26 April 2023
EMA/CHMP/779867/2022
Committee for Medicines for Human Use (CHMP)

Overview of comments received on 'Tadalafil film-coated tablets 2.5 mg, 5 mg, 10 mg and 20 mg product-specific bioequivalence guidance' (EMA/CHMP/315234/2014 Rev.2)

Interested parties (organisations or individuals) that commented on the draft document as released for consultation.

Stakeholder no.	Name of organisation or individual
1	AESGP
2	BEBAC; Institute of Medical Statistics, University of Vienna, Austria
3	GAP S.A.
4	Krka, d. d., Novo mesto
5	Medicines for Europe

Official address Domenico Scarlattilaan 6 ● 1083 HS Amsterdam ● The Netherlands
Address for visits and deliveries Refer to www.ema.europa.eu/how-to-find-us
Send us a question Go to www.ema.europa.eu/contact **Telephone** +31 (0)88 781 6000 An agency of the European Union

1. General comments – overview

Stakeholder no.	General comment (if any)	Outcome (if applicable)
1	AESGP is grateful for the opportunity to comment on this draft revised guideline. 'Comparable median (≤ 20% difference) and range for T_{max}' is mentioned on the template. Showing a relative or percentage difference would imply a statistical approach based on ratios. In theory T_{max} is a continuous variable, but practically it is not, as the timepoints are pre-defined. Therefore, assumptions on the distribution for a statistical approach based on ratios are not fulfilled. Our proposal would be to apply a non-parametrical approach on differences and compare these results on a numerical manner with the point estimates (e.g. median). In the absence of scientific rational, we oppose a T_{max} of 20%. We would like first of all to understand the reasons for proposing a T_{max} in the first place; we would be then open to consider and discuss a proposed T_{max} which is scientifically grounded and justified based on efficacy and safety considerations.	**Not accepted.** A comparable median T_{max} is required for drugs where the onset of action is clinically relevant. Only the point estimates of T_{max} are compared according to the Guideline on the Investigation of Bioequivalence, whereas the demonstration of bioequivalence for the non-parametric 90% CI of T_{max} was required in the past. The revision of the PSBGL intends to clarify the regulatory expectations by defining an objective criterion to avoid arbitrations. Note that the requirements for the comparison of the rate of absorption for drugs where the onset of action is clinically relevant are being harmonised in ICH M13. It is agreed that more sampling times are needed to characterise more accurately T_{max}, but the present approach is the less demanding approach amongst those available. The present approach is not based on ratios. The present approach is the following: If the reference median T_{max} is at 1.5 h, 20% of 90 minutes is 18 minutes. Therefore, if the test product has a median of 1.75 h (i.e. 105 minutes), the difference of 15 minutes is acceptable.
2	The clarification what is "meant by 'comparable' T_{max}" is appreciated.	**Accepted.** Comparable is defined by the acceptance range, which has been defined, i.e. differences ≤ 20% (and more precisely specified within 80–125%) of the value of the reference median.

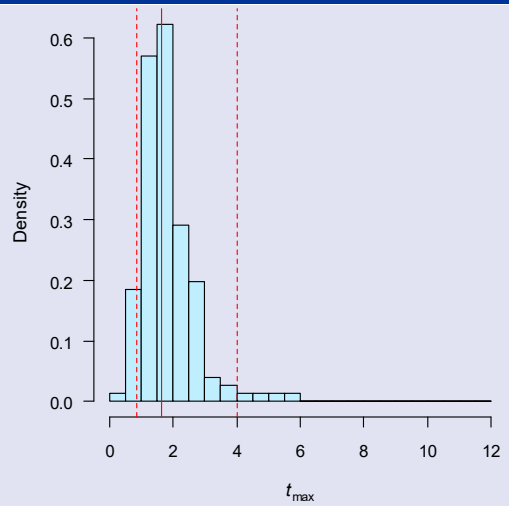
Stakeholder no.	General comment (if any)	Outcome (if applicable)
5	The specific comments to tadalafil draft product-specific bioequivalence guidance (EMA/CHMP/315234/2014 Rev.2*) provided below were also submitted to paracetamol (EMA/CHMP/356877/2022 Rev.1*) and ibuprofen (EMA/CHMP/356876/2017 Rev.1*) draft product-specific bioequivalence guidance(s) since all revisions concern the definition what is meant by 'comparable' T_{max} as an additional main pharmacokinetic variable in the bioequivalence assessment section of the guidance.	
5	The new proposal for acceptance criteria for median of T_{max} was introduced based on disagreement in registration procedure (IE/H/1132/001/DC) that involved ibuprofen formulations. In particular, referral for the Art. 10(1) application for an oral lyophilisate containing ibuprofen was triggered as it was considered by the objecting CMS that the bioequivalence requirements for T_{max} are not in line with the product-specific bioequivalence guideline (PSBGL) issued by PKWP. PKWP has been consulted during the referral procedure and confirmed that the presented T_{max} values are not to be considered "comparable", as mentioned in the PSBGL (CMDh minutes for the meeting on December 14 – 16, 2021, EMA/CMDh/89802/2022). Since this particular case represents a precedent for definition of general criteria, members of Medicines for Europe would appreciate if concrete data were made public. This would definitely contribute to transparency behind proposing a new criterion. Alternatively, example data sets of, in the PKWP point of view, comparable and non-comparable difference could be released, in order to permit stakeholder's review and further scientific discussion that must precede implementation of any new criteria affecting future submissions. These data shall include individual subject $T_{max(es)}$ along	**Partly accepted.** It is not considered necessary to make public any further data. Transparency on the criteria and how to apply it is given above in response to the first comment. It is not necessary to include individual subject $T_{max(es)}$ along with additional relevant information (e.g. period and sequence information in case of cross-over design) because the analysis is based on the medians of test and reference in a numerical subtraction.

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	with additional relevant information (e.g., period and sequence information in case of cross-over design).	

2. Specific comments on text

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
Table 'Requirements for bioequivalence demonstration' Line 'Bioequivalence assessment'	2	**Comments:** 'T' is the SI symbol for the absolute temperature. **Proposed change:** Use the correct SI symbol 't' for time, at least for consistency with the overarching guideline. [1] 1. EMA (CHMP). Guideline on the Investigation of Bioequivalence. CPMP/EWP/QWP/1401/98 Rev.1/Corr. London, 20 January 2010.	**Accepted.**
Table 'Requirements for bioequivalence demonstration' Line 'Bioequivalence assessment'	2	**Comment:** 'Comparable […] range for T_{max}'. Like the mean, the range has a breakdown point of zero, i.e. a *single* extreme value distorts the range. Hence, a *confirmatory* assessment of the range is not contained in the statistical toolbox. It must only be assessed in an *exploratory* data analysis. Let us consider three formulations (two tests T1, T2, and one reference R) in a study of an arbitrarily large [sic] sample size. All T_{max} values except one are identical: The sets of observed T_{max} values are R {1,…,1.25}, T1 {1,…,1.5}, T2 {1,…,1}. Their respective ranges are 0.25, 0.5, and 0. Are these ranges 'comparable', and if yes, why? If they are 'not comparable', why? Is T1 'worse' than R because its range is larger? Is T2 'better' than R because its range is smaller (actually zero)? Of course, such a comparison is absurd. Naturally, the medians are identical.	**Not accepted.** The wording of the current Guideline on the Investigation of Bioequivalence on this topic is difficult to implement: "*A statistical evaluation of T_{max} is not required. However, if rapid release is claimed to be clinically relevant and of importance for onset of action or is related to adverse events, there should be no apparent difference in median T_{max} and its variability between test and reference product*". The purpose of this updated PSBGL is to clarify how to assess or compare the medians with an objective acceptance range. The assessment of the range is more subjective. If all the values except one are the same, the ranges would be considered acceptable. Therefore, only if differences are evident and worse for the test product, the range could be used for a regulatory decision.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
Table 'Requirements for bioequivalence demonstration' Line 'Bioequivalence assessment'	2	**Comment:** The *true* T_{max} follows a continuous distribution indeed. Furthermore, it is on a ratio scale (i.e. with a true zero). However, due to the sampling schedule, the *observed* T_{max} gets discretized, i.e. results in data on an ordinal scale. The only [sic] allowed operations for ordinal data are addition, subtraction, and ranking. To be clear: Multiplication and division are *not* allowed. Hence, calculating a ratio (expressed as a percentage) is statistically flawed from the start. The distribution of observed T_{max} is skewed to the right, which "can be attributed to the asymmetry of the observed concentrations around the peak. The concentrations rise more steeply before the peak than they decline following the true maximum response. Consequently, it is more likely that large observed concentrations occur after than before the true peak time." [2] 2. Tóthfálusi L, Endrényi L. *Estimation of C_{max} and T_{max} in Populations After Single and Multiple Drug Administration.* J Pharmacokin Pharmacodyn. 2003; 30(5): 363–85. doi:10.1023/b:jopa.0000008159.97748.09. An example from our files; pooled IR data of seven studies:	**Not accepted.** The objective of the present review of the PSBGL is not to change the requirements of the existing Guideline on the Investigation of Bioequivalence, but to clarify how to interpret it. As the discussion of any difference in the context of the application is subjective, the present update of the PSBGL intends to define an objective criterion to avoid arbitrations. Regarding the comment on calculating the ratio of data on an ordinal scale is not an allowed operation. Hence, the '±20% difference in medians' criterion is statistically flawed, the ordinal scale is due to the discrete time schedule, whereas the continuous "true T_{max}" can be considered as the target of estimation. Hence, calculating a ratio as an estimation of the true T_{max} ratio still makes sense even if estimation may not be optimal due to the discrete sampling time points' estimation. Obviously, the denser the sampling schedule the more accurate the estimation will be, but for practical reasons the number of sampling time points is limited. However, as the '±20% difference in medians' criterion might still be considered as flawed since it violates the principle of symmetry (i.e. the requirement that test should be equivalent to reference if and only if reference is equivalent to test) it is therefore slightly modified (or more precisely specified) to an 80–125% rule.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		 Red line median, red dashed lines 2.5 and 97.5 percentiles. As expected, the distribution of T_{max} is heavily skewed to the right (skewness +0.771), thus confirming the theoretical considerations. [2] It must not be forgotten that comparative bioavailability of conventional PK metrics (*AUC*, C_{max},…) is based on a *clinically relevant difference* Δ, leading with the common 20% in a *multiplicative* statistical model to the BE-limits $\{\theta_1,\theta_2\}=\{100(1-\Delta),100(1-\Delta)^{-1}\}=\{80\%,125\%\}$. Consequently, a similar approach should be applied to T_{max}, i.e. if – and only if – clinically relevant, a certain Δ has to be pre-specified, which leads in an *additive* statistical model to the BE-limits $\{\theta_1,\theta_2\}=\{-\Delta,+\Delta\}$.	An acceptance range (delta) is pre-defined in this PSBGL for T_{max}, because T_{max} is compared only in those cases where it is clinically relevant for the onset of action. It is agreed that from an inferential/statistical point of view, the use of a non-parametric 90% confidence interval is more correct. But this correct statistical methodology is not implemented in the PSBGL because the PSBGL has to be in line with the overarching Guideline on the investigation of bioequivalence. The comparison of the medians does not intend to preserve the type 1 error but to exclude formulations with different onset of action. The definition of ≤ 20% (80–125%) as acceptance range intends not to reject products where T_{max} is not excessively rapid and the sampling time around T_{max} is very frequent. For example, if samples are taken every 5 minutes and T_{max} occurs after 2 h, a 10-minute difference in a non-adjacent sampling time is acceptable, but it would be rejected if the samples are required to be adjacent. This criterion reinforces the idea that sampling times around T_{max} should be frequent enough to characterise C_{max} appropriately. If T_{max} is expected

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Given the fact that data are discrete on an ordinal scale, the only valid statistical approach for comparing two formulations is by an appropriate nonparametric method. [3–8] 3. Hauschke D, Steinijans VW, Diletti E. *A distribution-free procedure for the statistical analysis of bioequivalence studies.* Int J Clin Pharm Ther Toxicol. 1990; 28(2): 72–8. PMID:2307548. 4. Basson RP, Cerimele BJ, DeSante KA, Howey DJ. *T_{max}: An Un-confounded Metric for Rate of Absorption in Single Dose Bio-equivalence Studies.* Pharm Res. 1996; 13(2): 324–8. doi:10.1023/A:1016019904520. 5. Basson RP, Ghosh A, Cerimele BJ, DeSante KA, Howey DC. *Why Rate of Absorption Inferences in Single Dose Bioequivalence Studies are Often Inappropriate.* Pharm Res. 1998; 15(2): 276–9. doi:10.1023/a:1011974803996. 6. Hauschke D, Steinijans V, Pigeot I. *Bioequivalence Studies in Drug Development.* Chichester: Wiley; 2007. p. 97–100. 7. Chow S-C, Liu J-p. *Design and Analysis of Bioavailability and Bioequivalence Studies*. Boca Raton: Chapman & Hall/CRC Press; 3rd ed. 2009. p. 109–19. 8. Jones B, Kenward MG. *Design and Analysis of Cross-Over Trials.* Boca Raton: Chapman & Hall/CRC Press; 3rd ed. 2015. p. 68–96. As an aside, a nonparametric test was recommended by the EM(E)A for 19 years and is currently recommended in Argentina, Japan, South Africa, and by the WHO. A statistical comparison of T_{max} was never – and is not – required by the FDA and Health Canada. The agency responded in [9] to a comment of a stakeholder asking for a specific Δ: "It is not possible to give an absolute	after 30 minutes, samples every 5 minutes are required. It is not the objective to take samples every 2-3 minutes, even if this is necessary for some orally inhaled products and it is known to be feasible. Regarding the comment on the tight sampling schedule, samples every 5 minutes are feasible. Obviously, the tighter the sampling schedule the powerful (and accurate) a statistical test will be. Nevertheless, and more important, the power of a statistical test (usually be performed using a confidence interval), and consequently the sample size needed, will depend on the requested equivalence range and significance level (the allowed type-1 error rate). Equivalence range could be wider than the range that is applied for point estimate. Also, the allowed type-1 error rate (or equivalently, the coverage probability of the confidence interval) may be less strict than for AUC and C_{max}. This would allow for assessing the consumers risk for T_{max} but on a different level than for AUC and C_{max}. Still an agreement on both, equivalence range and significance level to be used, may be difficult to achieve. It is agreed that assessing the consumer risk would require a statistical test corresponding to a confidence interval approach.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		value in minutes for the acceptance range of T_{max} because this depends on the result obtained for the T_{max} of the reference in the study under assessment. […] studies in fasted state submitted by different applicants have exhibited different T_{max} values, e.g. from 2.0 to 3.2 hours." 9. EMA (CHMP). *Overview of comments received on 'Tadalafil film-coated tablets 2.5 mg, 5 mg, 10 mg and 20 mg product-specific bioequivalence guidance'.* EMA/CHMP/644909/2017. 25 January 2018. This response was readily based on a misconception. Δ is the clinically relevant difference. If the agency considers it to be e.g. 30 minutes, then in the first case the limits would be 90 – 150 minutes and in the second 162 – 222 minutes. Nothing easier than that. In assessing BE of conventional PK metrics the observed values of the reference vary between studies as well. However, their absolute values are irrelevant because the confidence inclusion approach hinges on a clinically relevant Δ of 20%. Further down [9] stated "No specific statistical tool could be defined […] because the comparison is not based on any statistical test […] but simply on the numerical comparison of medians and range." Why could a 'specific statistical tool' not be recommended in this guidance but an ANOVA for conventional PK metrics in [1]? In the following we explored both the '±20% difference in medians' criterion as well as with the nonparametric CI inclusion approach, where $\theta_1 = -\Delta$ and $\theta_2 = +\Delta$ $H_0 : \mu_T - \mu_R \not\subset \{\theta_1, \theta_2\}$ vs $H_1 : \theta_1 < \mu_T - \mu_R < \theta_2$	It is agreed that the clinically relevant delta (acceptance range) should be fixed by the agency. It is considered that while the Hodges-Lehmann estimator is an adequate estimator to compare T_{max} of Test (generic) and Reference (innovator) products, it estimates the median difference as compared to the current approach of comparable median and range for T_{max} which estimates the difference in medians. The current approach has been a requirement of the ibuprofen product-specific guideline since 2018 and the present revision of the product specific guideline concerns better defining what is meant by comparable and not introducing a new method particularly one for which EMA experience in regulatory submissions is limited. Therefore, the continued use of the current approach is recommended until the BE requirements are updated with M13. The proposed ≤ 20% difference should be understood as 80-125% in order to be symmetrical. To conclude: • It is agreed that assessing the consumer risk would require a statistical test corresponding to a confidence interval approach. However, the guideline does not require the calculation of the non-parametric 90%CI because it would increase notably the required sample size.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		We simulated individual subject profiles of 24 subjects in 2,500 studies in a three-arm parallel design.* * Only for speed reasons. Runtime of a couple of hours on a work-station. Simulating a crossover design takes days. The absorption rate constants k_{01} of three formulations, i.e. R (reference), A (fast), and B (slow) were obtained by numerically solving $$\log_e\left(k_{01}\cdot t_{½}/\log_e(2)\right)/\left((k_{01}-\log_e(2)/t_{½}\right)-t_{max}=0$$ for $t_{½}$ = 17.5 h and T_{max} 2 h, 96 min, and 144 min, respectively. Elimination, fraction absorbed, and volume of distribution were identical. Error distributions were uniform for f (0.6–1), lognormal for V (CV 50%), k_{01} (CV 35%), k_{10} (CV 40%). Distribution of the analytical error was normal with a CV of 8% of the simulated concentration. The LLOQ was set to 5% of $C_{max(R)}$. The sampling schedule was every ten minutes until four hours, 4.25, 4.75, 6, 9, 12, 24, 36, 48, and 72 hours (34 time points). In the nonparametric test Δ was set to 24 min, mimicking the '±20% difference in medians' criterion. Since $\mu_A=\theta_1$ and $\mu_B=\theta_2$, the number of passing studies divided by the number of simulations represents the empiric Type I Error. As rightly stated [9] 'the comparison is not based on any statistical test'. Although the '±20% difference in medians' criterion is not a statistical test, one can expect for both test treatments an equal chance to pass or fail because $$\mu_A=0.8\times\mu_R\rightarrow\tilde{t}_{max(A)}<0.8\times\tilde{t}_{max(R)}\approx\tilde{t}_{max(A)}\geq0.8\times\tilde{t}_{max(R)}$$ as well as $$\mu_B=1.2\times\mu_R\rightarrow\tilde{t}_{max(B)}\leq1.2\times\tilde{t}_{max(R)}\approx\tilde{t}_{max(B)}>1.2\times\tilde{t}_{max(R)}$$	• The comparison of the medians is intended to exclude products with different onset of action, because a statistically sound method requires excessive sample size. • Samples every 5 minutes are feasible. • Asking for a non-parametric 90% CI is more restrictive. Although it is agreed that the non-parametric 90% CI for the T_{max} difference is more correct methodologically, its use was discarded by the Guideline on the Investigation of Bioequivalence and this PSBGL cannot implement it against the guideline. It is agreed that the clinically relevant delta (acceptance range) should be fixed by the agency. Specific equivalence ranges may indeed be discussed. Still, it appears useful to first establish a default range that could be adapted for specific substances. The definition of a clinically relevant acceptance range for each specific drug is not feasible and it is not in line with the Guideline on the investigation of bioequivalence. Requiring T_{max} as a primary PK metric in vivo is not inconsistent with the in vitro approach because when in vitro dissolution is used for a waiver of the in vivo study, it is assumed not only that C_{max}, and AUC will be equivalent but also T_{max}. In addition, the in vivo

		Confirming [2] and our observations of IR formulations, the distributions were positively skewed (R +0.626, A +0.772, B +0.599). The empiric Type I Errors were controlled (A vs R 0.0308, B vs R 0.0208; i.e. below the significance limit of the binomial test 0.0578). Surprisingly in the '±20% difference in medians' criterion passing-rates were substantially larger than the expected 50% (A 64.4%, B 60.1%). It is highly questionable, whether for a drug product with a T_{max} of two hours and onset of effect as early as 16 minutes [10] a Δ of 24 minutes has <u>any</u> clinical relevance at all. 10. Rosen RC, Padma-Nathan H, Shabsigh R, Saikali K, Watkins V, Pullman W. *Determining the earliest time within 30 minutes to erectogenic effect after tadalafil 10 and 20 mg: a multicenter, randomized, double-blind, placebo-controlled, at-home study.* J Sexual Med. 2004; 1: 193–200. doi:10.1111/j.1743-6109.2004.04028.x. Furthermore, sample size estimation would require subject simulations with an in-depth knowledge of not only the drug but also of the formulations (absorption rate constant, lag time). Whereas PK parameters might be in the public domain, their variances almost never are. It is a widespread misconception that the Wilcoxon signed-rank test (for paired samples) and the Mann–Whitney U test (for independent samples) compare medians. The former employs the Hodges-Lehmann estimator, whereas the latter compares the median of the difference between a sample from x and a sample from y. Both are permutation tests and thus, computationally intensive. Strictly speaking, they give unbiased estimates of a shift in location only if distributions are identical (though not necessarily symmetrical). However, in well-controlled studies this is likely the case. [3] In our simulations	approach is not considered an alternative approach in all settings, but only allowed in specific circumstances. When products have the same or similar T_{max}, it is expected that the sample size required to show equivalence in C_{max} will be able to provide an accurate estimation of T_{max}. Compliance with an arbitrary limit of 20% for the difference in medians is considered feasible and in line with the guideline on the investigation of bioequivalence. Obviously, the closer the assumed PK model to the data generating model the more precise the sample size estimation will be. However, sample size estimation is always based on assumptions. For specific active substances, it might be possible to assume a Population PK model that is reasonably close. The Guideline on the Investigation of Bioequivalence will be updated by the ICH in M13. The proposal could be considered in an updated version. The population median as a population parameter has no variability, since it is a fixed parameter. The empirical median as an estimation method is variable according to the sampling distribution, which can be described by the corresponding standard error of the median.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		distributions were similar (skewness +0.599 to +0.722). Recall that in parametric methods independent and identical distributions are assumed as well. Furthermore, in a crossover study evaluated by an ANOVA homoscedasticity (equal variances) is assumed. If these assumptions do not hold, the residual error is inflated, increasing the producer's risk – which is not a regulatory concern. The same is likely in nonparametric approaches. Alternatives not requiring identical distributions [11–13] have not been assessed for their operating characteristics in a BE-setting so far. 11. Brunner E, Munzel U. *The Nonparametric Behrens-Fisher Problem: Asymptotic Theory and a Small-Sample Approximation.* Biom. J. 2000; 42(1): 17–25. doi:10.1002/(SICI)1521-4036(200001)42:1%3C17::AID-BIMJ17%3E3.0.CO;2-U. 12. Neubert K, Brunner E. *A studentized permutation test for the non-parametric Behrens–Fisher problem.* Comput Stat Data Anal. 2007; 51(10): 5192–204. doi:10.1016/j.csda.2006.05.024. 13. Wilcox RA. *Introduction to Robust Estimation and Hypothesis Testing.* London: Academic Press; 4th ed. 2017. p. 192–8. To conclude: • Calculating the ratio of data on an ordinal scale is not an allowed operation. Hence, the '±20% difference in medians' criterion is statistically flawed. ○ Since it is not a valid statistical test, the consumer risk cannot be assessed. ○ It would require a tight sampling schedule, which is not realistic for products with an early T_{max}.	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		○ It is extremely restrictive and hence, would require prohibitively large sample sizes. • The confidence interval inclusion approach is based on a valid test for differences in T_{max}, and hence, controls the consumer risk. ○ The clinically relevant Δ should be fixed by the agency. ○ Sample size estimation requires *full* information of the PK of the drug / drug products and a suitable PK model in order to perform simulations. If this information is not available, strictly speaking the requirement "The number of subjects to be included in the study should be based on an appropriate sample size calculation" [1] cannot be fulfilled. Then a reasonably large pilot study has to be performed in order to establish a valid (Population) PK model. ○ It is an open question, which Δ might by clinically relevant for a drug product with such an early onset of effect. By the way, the statement "[…] if rapid release is claimed to be clinically relevant […] there should be no apparent difference in median T_{max} and its variability between test and reference product" in [1] deserves an update in its next revision as well. What might 'apparent' be? Furthermore, the median is a statistic (one of many *estimators* of location) and its value is an *estimate* (i.e. a certain number). It does not have a 'variability', only the sample has one.	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change: T_{max} should be compared by a nonparametric method. The 90% confidence interval should lie within ± *X* hours.* * The value of *X* to be stated in the guidance should be based on the clinically relevant Δ and depends on the PD property caused by the reference formulation. It can be at least one hour or – in our opinion probably better – a comparison not required at all.	
Line 21 Table/ Bioequivalence assessment	3	**Comment**: The draft guidance requires the following the assessment of bioequivalence: *"Comparable median (≤20% difference) and range for T_{max}"*. This requirement implies that T_{max} has a substantial effect on the efficacy of tadalafil. T_{max} is a poor predictor of any differences in the rate of absorption and an even worse one regarding the onset of action. No clear concentration-effect relationships have been established for any of the three phosphodiesterase-5 inhibitors (sildenafil, vardenafil and tadalafil)[1]. Tadalafil's efficacy in erectile dysfunction is characterized by a distinct nonlinear, saturable dose-response relationship that was characterized using an Emax model[2]. In an attempt to establish an effective dosage of tadalafil for continuous daily dosing the originator has evaluated all the existing evidence regarding the dose-response and concentration-response relationship for tadalafil[3]. The conclusion from this evaluation was that ***"although a direct***	**Not accepted.** See EMA Clinical Pharmacology Q&A 4.14 on bioequivalence requirements for tadalafil orodispersible tablet. In addition, note that the requirement of a similar T_{max} applies for tadalafil products indicated for erectile dysfunction. It does not apply to those products indicated for pulmonary arterial hypertension. T_{max} is used to ensure that the shape of the concentration time curve is similar, when the shape is similar, it can be expected that the time of onset of action will be also similar. Generics' acceptance range is defined arbitrarily to ensure a similar biopharmaceutical quality. Their acceptance range are not defined based on PK/PD relationships. When equivalence in biopharmaceutics quality has been demonstrated by means of rate and extent of

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		***correlation of plasma concentrations with efficacy has not been established, a total tadalafil plasma concentration of 55 ng/ml, which approximates 90% enzyme inhibition in vitro, constituted a reasonable pharmacodynamic target for clinical development with the aim of maintaining these concentrations throughout the dosing interval*.”** Based on this it can be concluded that the onset of action of tadalafil does not primarily depend on T_{max} (the time for achieving maximum concentration) but on the time needed to achieve concentrations above 55 ng/ml. This conclusion is supported by the dosage recommendation of the originator that the product “*may be taken at least 30 minutes prior to sexual activity” and that “Tadalafil demonstrated statistically significant improvement in erectile function and the ability to have successful sexual intercourse up to 36 hours following dosing”.* In support of this we would like to present real data from a bioequivalence trial performed for a generic tadalafil product under fed conditions. The mean concentration vs. time curves from this trial are presented in the following table. A graphical presentation of the efficacy threshold in relation to the mean concentration vs. time curves is presented in Figure 1.	absorption, a bridge can be made to preclinical tests and of clinical trials associated with the reference medicinal product. When differences in C_{max}, AUC_t and T_{max}, in case on set of action is important, are different between two test and reference product, the biopharmaceutical quality is not the same and per definition the product cannot be considered a generic.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome

Sample no. / Time [h]		Test		Reference	
		Mean	Std	Mean	Std
1	0.00	0.36	1.39	0.10	0.43
2	0.17	0.43	1.66	0.79	1.80
3	0.33	20.48	35.26	23.40	35.60
4	0.50	101.87	122.96	91.50	112.78
5	1.00	363.52	161.82	231.17	155.99
6	1.50	417.50	136.90	293.40	158.37
7	2.00	408.56	122.05	322.97	131.57
8	2.50	405.47	122.46	337.68	119.61
9	3.00	393.99	120.11	357.36	97.49
10	3.50	381.15	122.31	356.82	88.79
11	4.00	370.52	107.94	360.45	94.02
12	4.50	368.78	107.87	358.09	90.08
13	5.00	317.03	101.46	322.85	87.50
14	8.00	268.21	100.79	263.47	88.02
15	12.00	216.45	84.03	223.52	73.80
16	24.00	148.90	77.65	146.11	54.67
17	48.00	65.82	53.84	60.98	39.50
18	72.00	28.38	28.65	28.49	25.60

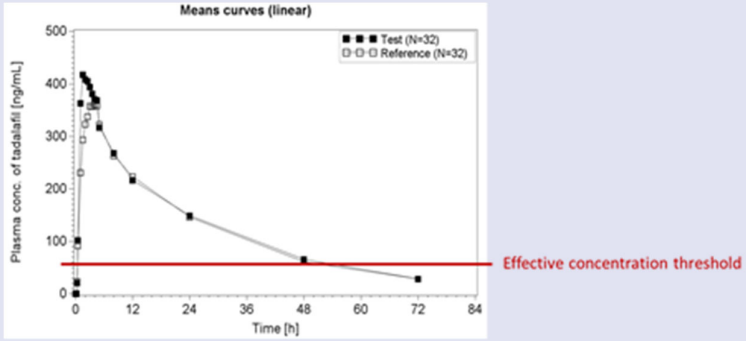


Figure 1. Mean (arithmetic) tadalafil plasma concentration-time profile (linear)

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		In the same bioequivalence trial, the 90% confidence intervals for the endpoints AUC_{0-72h} and C_{max} were as follows: **TADALAFIL (n=32)** **Variable** / **method** / **point estimator** / **confidence intervals** / **CV(%)** AUC(0-72h) (ratio test/reference) / ANOVA-log / 100.51% / 95.01% - 106.33% / 13.32% Cmax (ratio test/reference) / ANOVA-log / 111.96% / 105.82% - 118.45% / 13.35% The median T_{max} after administration of the test and the reference product was 1.5 and 2.5 hours, respectively. This means that according to the suggested Draft Guidance both products will not be regarded as bioequivalent only due to the median difference in T_{max} although both the onset and the duration of effect of the products are identical (Figure 1). 1. Mehrotra N, Gupta M, Kovar A, Meibohm B. The role of pharmacokinetics and pharmacodynamics in phosphodiesterase-5 inhibitor therapy. Int J Impot Res. 2007 May-Jun;19(3):253-64. 2. Staab A, Tillmann C, Forgue ST, Mackie A, Allerheiligen SR, Rapado J, Trocóniz IF. Population dose–response model for tadalafil in the treatment of male erectile dysfunction. Pharm Res. 2004 Aug;21(8):1463-70. 3. Wrishko R, Sorsaburu S, Wong D, Strawbridge A, McGill J. Safety, efficacy, and pharmacokinetic overview of low-dose daily administration of tadalafil. J Sex Med. 2009 Jul;6(7):2039-48.	
Line 21 Table/ Bioequivalence assessment	3	The estimation of T_{max} very much depends on the sampling schedule which has its own limitations. With the sampling schedule used in the trial described above (which fulfills all the requirements for a sampling schedule as defined in the Guideline on the Investigation of Bioequivalence.	**Partly accepted.** It is acknowledged that the estimation of T_{max} depends on the sampling schedule. As is indicated in the Guideline on the Investigation of Bioequivalence. CPMP/EWP/QWP/1401/98 Rev. 1/ Corr **) the

<table>
<tr><th>Line no.</th><th>Stakeholder no.</th><th>Comment and rationale; proposed changes</th><th>Outcome</th></tr>
<tr><td></td><td></td><td>CPMP/EWP/QWP/1401/98 Rev. 1/ Corr **) 30 minutes already represent 20% of the median T_{max} of the reference product.</td><td>sampling time of the study should be adapted to be sufficiently frequent around the expected T_{max} of the reference product. Indeed, in some cases even 10 min already represents 20% of the median T_{max} of the reference product. The schedule should depend on the expected T_{max} of the reference product.</td></tr>
<tr><td>Line 21
Table/
Bioequivalence assessment</td><td>3</td><td>The median depends only on the number of observations. This means that only one single value could decide on the median difference. A simple example is provided in the following table.
<table><tr><th>Volunteer</th><th>T_{max} Test (h)</th><th>T_{max} Reference (h)</th></tr><tr><td>1</td><td>0.5</td><td>0.5</td></tr><tr><td>2</td><td>0.5</td><td>0.5</td></tr><tr><td>3</td><td>1.0</td><td>1.0</td></tr><tr><td>4</td><td>1.0</td><td>1.0</td></tr><tr><td>5</td><td>1.5</td><td>2.0</td></tr><tr><td>6</td><td>2.0</td><td>2.0</td></tr><tr><td>7</td><td>2.0</td><td>2.0</td></tr><tr><td>8</td><td>2.5</td><td>2.5</td></tr><tr><td>9</td><td>2.5</td><td>2.5</td></tr></table>
The median T_{max} of the test product in the example above is 1.5 hours and that of the reference product is 2 hours. This means that the difference between both products is more than 20% due to one single value. If the suggested guidance is followed both products would not be regarded as bioequivalent although from the medical perspective such a conclusion would be at least questionable.</td><td>Not accepted.
This example is not considered valid because the sampling frequency is not frequent enough (see previous comment). Adjacent samples should differ less than 20%. Consequently, adjacent medians would be considered equivalent.</td></tr>
</table>

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		In conclusion we regard the wording suggested in the draft guidance referring to the use of T_{max} for the assessment of bioequivalence of products containing tadalafil as highly problematic and not scientifically justified. **Proposed change:** Such specific Tadalafil Bioequivalence Guidance should remain as is currently "Rev. 1" and not proceed to the "Rev. 2".	
Bioequivalence assessment, Main PK variables (in the table)	4	**Comment:** PK parameter T_{max} is listed as one of the main PK variables (together with C_{max} and AUC_{0-t}). In our opinion, the inclusion of T_{max} in the primary endpoint analysis is not justified for tadalafil and should be deleted. It is well known that PK parameter T_{max}: -is very sensitive parameter; -is highly variable and -has low statistical power. Furthermore, the sample size of bioequivalence study is not estimated to have enough statistical power for comparative T_{max} analysis. Thus, it is recommended to keep the requirements as presented in the guideline CPMP/EWP/QWP/1401/98 Rev.1/Corr**, that is the statistical evaluation of T_{max} should not be required unless rapid release is claimed to be clinically relevant and of importance for onset of action or is related to adverse events. This is not applicable for tadalafil and thus there is no need to require T_{max} as pivotal PK parameter in bioequivalence studies.	**Not accepted.** Even if it is highly variable and sensitive it is important to ensure that T_{max} is similar to confirm that the products will have a similar onset of action and are interchangeable. To avoid an excessive power, the 90% CI of T_{max} is not required, but only a comparison of the point estimates. For some indications of tadalafil the rapid onset of action is clinically relevant. Therefore, its T_{max} has to be comparable.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change: Main pharmacokinetic variables: C_{max}, AUC_{0-t}	
Bioequivalence assessment, 90% confidence interval (in the table)	4	**Comment:** In this section, the comparable median (≤ 20% difference) and range for T_{max} are proposed. T_{max} is categorical variable that can only take values based on the planned sampling scheme. Expected values for T_{max} are consequently confined to some preselected categories and therefore, median T_{max} depends more on the study design and less on the formulation of the drug. Because median and not the average value is reported, T_{max} of only one subject can determine the position at which T_{max} will be. If we have active ingredient with T_{max} 0.5 hours post-dose, difference of more than 6 minutes already exceed 20%. In a case of T_{max} at 1 hour post-dose, 20 % occurs at difference of 12 minutes and in case of median T_{max} at 2 hours post-dose, 20% difference corresponds to 24 minutes. We can conclude that with normal sampling schedule the difference of median T_{max} for just one sampling time already exceeds 20%. There are no literature data supporting that 20% difference in median T_{max} will be clinically significant for onset of action for tadalafil. Furthermore, it does not make sense to limit T_{max} in the direction of smaller values, since these values are said to provide faster onset of action of tadalafil.	**Not accepted.** The sampling times should be defined in order to ensure that a difference larger than 20% can be discarded. T_{max} is expected to occur in the same sampling time for test and reference for drugs where the onset of action is clinically relevant. As the T_{max} of only one subject can determine the position at which T_{max} will be, the 20% acceptance range gives some flexibility and do not punish those studies with frequent sampling schedules. The generic products are considered interchangeable in all potential patients because their biopharmaceutical quality is equivalent to that of the reference product. In those cases where the onset of action is clinically relevant T_{max} should be considered also to ensure that the biopharmaceutical quality is sufficiently similar. T_{max} needs to be considered to ensure that the shape of the concentration time curve is sufficiently similar. The 20% value is defined arbitrarily to ensure a similar biopharmaceutical quality. In the same way

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		For on-demand dosing of 10 or 20 mg tadalafil SmPC of tadalafil suggests taking the medication (at least 30 minutes) prior to anticipated sexual activity and with or without food. On the other hand, application of once daily tadalafil (2.5 or 5 mg) should be at approximately the same time of the day. Additionally, in the Pharmacokinetic properties it is explained, that the rate and extent of absorption are not influenced by food, thus tadalafil may be taken with or without food. Even time of dosing (morning versus evening after a single 10 mg administration) is said to have no clinical relevant effects on the rate and extent of absorption (1). The magnitude of the food effect on the pharmacokinetics of tadalafil was further investigated in literature. In the figure below there are presented fasting and fed T_{max} values for reference Cialis 20 mg from different studies. The overall median T_{max} from all the listed studies in fasting state is 2.25 hours and in fed state 3.67 hours. The difference between both conditions is 1.45 hours, which corresponds to 64.4% difference relative to fasting state. The smallest	that a 20% acceptance range is defined for the 90% CI of C_{max} and AUC, but in this case the assessment is not based on 90% CI. Even if the SmPC indicates that the intake can be with or without food, bioequivalence was not shown between fasted and fed state intake. The differences allowed for addressing the food effect are not acceptable for generics demonstrating bioequivalence.

Product	Fasting median T_{max} and range in hours	Fed median T_{max} and range in hours
Cialis 20 mg (2)	2.0 (0.75-8.0)	3.5 (1.67-6.0)
Cialis 20 mg (3)	2.0 (0.5-4.0)	2.5 (1.0-4.0)
Cialis 20 mg (4)	2.67 (0.67-4.5)	3.67 (1.67-8.0)
Cialis 20 mg (5)	2.67 (0.67-4.5)	3.67 (1.67-8.0)
Cialis 20 mg from in house data	2.25 (0.67-4.5)	3.75 (0.70-10.0)

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		difference of 0.5 hours was observed in food effect study, conducted by Forgue ST et al. It was concluded that there was no significant difference between T_{max} values between both conditions (3), but if we calculate the difference of median T_{max} in fasting and fed state, it is 25%. Studies were made on lower doses of Cialis as well. Difference between T_{max} of fasting and fed state for 2.5 mg tadalafil is 1.83 hours, which corresponds to 109.6% (4). Comparison of fasting in fed state of 5 mg tadalafil shows us that T_{max} in fed state is delayed for 0.9 hours, which corresponds to 45% difference vs. fasting state (2). According to literature data, tadalafil is unaffected by fatty meal as it does not have significant effect on the rate and extent of absorption, and can therefore be taken with meals, without a decrease in efficiency. In addition, the advantage of tadalafil is the elimination of the need to coordinate the timing of the meals around sexual activity (6, 7). Information, obtained from the studies shows us that difference in T_{max} between fasting and fed range from 25% to 109.6%. If we summarize, based on the fact that tadalafil can be taken with or without food and based on the high differences in T_{max} between fasting and fed states, up to 20% allowed difference in T_{max} will not have obvious and significant consequences and it is undoubtedly too strict criteria for T_{max} parameter of tadalafil. Many factors influence the in vivo performance of orally administered drugs and dosage forms. Apart from the properties of the dosage form and the drug, the physiological environment is of high importance (8). Physiological factors that lead to	To avoid the influence of external factors the studies are standardised and cross-over.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		variability in drug absorption can be the volume and the pH value of residual gastric contents, the motility of the stomach, the kinetic of gastric emptying of the co-administered water and the transit time of the drug product (9). Gastric emptying is generally accepted to be one of the most critical physiological process for oral drug delivery (8). Moreover, in bioequivalence studies in fasting conditions, drug is administered randomly relative to motility cycle and, therefore, we introduce this random variable, independent of dosage form, into classic bioequivalence studies (10). A drug will remain in the stomach for unknown lengths of time and consequently, T_{max} can significantly differ within and between subjects even under highly standardized conditions. T_{max} values can be highly distributed and even one subject can be the reason, why it comes to the differences in Median T_{max} between test and reference formulation. We can conclude, that up to 20% difference in T_{max} between two formulations is undoubtedly too strict criteria for T_{max} parameter of tadalafil. **Proposed change:** In the table, section 'Bioequivalence assessment', modify text as to following: 90% confidence interval: 80.00 – 125.00 % for AUC_{0-t} and C_{max}. Statistical evaluation of T_{max} is not required unless applicable in the context of the application, e.g. if rapid release is claimed to be clinically relevant. In that case, comparison of T_{max} should be based on non-parametric methods and should be applied to untransformed data. If comparison of T_{max} difference is still required, we propose assessment by non-	The comparison based on medians avoids the influence of outlier values.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		parametrical test and statistically significant difference further assessed for clinical significance, where it is evident that 1.5 hours difference in T_{max} (fasting vs. fed T_{max} difference for median T_{max}) is considered clinically not relevant. **References:** 1. CIALIS®2.5 mg, 5 mg, 10 mg and 20 mg film-coated tablets, SPC. Last update: 28.1.2021. Cited: 14.7.2022. Available from: https://www.medicines.org.uk/emc/product/7432/smpc#gref 2. https://file.wuxuwang.com/hma/NL_H_3543_002_PAR.pdf accessed: 7. 6. 2022 3. Forgue TS, Patterson BE, Bedding AW, Payne CD, Phillips DL, Wrishko RE, et al. Tadalafil pharmacokinetics in healthy subjects. Br J Clin Pharmacol. 2005; 61(3):280-8. 4. https://file.wuxuwang.com/hma/NL_H_3650_004_PAR.pdf accessed: 7. 6. 2022 5. https://www.geneesmiddeleninformatiebank.nl/pars/h122680.pdf accessed: 7. 6. 2022 6. Mehrotra N, Gupta M, Kovar A, Meibohm B. The role of pharmacokinetics and pharmacodynamics in phosphodiesterase-5 inhibitor therapy. Int J Impot Res. 2007; 19(3):253-64. 7. Coward RN, Carson CC. Tadalafil in the treatment of erectile dysfunction. Ther Clin Risk Manage. 2008; 4(6):1315-30. 8. Grimm M, Scholz E, Koziolek M, Kühn JP, Weitschies W. Gastric Water Emptying under Fed State Clinical Trial	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Conditions Is as Fast as under Fasted Conditions. Molecular Pharmaceutics. 2017; 14(12):4262-4271. 9. Grimm M, Koziolek M, Kuhn JP, Weitschies W. Interindividual and intraindividual variability of fasted state gastric fluid volume and gastric emptying of water. Eur J Pharm Biopharm. 2018; 127:309-17. 10. Hens B, Tsume Y, Bermejo M, Paixao P, Koenigsknecht MJ, Baker JR, et al. Low buffer capacity and alternating motility along the human gastrointestinal tract: Implications for in vivo dissolution and absorption of ionizable drugs. Mol Pharm. 2017; 14(12):4281-94.	
Table Requirements for bioequivalence demonstration (PKWP)/ Bioequivalence assessment	5	**Comment:** The draft guidance EMA/CHMP/315234/2014 Rev.2* is introducing a proposal for assessment of comparability of main additional pharmacokinetic variable T_{max}. More specifically, apart from the previously implemented requirement for a 'comparable median and range for T_{max}', newly, comparable median for T_{max} is to be concluded only if the difference between the test and reference median is less than or equal to 20%. While the motivation of PKWP to introduce acceptance criteria for T_{max} to conclude similarity in biopharmaceutical quality for generic medicines is understood, the current proposal is considered not acceptable for statistical and ethical reasons, as described in details in the below paragraphs. In line with the current version of EMA bioequivalence guideline (CPMP/EWP/QWP/1401/98 Rev.1/Corr), the sampling schedule should include frequent sampling around predicted T_{max} to provide a reliable estimate of peak exposure. However, the proposed 20% difference may easily lead to conclusion of non-	**Partly accepted.** The sampling times should be defined based on the expected T_{max} of the reference product. If it is 30 or 45 minutes, samples every 5 minutes should have been defined. For example, at 0.17, 0.33, 0.5, 0.58, 0.67, 0.75, 0.83, 0.92, 1, 1.25, 1.5, 1.75, 2, 2.5, 3, 4, 6, 8, 10, 12, 16, and 24 hours post-dose.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		comparability between formulations $T_{max(es)}$ even in cases where the calculated medians differ just by one sampling interval. For instance, in a study with sampling intervals of every 20 minutes, median achieved at 1.67 hours and 1.33 hours for test and reference, respectively, represents a difference of 26% (expressed as percentage of reference median, i.e. 100 × (1.67-1.33)/1.33 [%]). The situation becomes even more difficult for molecules with a shorter T_{max}, such as fast dissolving ibuprofens or paracetamol-containing products, where typically sampling intervals are more extensive in the first hour following the dosing. Here, sampling intervals of every 10 minutes for a product with an expected median of 0.5 hour means that a median difference in one sampling interval grossly fails the 20% acceptance criterion (observed difference would equal to 33%). Obviously, with median of 0.5 hour, one would have to sample at least every 6 minutes to satisfy the 20% difference. This would lead to extensive sampling intervals in the first hour after dosing (= 12 samples), since equidistant sampling intervals are typically required to achieve similar precision of C_{max} capture for majority of subjects. Not surprisingly, this is considered unrealistic due to the need of additional samples to describe the entire PK profile, logistical issues, but more importantly, due to excessive and unnecessary subject burden. Finally, there is no reasonable way of designing the study to decrease the sponsor risk or, increase power to pass this criterion, as it is feasible for other PK metrics such as C_{max} or AUC. The above examples are not only theoretical, but are based on real studies conducted by members of the association. A representative example is summarized in the	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		following paragraph; the study was conducted in a well-established CRO located in Canada (data available on request). A randomized, 2-period, 2-sequence, single-dose, cross-over bioequivalence study under fasting conditions in 26 volunteers was designed for a generic formulation containing 500 mg of paracetamol; sampling intervals were employed as following: (0-hour) and at 0.17, 0.33, 0.5, 0.67, 0.83, 1, 1.25, 1.5, 1.75, 2, 2.5, 3, 4, 6, 8, 10, 12, 16, and 24 hours post-dose. The resulting test-to-reference ratios along with the 90% confidence intervals for Cmax and $AUC_{(0-t)}$ were as following: 101.02 (91.69 - 111.31) and 100.80 (97.68 - 104.03), respectively. With respect to T_{max}, the test formulation displayed a median of 0.50 hours (min-max: 0.33-2.00 hours) and the reference displayed a median of 0.75 hours (min-max: 0.33-3.00 hours). The calculated medians from this study do not satisfy the proposed 20% criterion (difference of 33%) and thus might appear different, however, a statistical evaluation of within-subject (period) differences in T_{max} reveals otherwise (refer for details further below). In PK studies, concentrations are only taken typically at a set of predetermined times, and so T_{max} is an inherently discrete random variable (Patterson & Jones, 2006). While T_{max} is continuous in theory (Willavize et al., 2008), its distribution, either on the original scale (or on the log-scale), rarely follows a normal distribution (Chow & Liu, 2009). Consequently, statistical analysis of discrete variables like T_{max} requires the use of non-parametric (distribution-free) procedure. In fact, non-	Based on the present guideline on the investigation of bioequivalence the rapid release when onset of action is clinically relevant has to be assessed based on median T_{max} values. We cannot change the requirements of the guideline in this PSBGL, but to clarify the acceptance range. A larger sample size may be necessary. At least the present approach does not require that the complete non-parametric 90% confidence interval is contained within the 80-120% acceptance range. Only the point estimate should be within the 20% limits (80-125%). It is agreed that the proposed approach is not able to preserve the type 1 error. But this is the approach

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		parametric analysis was implemented in the earlier version of the EM(E)A bioequivalence guideline and is still applicable as per the current WHO guideline (WHO, 2017). Construction of non-parametric confidence interval in 2x2 cross-over designs is based on period differences (Hauschke et al., 1990). For the above study with paracetamol, the treatment difference (Hodges-Lehmann estimate) was – 0.125 hours (–7.5 minutes) along with 90%-confidence intervals (exact) ranging from –0.245 to 0.000 hours. The analysis detected no significant differences between test and reference (p=0.2234 for Wilcoxon-Mann-Whitney test; confidence interval includes zero). Clearly, this example illustrates that the newly proposed criterion concludes a difference where there is none based on statistical analysis appropriate for a discrete variable (and study design). Of note, the use of non-parametric analysis for T_{max} is not against general principles of the EMA bioequivalence guideline (CPMP/EWP/QWP/1401/98 Rev.1/Corr); non-parametric analysis is stated as not acceptable for analysis of PK parameters that are analysed following logarithmic transformation, i.e., applies to C_{max}, $AUC_{(0-t)}$ and/or $AUC_{(0-inf)}$. The difficulty in application of the new criterion may further be demonstrated by means of Monte Carlo simulation (e.g., by utilizing the sample function in R-software, R Core team, 2022). In this exercise, 26 values (sample size of the paracetamol study) were randomly sampled from population to obtain two sets of T_{max} values, one for test and one for reference. The population to sample from (for both products) exactly matched the T_{max} distribution observed in reality for the reference	defined in the Guideline on the Investigation of Bioequivalence. The sponsor should define the sampling times with enough frequency to ensure that a difference higher than 20% can be discarded. The protocol should predefine the methodology by simply considering the difference between medians. This approach is based on the present requirements of the Guideline on the Investigation of

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		product in the above paracetamol study. For each of the 100'000 simulation runs, the test and reference medians along with their percent difference was computed and proportion of studies passing the 20% difference was evaluated. The results revealed that only 50% of simulated studies passed the proposed criterion of less than or equal to 20% difference despite the fact that population medians for both products were absolutely identical. Based on real data, this simulation demonstrates that power of the newly proposed acceptance criterion is low. Moreover, being a decision procedure based on point estimate only, it is not likely that the type I error is adequately controlled. Concerning assessment of 'comparable range' for T_{max}, in the past, draft product-specific bioequivalence guidance(s) for paracetamol, ibuprofen, tadalafil or dimethylfumarate were commented by stakeholders in the sense that it is not clearly defined and it is questionable how it should be practically evaluated. Unfortunately, these comments were not adequately addressed by PKWP, moreover, newly proposed PSBG revisions maintain the same uncertainty. Since objective rules when 'simply the numerical comparison' (the term used by PKWP in overview of comments EMA/CHMP/644909/2017) would or would not conclude similarity are lacking, unclear acceptance criterion referred to as 'comparable' has no place in a modern guidance. As stated by the PKWP in the response to stakeholder comments (EMA/CHMP/729976/2017), '*the use of Tmax as*	Bioequivalence, since the PSBGL cannot define different approaches. The rate of absorption is considered relevant for onset for action. Therefore, as stated the T_{max} is not assessed with a statistical approach based on non-parametric 90% CI, but only with medians. 20% has been defined arbitrarily in the same way that 20% is used by default for C_{max} and AUC of all drugs, except HVDP and NTID. It could also be argued that the acceptance range should be defined per drug and even per dosage form for C_{max} and AUC, instead of using 80-125%. But as a measurement of biopharmaceutical quality a 20% acceptance range is used.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		pivotal variable is only applicable in certain situations. Unless the rate of absorption is important with regard to for instance efficacy, statistical evaluation of Tmax is not required.' Accordingly, this shall be implemented in the revised guidance text. **Proposed change:** In the table, section 'Bioequivalence assessment', modify text as to following: 90% confidence interval: 80.00 – 125.00% for AUC_{0-72h} and C_{max}. Comparison of T_{max} should be based on non-parametric methods and should be applied to untransformed data. **References:** • CMDh minutes, EMA/CMDh/89802/2022 • EMA guideline, CPMP/EWP/QWP/1401/98 Rev.1/Corr • Hauschke D et al. (1990). Int J Clin Pharmacol Ther Toxicol. 28(2): 72-8 • Chow SC & Liu JP (2009). 3rd edition, Chapman & Hall/CRC, Boca Raton • Overview of comments, EMA/CHMP/644909/2017 & EMA/CHMP/729976/2017 • Patterson S & Jones B (2006). Chapman & Hall/CRC, Boca Raton • R Core team (2022). R Foundation for Statistical Computing, Vienna, Austria • WHO (2017). WHO Technical Report Series, No. 1003, Annex 6 • Willavize SA & Morgenthien EA. (2008). Pharm Stat. 7(1): 9-19	