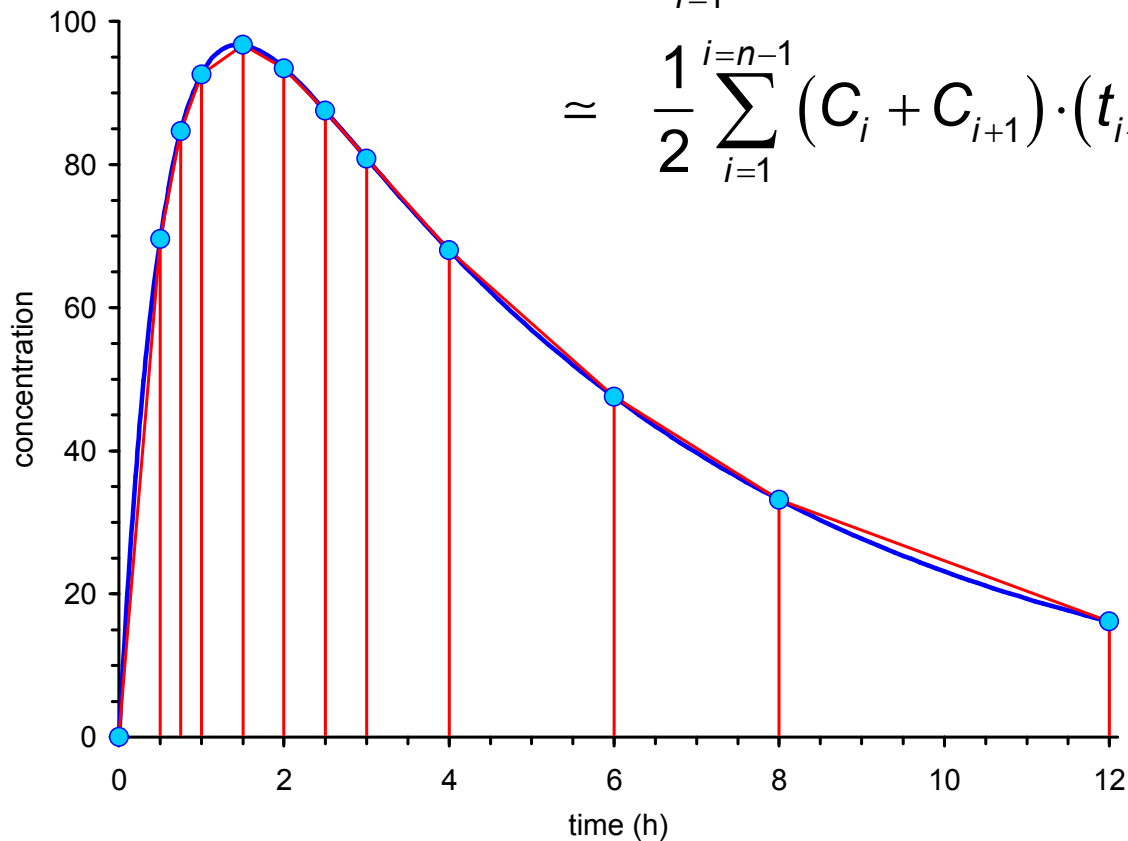# *AUC* | linear trapezoidal method

- Linear interpolation between data points
- Sections are represented by trapezoids
- Sides *a*, *b* are two neighbouring concentrations
- *h* is the time interval
- Area of one trapezoid $\quad A = \dfrac{a+b}{2} h$
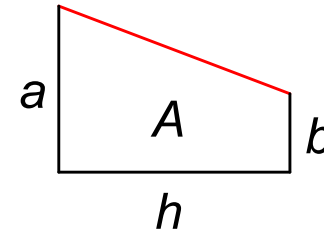
# *AUC* | linear trapezoidal method

- Total Area $AUC_{0-t_n} \simeq \displaystyle\sum_{i=1}^{i=n-1} \frac{C_i + C_{i+1}}{2}(t_{i+1} - t_i)$

$$\simeq \frac{1}{2}\sum_{i=1}^{i=n-1}(C_i + C_{i+1}) \cdot (t_{i+1} - t_i)$$

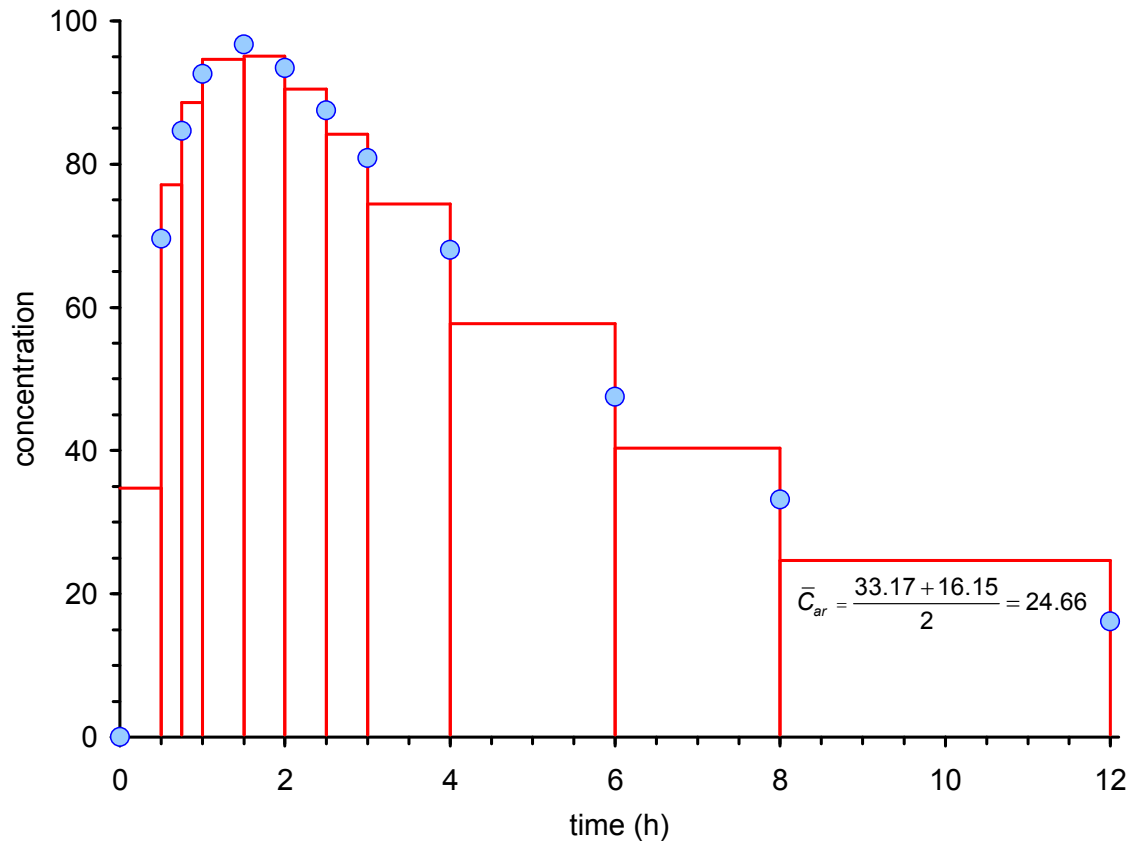# *AUC* | linear trapezoidal method



arithmetic means of neighbouring concentrations

$$\bar{C}_{ar} = \frac{33.17 + 16.15}{2} = 24.66$$

# *AUC* | linear trapezoidal method

- **Positive bias**
  - Overestimates *AUC* in both the absorption and distribution / elimination phases

- **Originated in the dark ages**
  - when profiles were plotted on paper, cut out, weighed on an analytical scale, and compared to the paper-weight of known area (*e.g.*, A4 of 80 g/m$^2$: 4.9896 g / 623.7 cm$^2$)

- **Should have been thrown into the scientific waste-can with the invention of pocket calculators decades ago**

- **In general elimination follows an exponential decrease**

$$C(t) = \frac{f \cdot D}{V} \frac{k_a}{k_a - k_e} \left( e^{-k_e \cdot t} - e^{-k_a \cdot t} \right)$$

# *AUC* | lin-up / log-down trapezoidal method

- Much better alternative:
  Linear-up / logarithmic-down trapezoidal method
- Sections with *increasing or equal* concentrations
  ($C_{i+1} \geq C_i$) calculated by the linear trapezoidal method
- Sections with *decreasing* concentrations
  ($C_{i+1} < C_i$) calculated by the logarithmic-linear trapezoidal
  method, *i.e.*,

$$AUC_{t_i - t_{i+1}} \simeq \frac{C_{i+1} - C_i}{\ln \dfrac{C_{i+1}}{C_i}} \left( t_{i+1} - t_i \right)$$
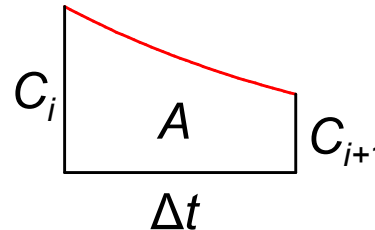
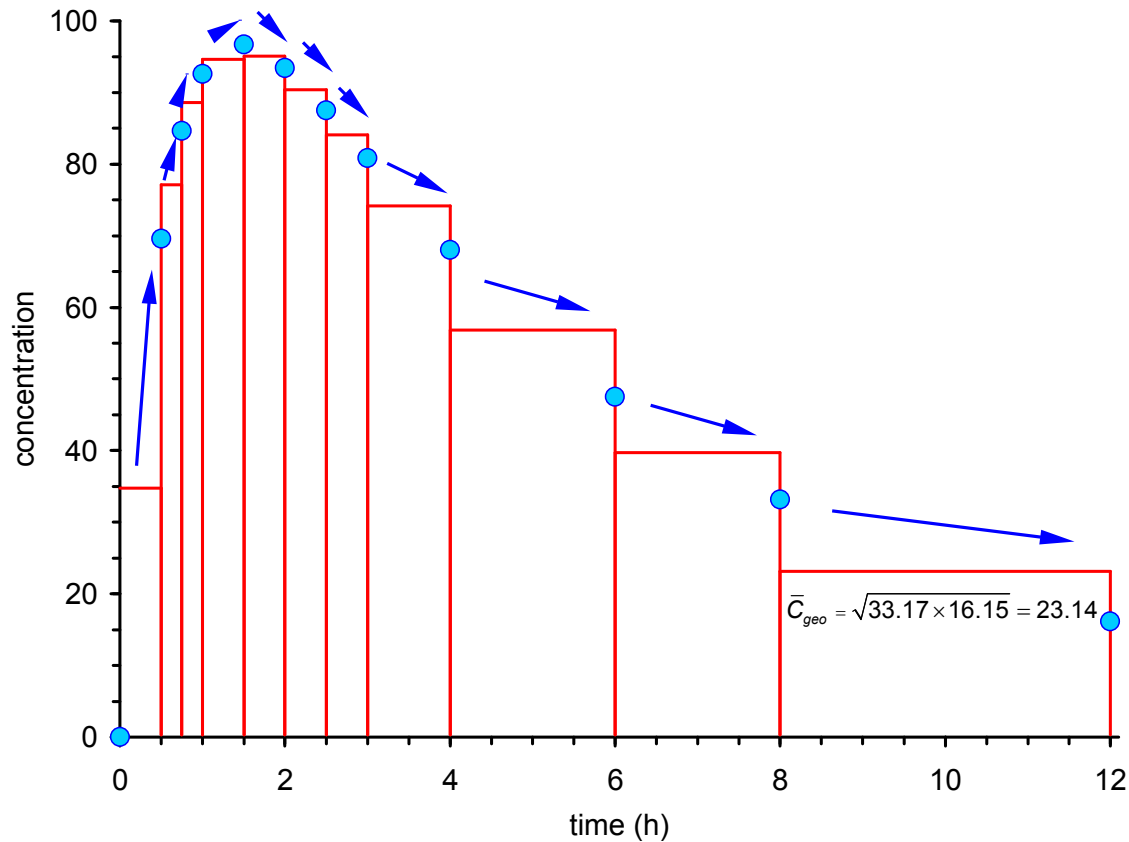# *AUC* | lin-up / log-down trapezoidal method
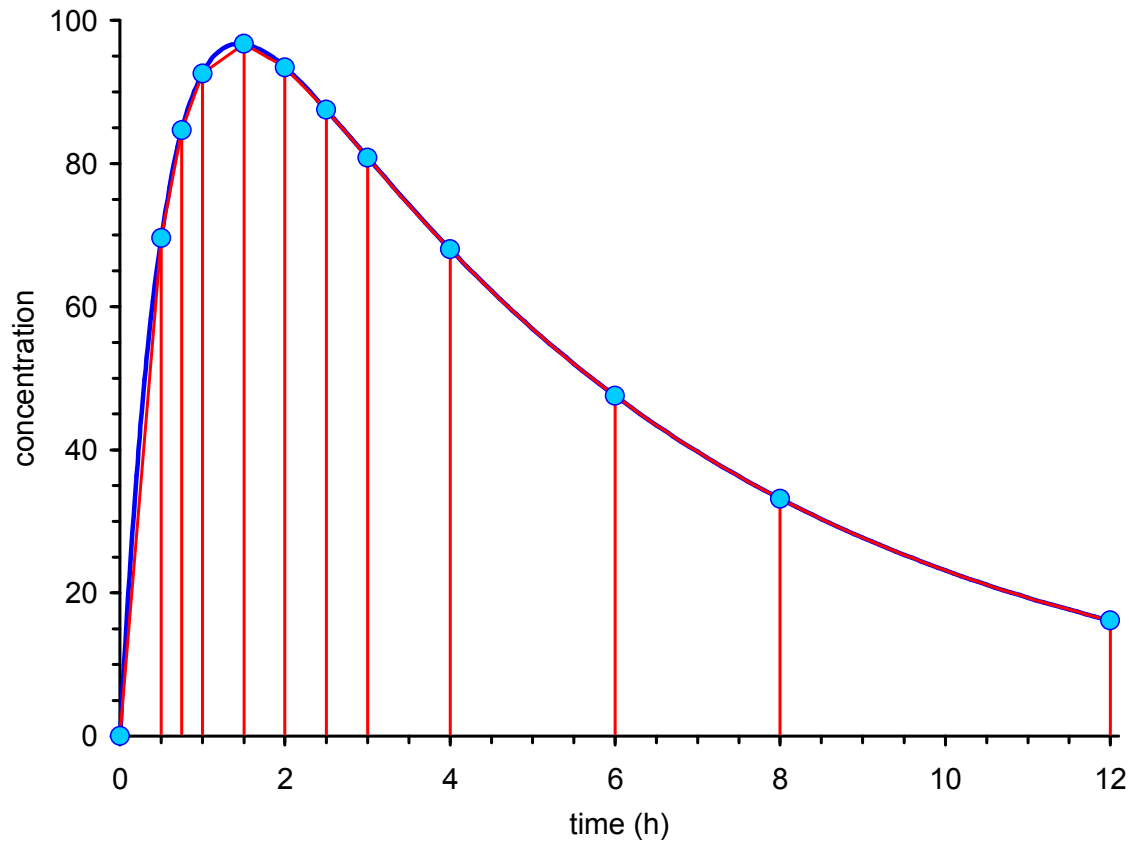
# *AUC* | lin-up / log-down trapezoidal method

*arithmetic* / *geometric* means of neighbouring concentrations



The figure shows a concentration vs. time (h) plot with cyan data points, red step/trapezoidal bars, and blue arrows. The equation shown on the plot:

$$\bar{C}_{geo} = \sqrt{33.17 \times 16.15} = 23.14$$

# *AUC* | lin-up / log-down trapezoidal method

- Avoids positive bias in distribution / elimination phases
- Suitable for both i.v. and e.v. administrations
- Suitable for multiphasic profiles
  - Secondary peaks due to enterohepatic recycling
  - Pulsatile release products
  - If *AUC* of more than one profile has to be calculated (*e.g.*, two doses with $\tau$ 12 h and $AUC_{0-24}$ is required due to circadian variation in PK)
- Implemented in standard PK software for decades
- Only exception where the method performs worse than the linear trapezoidal
  - Drugs following Michaelis-Menten PK (*e.g.*, alcohol)

# $AUC_{0-t}$ | Problem 1

- Recap: In most jurisdictions the PK metric for BE is $AUC_{0-t}$, where $t$ is the last time point with a quantifiable concentration

- Ideally we are able to calculate $AUC_{0-t}$
  - for all treatments
  - in all subjects

- What if
  - a sample was missing (*e.g.*, vial broken in centrifugation)?

- Example
  - True T/R-ratio 95%, 12 h sample (R) missing
  - Comparison of linear and lin-up / log-down trapezoidal methods

# $AUC_{0-t}$ | Problem

AUC∞ (R) 725, AUC∞ (T) 671, T/R 92.5%, bias -2.60%
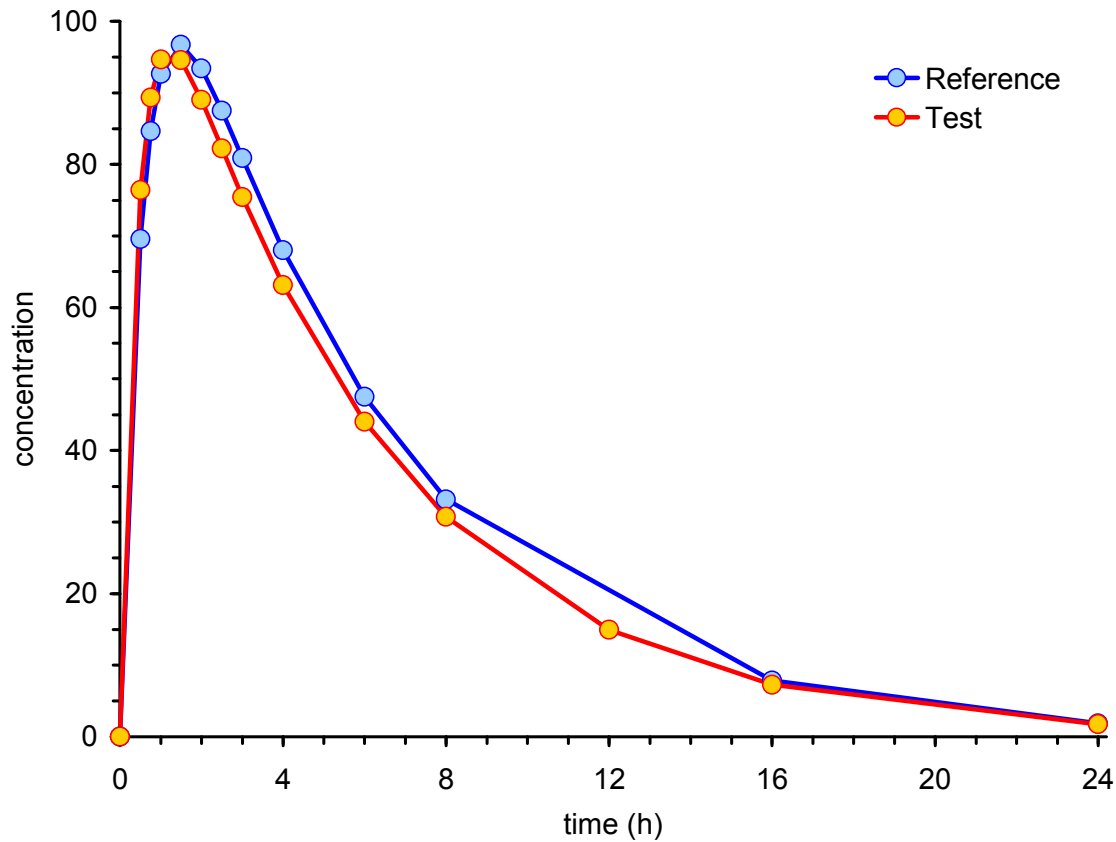
# $AUC_{0-t}$ | Solution



AUC∞ (R) 694, AUC∞ (T) 658, T/R 94.9%, bias -0.15%
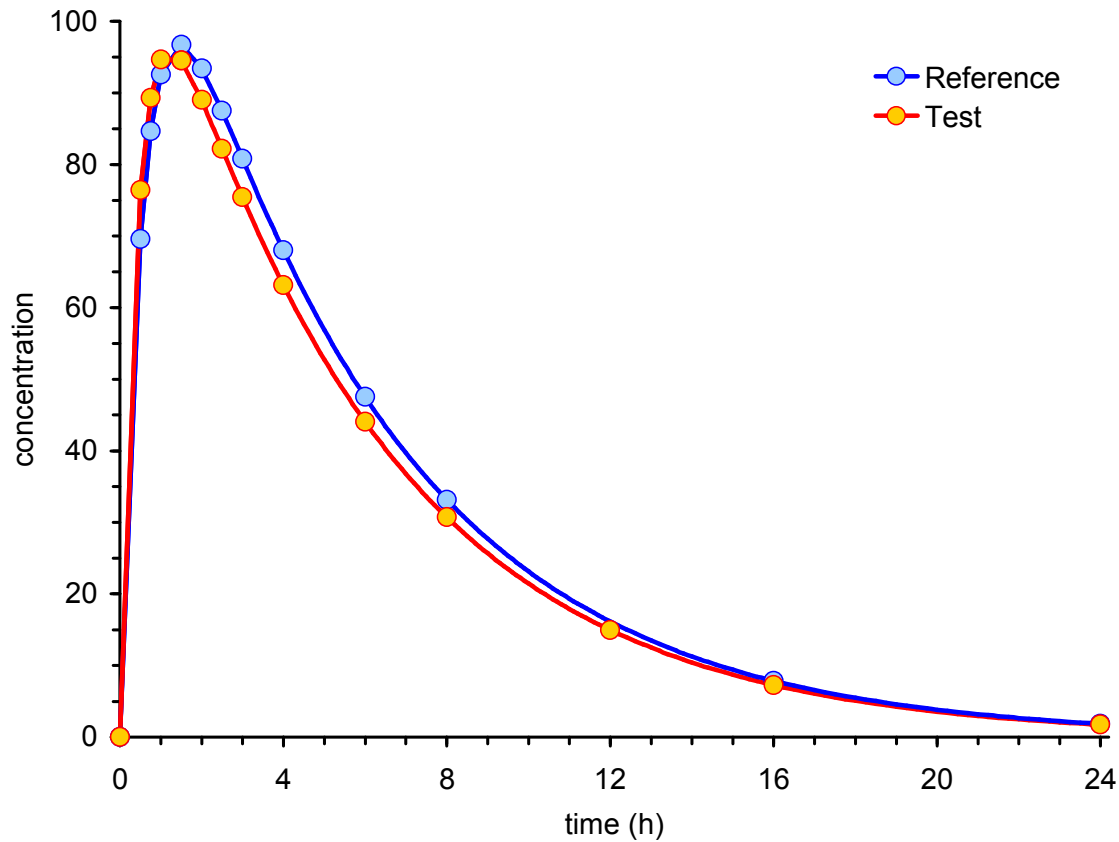
# $C_{max}$ | Problem & Solutions

- ## What if
  - samples in the area of $t_{max}$ are missing?
- ## Exclude the subject from the comparison of $C_{max}$
  - Power depends on the *CV* (coefficient of variation), the *GMR* (geometric mean ratio), and *n* (sample size) where the rank order of their influence on power is
    $GMR \gg CV > n$
  - Power will be compromised but to a much lesser degree than many people expect

- Recap: To obtain a reliable estimate of the apparent elimination $\lambda_z$, *at least* three samples required
  - The automatic algorithm based on maximizing $R^2_{adj}$ is known to be 'greedy' (*i.e.*, reaches for too early time points) and
    - has difficulties with 'flat' profiles (*e.g.*, ill-defined $C_{max}$ of controlled release products) and
    - regularly fails completely for multiphasic release products
  - Alternative: TTT method *
    - Implemented in the open source package <u>bear</u> for <u>R</u>
  - Visual inspection of fits by a pharmacokineticist (with optional correction) is mandatory in all methods

* Scheerans C, Derendorf H, Kloft C. *Proposal for a Standardised Identification of the Mono-Exponential Terminal Phase for Orally Administered Drugs.* Biopharm Drug Dispos. 2008;29(3):145–57. <u>doi:10.1002/bdd.596</u>.

# Statistical Evaluation

- **Parallel Designs**
  - One group is treated with the test formulation and another group with the reference
  - Quite common that – due to drop-outs – the data set of eligble subjects is imbalanced, *i.e.*, $n_1 \neq n_2$
  - Equal variances should never be assumed (FDA 2001)
    - Treatment effect might be biased and
    - patient's risk inflated
    - In some software (*e.g.*, Kinetica, ThothPro) either wrong calculation or not possible at all

| Subj. | Group 1 (T) | Group 2 (R) |
|-------|-------------|-------------|
| 1-13  | 100         | 110         |
| 2-14  | 103         | 113         |
| 3-15  | 80          | 96          |
| 4-16  | 110         | 90          |
| 5-17  | 78          | 111         |
| 6-18  | 87          | 68          |
| 7-19  | 116         | 111         |
| 8-20  | 99          | 93          |
| 9-21  | 122         | 93          |
| 10-22 | 82          | 82          |
| 11-23 | 68          | 96          |
| 12-24 | dropout     | 137         |
| $n$   | 11          | 12          |
| mean  | 95          | 100         |
| $s^2$ | 298         | 314         |
| $s$   | 17.3        | 17.7        |

# Statistical Evaluation

| Subj. | Group 1 (T) | ln (T) | Group 2 (R) | ln (R) |
|---|---|---|---|---|
| 1-13 | 100 | 4.605 | 110 | 4.700 |
| 2-14 | 103 | 4.635 | 113 | 4.727 |
| 3-15 | 80 | 4.382 | 96 | 4.564 |
| 4-16 | 110 | 4.700 | 90 | 4.500 |
| 5-17 | 78 | 4.357 | 111 | 4.710 |
| 6-18 | 87 | 4.466 | 68 | 4.220 |
| 7-19 | 116 | 4.754 | 111 | 4.710 |
| 8-20 | 99 | 4.595 | 93 | 4.533 |
| 9-21 | 122 | 4.804 | 93 | 4.533 |
| 10-22 | 82 | 4.407 | 82 | 4.407 |
| 11-23 | 68 | 4.220 | 96 | 4.564 |
| 12-24 | dropout | – | 137 | 4.920 |
| $n$ | 11 | 11 | 12 | 12 |
| mean | 95 | 4.539 | 100 | 4.591 |
| $s^2$ | 298 | 0.03418 | 314 | 0.03231 |
| $s$ | 17.3 | 0.1849 | 17.7 | 0.1798 |

Assuming equal variances

$$\nu = n_1 + n_2 - 2 = 21$$
$$t_{1-\alpha,21} = 1.7207$$

90% CI: 83.28% – 108.20%

Adjusting for unequal variances by Satterthwaite's degrees of freedom

$$\nu = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\dfrac{\left(s_1^2/n_1\right)^2}{n_1-1} + \dfrac{\left(s_2^2/n_2\right)^2}{n_2-1}} = 20.705$$
$$t_{1-\alpha,20.705} = 1.7219$$

90% CI: 83.26% – 108.23%

Minor difference in the CIs but only little imbalance in the data and variances quite similar. However, the simple *t*-test is always liberal, *i.e.*, compromises the patient's risk.

# Statistical Evaluation

- **Crossover Designs (2×2×2)**
  - Every subject is treated with both the test and the reference formulation
  - Subjects randomized to two sequences TR and RT
  - Treatment periods separated by washout
  - Potential period effects are accounted for in the analysis (mean out)
  - Evaluation by
    - Analysis of Variance (ANOVA) – WHO, EMA, …
    - Linear mixed effects model – FDA, Health Canada
    - Results are identical for balanced datasets (equal number of subjects in both sequences) and differ only slightly for imbalanced ones

# Statistical Evaluation

- ## Crossover Designs (2×2×2) – Example

| subject | T | R |
|---|---|---|
| 1 | 28.39 | 35.44 |
| 2 | 39.86 | 49.42 |
| 3 | 32.75 | 36.78 |
| 4 | 33.36 | 33.40 |
| 5 | 34.97 | 34.81 |
| 6 | 24.29 | 24.65 |
| 7 | 28.61 | 31.77 |
| 8 | 45.44 | 45.54 |
| 9 | 59.49 | 65.29 |
| 10 | 27.87 | 28.23 |
| 11 | 24.26 | 25.71 |
| 12 | 42.30 | 37.01 |

| sequence RT | | | sequence TR | | |
|---|---|---|---|---|---|
| subject | P I | P II | subject | P I | P II |
| 2 | 39.86 | 49.42 | 1 | 28.39 | 35.44 |
| 3 | 32.75 | 36.78 | 4 | 33.36 | 33.40 |
| 5 | 34.97 | 34.81 | 6 | 24.29 | 24.65 |
| 8 | 45.44 | 45.54 | 7 | 28.61 | 31.77 |
| 10 | 27.87 | 28.23 | 9 | 59.49 | 65.29 |
| 11 | 24.26 | 25.71 | 12 | 42.30 | 37.01 |

Ordered by treatment sequences (RT|TR)

ANOVA on log-transformed data $\rightarrow$

# Statistical Evaluation

- ## Crossover Designs (2×2×2) – Example cont'd

| Sequence | Period 1 | | Period 2 | | Sequence mean | |
|---|---|---|---|---|---|---|
| 1 | 1R = $X_{\cdot11}$ | 3.5103 | 1T = $X_{\cdot21}$ | 3.5768 | $X_{\cdot\cdot1}$ | 3.5436 |
| 2 | 2T = $X_{\cdot12}$ | 3.5380 | 2R = $X_{\cdot22}$ | 3.5883 | $X_{\cdot\cdot2}$ | 3.5631 |
| Period mean | $X_{\cdot1\cdot}$ | 3.5241 | $X_{\cdot2\cdot}$ | 3.5826 | $X_{\cdot\cdot\cdot}$ | 3.5533 |

| | | | | |
|---|---|---|---|---|
| RT = | $n_1$ = 6 | | | |
| TR = | $n_2$ = 6 | $1/n_1 + 1/n_2$ | 0.3333 | |
| *balanced* | n = 12 | $1/n$ 0.0833 | $n_1 + n_2 - 2$ | 10 |

**Analysis of Variance**

| Source of variation | df | SS | MS | F | P-value | CV |
|---|---|---|---|---|---|---|
| *Inter*-subjects | | | | | | |
| Carry-over | 1 | 0.00230 | 0.00230 | 0.0144 | 0.90679 | |
| Residuals | 10 | 1.59435 | 0.15943 | 29.4312 | 4.32E-6 | 28.29% |
| *Intra*-subjects | | | | | | |
| Direct drug | 1 | 0.00040 | 0.00040 | 0.0733 | 0.79210 | |
| Period | 1 | 0.02050 | 0.02050 | 3.7844 | 0.08036 | |
| Residuals | 10 | 0.05417 | 0.00542 | | | 7.37% |
| Total | 23 | 1.67172 | | | | |

$\delta_{ML}$ **1.0082** *MLE (maximum likelihood estimator) of Delta-ML*

$X_R$ 3.5493 *LS (least squares mean for the reference formulation)*     $\exp(X_R)$ 34.79

$X_T$ 3.5574 *LS (least squares mean for the test formulation)*     $\exp(X_T)$ 35.07

# Statistical Evaluation

- Crossover Designs (2×2×2) – Example cont'd

**Classical (Shortest) Confidence Interval**

± x rule:      20 [ 100 - x; 1 / (100 - x) ]

| | | | | |
|---|---|---|---|---|
| $\theta_L$  -0.2231 | $\theta_U$  +0.2231 | $\alpha$  0.0500 | p=1-2·$\alpha$  0.9000 |
| $\delta_L$  **80%** | $\delta_U$  **125%** | $t_{2 \cdot \alpha, df}$  1.8125 | |
| $L_1$  -0.0463 | $U_1$  0.0626 | *difference within Theta-L AND Theta-U; bioequivalent* | |
| $L_2$  **95.47%** | $U_2$  **106.46%** | ***difference within Delta-L AND Delta-U; bioequivalent*** | |

$\delta_{ML}$  ☞ **100.82%** ☞  *MLE; maximum likelihood estimator*

$\delta_{MVUE}$   100.77%      *MVUE; minimum variance unbiased estimator*

$\delta_{RM}$   100.98%      *RM; ratio of formulation means*

$\delta_{MIR}$   101.44%      *MIR; mean of individual subject ratios*

# Statistical Evaluation

- Interpreting ANOVA Tables – Example cont'd

**Analysis of Variance**

| Source of variation | df | SS | MS | F | P-value | CV |
|---|---|---|---|---|---|---|
| *Inter*-subjects | | | | | | |
| Carry-over | 1 | 0.00230 | 0.00230 | 0.0144 | 0.90679 | |
| Residuals | 10 | 1.59435 | 0.15943 | 29.4312 | 4.32E-6 | 28.29% |
| *Intra*-subjects | | | | | | |
| Direct drug | 1 | 0.00040 | 0.00040 | 0.0733 | 0.79210 | |
| Period | 1 | 0.02050 | 0.02050 | 3.7844 | 0.08036 | |
| Residuals | 10 | 0.05417 | 0.00542 | | | 7.37% |
| Total | 23 | 1.67172 | | | | |

Should not be tested: Design – washout!

Not surprising: different subjects!

Irrelevant: Significant value would only mean that 100% is not included in the 90% CI.

Irrelevant; both formulations affected to the same degree

$$CV_{within} = 100\sqrt{e^{MS_w} - 1}$$

# Statistical Evaluation

- ## Statistical significant ≠ clinically relevant
  - For any given T/R-ratio and variability one will get a significant treatment effect (in the ANOVA $p < 0.05$) if the sample size is only large enough
    - The confidence interval narrows with $\sqrt{N}$, *i.e.*, if one uses a four times larger sample size, the CI will be ~half as wide
    - If the CI does not include 100% any more, treatments will *statistically significant* differ
    - However, if the 90% CI is within the BE-limits, this difference is *clinically not relevant*

# Statistical Evaluation

- **General Procedure (all Designs)**
  - Based on the design set up a statistical model
  - Log-transform the PK metrics of interest
  - Calculate for T and R
    - Balance sequences:      Geometric mean
    - Imbalanced sequences: Adjusted mean (a.k.a. least squares mean)
  - Calculate the ratio of means
  - Calculate the 90% confidence interval (CI) around the ratio
  - The *width* of the CI depends on the variability observed in the study
  - The *location* of the CI depends on the observed test/reference-ratio

# Statistical Evaluation

- ## BE Assessment (all Designs)
  - Decision rules based on the CI and pre-specified BE-limits
    - CI entirely outside the BE-limits →
      <span style="color:red">Bioinequivalence proven</span>
    - CI overlaps the BE-limits (lies not entirely within the limits) →
      <span style="color:orange">Bioequivalence not proven</span> (indecisive)
    - CI entirely inside the BE-limits →
      <span style="color:green">Bioequivalence proven</span>
  - Methods for reference-scaling
    - The BE-limits depend on the $CV_{wR}$ observed in study
    - Only the method pre-specified in the protocol

# Statistical Evaluation