26 April 2023
EMA/758781/2022
Committee for Medicines for Human Use (CHMP)

Overview of comments received on 'Paracetamol oral use immediate release formulations product-specific bioequivalence guidance' (EMA/CHMP/356877/2022 Rev.1)

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	Angelini Pharma S.p.A.
3	BEBAC; Institute of Medical Statistics, University of Vienna, Austria
4	Krka, d. d., Novo mesto
5	Medicines for Europe
6	PAEDIATIS SAS - France

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. 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. Considering the short T_{max} of paracetamol fast acting formulations (0.40 – 0.88h), we need for example to sample every 5 minutes for the two first hours to improve the statistical comparison related to ±20% maximum difference which pause feasibility and ethical concerns. 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 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 (80.00–125.00%). 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.
3	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 with 80–125%) of the value of the reference median.

5	The specific comments to paracetamol draft product- specific bioequivalence guidance (EMA/CHMP/356877/2022 Rev.1*) provided below were also submitted to ibuprofen (EMA/CHMP/356876/2017 Rev.1*) and tadalafil (EMA/CHMP/315234/2014 Rev.2*) draft product-specific 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 with additional relevant information (e.g., period and sequence information in case of cross-over design).	**Partly accepted.** It is not considered necessary to make public any further data. Transparency on the criteria and how to apply them 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.

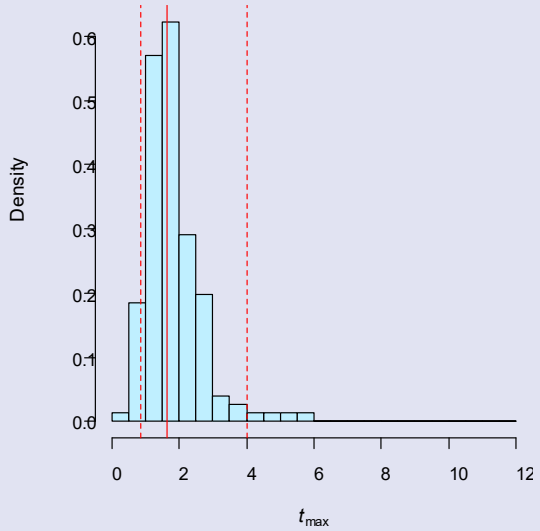
6	The specific comments to paracetamol (EMA/CHMP/356877/2022 Rev.1*) draft product-specific guidance(s) and specifically on usefulness of T_{max} as an additional main pharmacokinetic variable in the bioequivalence assessment section of the guidance.	**Noted**

2. Specific comments on text

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
20 "Bioequivalence assessment" header	2	**Comments:** The updated product specific bioequivalence guideline (PSBGL) for Paracetamol proposes the introduction of the definition of comparable median for T_{max} as ≤ 20% difference, still recommending the comparability of the range of T_{max}. Paracetamol is rapidly absorbed when administered orally, with T_{max} generally attained by 1.5 h after administration of solid formulations. However, T_{max} is, by definition, the time to reach maximum concentration (C_{max}), not the time of onset of drug efficacy. To account for the evaluation and the comparison of the actual onset of action among different paracetamol formulations, higher relevance should be given to the time required to achieve acknowledged active plasma concentrations (i.e., T_{onset}) for the proposed indication. Relevant scientific literature should be considered as a suitable basis to be used as a reference for the scope. For instance, Gibb and Anderson (2007) report paracetamol to be effective as antipyretic at plasma concentrations of 10–20 mg/l and to be (probably) effective as analgesic at about 5-20 mg/l. These and further available sound references could be used to identify univocal indication-specific thresholds/ranges for the computation of T_{onset}, to be compared among the different paracetamol formulations (especially for the fast-acting ones) to support T_{max} evaluation.	**Not accepted.** It is agreed that T_{max} and T_{onset} are different, but it is considered easier to compare T_{max} than to compare the time when minimum effective concentrations are achieved. If the median T_{max} values are similar and C_{max} and AUC are equivalent, the shape of the C-T curve is considered sufficiently similar to ensure that the time of onset of action will be equivalent.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
20 "Bioequivalence assessment" header	2	**Comments:** The updated product specific bioequivalence guideline (PSBGL) for Paracetamol sets as comparable median T_{max} values with a difference ≤ 20%, with no clear definition of how to evaluate the comparability of the ranges of T_{max}. From a statistical point of view such an evaluation would hardly be sound, being guided by subjective assessments. More precise criteria (e.g., subjects' distribution within the range) would be helpful for the scope.	**Not accepted.** 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.
Table 'Requirements for bioequivalence demonstration' Line 'Bioequivalence assessment'	3	**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'	3	**Comments:** '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	**Not accepted.** The wording of the current guideline on the investigation of bioequivalence in 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*".

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		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. **Proposed change:** ~~Comparable […] range for T_{max}.~~	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.
Table 'Require-ments for bio-equivalence demonstration' Line 'Bioequivalence assessment'	3	**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.	**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.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		An example from our files; pooled IR data of seven studies:  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 $\{\theta_1,\theta_2\}=\{-\Delta,+\Delta\}$.	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. 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.

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 Unconfounded Metric for Rate of Absorption in Single Dose Bioequivalence 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 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 taking 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 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

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		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 $\theta_2 = +\Delta$ $H_0 : \mu_T - \mu_R \subset \{\theta_1, \theta_2\}$ vs $H_1 : \theta_1 < \mu_T - \mu_R < \theta_2$ 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 $\log_e(k_{01} \cdot t_{½} / \log_e(2)) / ((k_{01} - \log_e(2) / t_{½})) - t_{max} = 0$ for k_{01} with $t_{½}$ = 1.8 h and T_{max} 75 min, 60 min, and 90 min, respectively. Elimination, fraction absorbed, and volume of distribution were identical. Error distributions were uniform for f (0.7–0.9), lognormal for V (CV 50%), k_{01} (CV 35%), k_{10} (CV 30%). Distribution of the analytical error was normal with a CV of 5% of the simulated concentration. The LLOQ was set to 5% of $C_{max(R)}$. The sampling schedule was 5 minutes, every ten minutes until 2.5 hours, 2.75, 3, 3.25, 3.5, 4, 5, 6, 8, 12, 16, and 24 hours (28 time points). In the nonparametric test Δ was set to 15 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. The '±20% difference in medians' criterion is not a statistical test. However, one can expect for both test treatments an equal chance to pass or fail because	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 the equivalence range and significance level to be used may be difficult to achieve.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		$\mu_A = 0.8 \times \mu_R \rightarrow \tilde{t}_{max(A)} < 0.8 \times \tilde{t}_{max(R)} \approx \tilde{t}_{max(A)} \geq 0.8 \times \tilde{t}_{max(R)}$ as well as $\mu_B = 1.2 \times \mu_R \rightarrow \tilde{t}_{max(B)} \leq 1.2 \times \tilde{t}_{max(R)} \approx \tilde{t}_{max(B)} > 1.2 \times \tilde{t}_{max(R)}$ Confirming [2] and our observations of IR formulations, the distributions were positively skewed (R +0.704, A +0.749, B +0.652). The empiric Type I Errors were controlled (A vs R 0.0292, B vs R 0.0328; i.e. below the significance limit of the binomial test 0.0578). Surprisingly in the '±20% difference in medians' criterion passing-rates were larger than the expected 50% (A 54.0%, B 54.6%). It is highly questionable, whether for a drug product with a T_{max} of 75 minutes a Δ of 15 minutes has any clinical relevance. For an ODT reference with a T_{max} of 30 minutes a Δ of six minutes is practically unachievable even with an logistically unrealistic sampling every two minutes. 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	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. Regarding the comment on the inconsistency with the in vitro approach, we do not agree necessarily since the in vivo approach is not considered an alternative approach in all settings but only allowed in specific circumstances. To conclude: • It is agreed that assessing the consumer risk would require a statistical test corresponding to a confidence interval approach. However, the

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		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 distributions were similar (skewness +0.652 to +0.749). 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 [9–11] have not been assessed for their operating characteristics in a BE-setting so far. 9. 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. 10. 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. 11. Wilcox RA. *Introduction to Robust Estimation and Hypothesis Testing.* London: Academic Press; 4th ed. 2017. p. 192–8.	guideline does not require the calculation of the non-parametric 90% CI because it would increase notably the required sample size. • 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

		Is somewhat surprising to require an additional PK metric for a drug which was introduced in clinical practice in 1883 and is the most commonly taken analgesic worldwide. Furthermore, a BCS-based biowaiver is acceptable as an alternative to the *in vivo* approach. Paracetamol is covered in one of the FIP's biowaiver monographs. [12] 12. Kalantzi L, Reppas C, Dressman JB, Amidon GL, Junginger HE, Midha KK, Shah VP, Stavchansky SA, Barends DM. *Biowaiver Monographs for Immediate Release Solid Oral Dosage Forms: Acetaminophen (Paracetamol).* J Pharm Sci. 2006; 95(1): 4–14. doi:10.1002/jps.20477. In other words, in the *in vitro* approach it would be readily acceptable to assess the risk of bioinequivalence based on studies in the public domain without T_{max} as a primary PK metric. Paracetamol was considered a 'model drug' for influence of gastric emptying on absorption. [13] Hence, an observed difference in T_{max} values is to a great part not caused by intrinsic properties of formulations themselves but a nuisance, confounding their true release characteristics. 13. Willems M, Quartero AO, Numans ME. *How Useful Is Paracetamol Absorption as a Marker of Gastric Emptying? A Systematic Literature Study.* Dig Dis Sci. 2001; 46(10): 2256–62. doi:10.1023/a:1011935603893. 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}.	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 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,

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 an early onset of effect. ○ Requiring T_{max} as a primary PK metric in the *in vivo* approach is inconsistent with the acceptable risk assessment in the *in vitro* approach. ○ Differences in T_{max} do not directly reflect release characteristics of formulations but are confounded by gastric emptying times. 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*	which can be described by the corresponding standard error of the median.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		(i.e. a certain number). It does not have a 'variability', only the sample has one. **Proposed change:** ~~Comparable median (≤ 20% difference) […] for T_{max}.~~ t_{max} should be compared by a nonparametric method. The 90% confidence interval should lie within ± *X* minutes.* * 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. However, for consistency with the in vitro approach and the confounding by gastric emptying we suggest to remove the comparison of T_{max} completely.	
Table Requirements for bioequivalence demonstration (PKWP) / Bioequivalence assessment	6	**Comment:** The draft guidance EMA/CHMP/356877/2022 Rev.1*) is introducing in place of 'comparable median and range for T_{max}', a new condition: equivalence of T_{max} is to be concluded only if the difference between the test and reference median is less than or equal to 20%. This approach could be understood for products in which the T_{max} is long and not solely dependent of physiological factors such as gastric emptying. Paracetamol is currently classified as a class I drug and its absorption dependent of the gastric emptying (Sanaka, Willems). In other words, for IR formulation T_{max} is highly dependent of the gastric emptying. In this case any factor which modifies slightly the gastric emptying time such as stress	**Not accepted.** The formulation may cause gastric emptying to differ due to excipients or that dissolution is not complete before gastric emptying. It is not agreed that for all paracetamol products T_{max} only depends on gastric emptying.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		will impact T_{max} of paracetamol. This will increase the variability of the results. Sanaka M, Koike Y, Yamamoto T, Mineshita S, Yamaoka S, Hirama S, Tanaka H, Kuyama Y, Yamanaka M. A reliable and convenient parameter of the rate of paracetamol absorption to measure gastric emptying rate of liquids. Int J Clin Pharmacol Ther. 1997 Nov;35(11):509-13. PMID: 9401832. Willems M, Quartero AO, Numans ME. How useful is paracetamol absorption as a marker of gastric emptying? A systematic literature study. Dig Dis Sci. 2001 Oct;46(10):2256-62. doi: 10.1023/a:1011935603893. PMID: 11680606. As a consequence of the new requirements the sampling schedule should include frequent sampling around expected T_{max} to provide data allowing a maximum difference of 20% between medians. Literature indicated, for various solid formulation of Paracetamol, peak plasma concentrations within 0.17–1.2 h postdosing (Kalantzi). A dispersion has observed T_{max} of 0.5h as median and range of 0.25 to 2h (García Aguirre), median of 0.42 h or 0.75h ranges of 0.17-1.50 and 0.33-4h (Portoles). Values of T_{max} for tablets is around 0.5h median values reported between 0.5 and 2h with a frequent value of 0.5h (Dubray). Kalantzi L, Reppas C, Dressman JB, Amidon GL, Junginger HE, Midha KK, Shah VP, Stavchansky SA, Barends DM. Biowaiver monographs for immediate release solid oral dosage forms: acetaminophen (paracetamol). J Pharm Sci. 2006 Jan;95(1):4-14. doi: 10.1002/jps.20477. PMID: 16307451. García Aguirre L, Bohorquez Nassar C, Ruiz Olmedo I, Dennie L, Medina Nolasco AG. Bioequivalence Evaluation of Two Oral Formulations of Acetaminophen in Healthy Subjects: Results From a Randomized,	It is agreed that frequent sampling will be necessary for a proper characterisation of T_{max}, but the expected median should be known, and it is not expected that samples every 5 minutes are taken in range from 0.5 h to 2 h.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Single-Blind, Crossover Study. Clin Pharmacol Drug Dev. 2019 Jan;8(1):9-15. doi: 10.1002/cpdd.469. Epub 2018 May 11. PMID: 29749718; PMCID: PMC6667897. Dubray C, Maincent P, Milon JY. From the pharmaceutical to the clinical: the case for effervescent paracetamol in pain management. A narrative review. Curr Med Res Opin. 2021 Jun;37(6):1039-1048. doi: 10.1080/03007995.2021.1902297. Epub 2021 Apr 5. PMID: 33819115. Portolés A, Puerro M, Terleira A, Rodríguez A, Caturla MC, Fernández N, Vargas E. A new high-absorption-rate Paracetamol 500-mg formulation: a comparative bioavailability study in healthy volunteers. Curr Ther Res Clin Exp. 2003 Jul;64(7):401-11. doi: 10.1016/S0011-393X(03)00110-3. PMID: 24944391; PMCID: PMC4053016. Taking a 0.5h (30 minutes) as a median T_{max} and the range observed of 0.25 to 2h, a 20% difference in the median is of 6 minutes. The sampling time must be able to show a possible difference less than 20%: 0.1h (6 minutes). As the median is corresponding to observed values (or a mean of two of them in case of even number of observations). That would lead to extremely frequent samplings in zone of T_{max} (range 0.25-2h). That would be questionable for number of points, possibility to handle very frequent sampling, and total volume of blood within the study. As the range is between 0.25 and 2h that would lead to sample every 5 minutes for more than 2h i.e. 24 samples in 2h if T_{max} must be caught for all subjects! Another option is to have a low number of samples, even if real T_{max} and Cmax are not caught, in order to be similar i.e., to have a sample every 0.25h or even less frequent up to 2h.	The option of infrequent sampling is not acceptable. Products with different intake instructions are not generics, but hybrids.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		That lead to another point a 20% median difference on T_{max}, when this latest is early (i.e. 0.5h), will lead to acceptance limits of C_{max} that would be, in practice, tighter than official limits. Based on simulations; a 20% median difference for T_{max} observed at 0.5h leads only to 3% or even less of differences in mean C_{max} values! Another aspect is the possibility to put on the market ODT formulation with and without water. Kambayashi (2019) studied the transit time of ODTs with and without water. ODTs containing soluble API (such as paracetamol) have a rapid gastric emptying median time 0.2h (12 minutes) the concomitant intake of water decreases the inter-individual variability significantly: the interquartile range is 0.2h without water against 0.1h with an intake with water. There is therefore no difference in the median gastric emptying time as soon as it begins between ODT with and without water, however the lag time is different. For ODTs containing soluble products, the lag of gastric emptying is negligible in case of concomitant water intake and is longer (estimated median 0.15h) and more variable without water intake (interquartile range 0.2h, without water against 0.1h taken with water). So the intake without water would lead to a T_{max} delayed by 0.15h (9 minutes) at least compared to the median T_{max} of 0.5h when taken with water leading to impossibility to conclude to equivalence as difference is greater than 20%. That leads to the conclusion that based on the current draft guidance it will not be possible to prove equivalence of intake	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		with and without water based on the current proposed T_{max} criteria. Kambayashi A, Sako K, Kondo H. Characterization of the buccal and gastric transit of orally disintegrating tablets in humans using gamma scintigraphy. Int J Pharm. 2020; 576:118937. Paracetamol could be subjected to BCS based biowaiver for solid oral dosage form as it is classified as a BCS class I, if composition requirements and dissolutions fulfil the criteria. For this latest either • more than 85% for both test and reference (for example 86% and 100%) in less than 15 minutes are observed and biowaiver could be directly granted irrespective of the impact on T_{max}. • more than 85% is dissolved in 30 minutes but not in 15 minutes associated with a successful F2 test, i.e. for the limit of F2 close to 50% that leads to 10% difference in dissolution curves, and biowaiver is granted irrespective of possible impact on T_{max}. That indicated that the possible differences on T_{max} in vivo are not expected to be of importance in term of rate of absorption. Paracetamol oral use immediate release formulations product-specific bioequivalence guidance" EMA/CHMP/356877/2022 Rev 1* **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 only required if a faster or slower release is claimed to be clinically relevant.	It is agreed that if the conditions for a BCS biowaiver are met the T_{max} is assumed to be similar.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
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, inclusion of T_{max} in the primary endpoint analysis is not justified for paracetamol 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 T_{max} comparative 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 should be the case of special formulations with the claim of extremely rapid release (i.e. faster that the release of standard IR formulations) which is further claimed (with clinical studies) to be clinically relevant (i.e. faster onset of action in comparison to standard IR formulations). Furthermore, as paracetamol is BCS class I drug substance, in vitro approach based on a BCS biowaiver can be conducted instead of in vivo studies. In that case, the only concerning factor is that 85 %	**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. As analgesic and antipyretic drug, onset of action is clinically relevant. Not only > 85% of paracetamol has to be dissolved in 30 min, but the profiles also have to differ less than 10%. This ensures that T_{max} will not differ. If

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		of the drug (from test and reference formulation) must be dissolved in less than 30 minutes. When in vivo studies may be waived, T_{max} comparison in such cases remains unknown. Thus, it is not justified to include T_{max} as pivotal PK parameter in bioequivalence studies for paracetamol. **Proposed change:** Main pharmacokinetic variables: C_{max}, AUC_{0-t}	the reference dissolves in 15 minutes, the test also has to dissolve in 15 minutes.
Bioequivalence assessment, 90% confidence interval (in the table)	4	**Comment:** In this section, the comparable median (≤ 20 % difference) and range for T_{max} is 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. Paracetamol is an active ingredient with very rapid T_{max} also for standard IR formulations. 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} of about 30 minutes 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	**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% (80–125%) acceptance range gives some flexibility and do not punish those studies with frequent sampling schedules.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		normal sampling schedule the difference of median T_{max} for just one sampling time already exceeds 20 %. In case of paracetamol, in standard IR formulations T_{max} is usually expected to be between 0.5 and 1 hour post-dose, which would mean sampling schedule would have to be extremely intense in the first hour to be able to comply with the proposed requirements for T_{max}, which would cause not only logistic challenges but would also pose unnecessary and unjustified ethical burden to the subject. Moreover, there is lack of PKPD analysis of paracetamol and as per our knowledge there is no known clinical studies or literature data proving that more than 20 % difference between two formulations (which corresponds to 6 minutes at median T_{max} of 0,5 hours and 12 minutes at median T_{max} of 1 hour) will be clinically significant for onset of action or analgesic effect. Moreover, even if the correlation between T_{max} and onset of action would be showed in clinical studies, it would not make sense to limit T_{max} in the direction of smaller values, since these values would provide a faster onset of action of analgesics, i.e. a faster analgesic effect. T_{max} of paracetamol depends on the pharmaceutical forms; oral solutions are absorbed more rapidly than ordinary tablets and in case of effervescent tablets of paracetamol containing sodium bicarbonate, of which prokinetic action promotes gastric emptying and thus the arrival of the drug in the small intestine, the greatest C_{max} and shortest T_{max} is provided (1). According to different PK characteristics (T_{max} and C_{max}) of different pharmaceutical forms of paracetamol, they should be treated individually case-by-case if differences are showed to be clinically relevant for onset of action or analgesic effect.	This example illustrates that when T_{max} is expected at 0.5 h, sampling times should be every 5 minutes around T_{max}. If T_{max} is expected 1h after dosing, samples should be taken every 10 minutes around T_{max}. And samples every 20 minutes around T_{max} should be taken if T_{max} is expected 2 h after dosing. For fast onset of action, the drug should be taken on empty stomach. Therefore, the SmPC recognises that the T_{max} difference caused by the food intake is clinically relevant in the case of paracetamol.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		The relevance of the newly proposed allowed difference of 20% in median T_{max} value was further evaluated in the light of food effect of the concerned drugs. As per SmPCs of different paracetamol products in most cases, there are no recommendations for the administration of paracetamol with regard to food intake. However, one SmPC of Panadol 500 mg film-coated tablets adds that food affects T_{max}, however with no influence on its effectiveness and can be taken with or without food and SmPC of Lekadol states, that food affects absorption of the drug and for faster effect suggest taking the drug 1 hour before a meal (2, 3). Furthermore, there are no special recommendations in these SmPCs on how the drug should be taken in regard to treatment of acute or chronic pain. As the SmPCs in most cases do not define the magnitude of the food effect on the pharmacokinetics of paracetamol, a literature review was performed. Food does not affect its bioavailability as measured by AUC but it reduces C_{max} and considerably affects T_{max}. Listed data are supported by results of studies: generally, median T_{max} values of fasting state ranged from 0.33 to 0.81 hours, meanwhile T_{max} values of fed state ranged from 1.25 to 2.2 hours. Increase in T_{max} under fed condition relative to the fast condition ranged from 66.7 % to 354.5 % (from 0.81 to 1.71 hours and from 0.69 to 2.2 hours in study by Divoll et al. (109.9 % and 218.8 % difference), from 0.33 to 1.5 hours in study by Forrest et al. (354.5 % difference) and from 0.75 to 1.25 in study by Stillings et al. (66.7 % difference)). Furthermore, fed C_{max} was only between 50.63 % and 62.4 % of fasting C_{max} (4-7). Despite the considerable effect of food on the pharmacokinetics of paracetamol, according to SmPCs it can be taken independently	To avoid the influence of external factors the studies are standardised and cross-over. The comparison based on medians avoids the influence of outlier values.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		of the meal intake. As the literature data shows, the delay in T_{max} between fasting and fed state obviously exceed stated 20 %, with values ranging from 66.7 % to 354.5 %. Therefore one can conclude, that up to 20 % allowed difference in median T_{max} between two formulations is undoubtedly too strict criteria for T_{max} parameter of standard IR paracetamol formulations. 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 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). For highly permeable drugs (BCS I) like paracetamol, gastric emptying, which is highly variable, is the rate limiting step for drug absorption from the small intestine and it dictates the onset of drug plasma concentrations (9, 10). 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 % allowed difference in T_{max} between two formulations is undoubtedly too strict criteria for T_{max} parameter of paracetamol.	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		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. in case of special formulations with the claim of extremely rapid release (i.e. faster that the release of standard IR formulations) which is further claimed to be clinically relevant (i.e. faster onset of action in comparison to standard IR formulations). In that case, comparison of T_{max} should be based on non-parametric methods and should be applied to untransformed data. In case of standard IR formulations, only T_{max} range should be comparable. References: 1. Bannwarth B, Pehourcq F. Pharmacologic basis for using paracetamol: pharmacokinetic and pharmacodynamic issues. Drugs. 2003; 63(2):5-13. 2. Panadol 500 mg, filmomhulde tabletten SPC. Last update: 3. 2. 2022. Cited: 14. 7. 2022. Available from: https://www.geneesmiddeleninformatiebank.nl/smpc/h112182_smpc.pdf 3. Lekadol 500 mg, filmsko obložene tablete. SPC. Last update: 11. 5. 2021, Cited: 21. 7. 2022. Available from: http://www.cbz.si/cbz/bazazdr2.nsf/o/3B69BC766DE22686C12579EC0020002B/$File/s-025395.pdf 4. Moore RA, Derry S, Wiffen PJ, Straube S. Effects of food on pharmacokinetics of immediate release oral formulations of aspirin,	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		dipyrone, paracetamol and NSAIDs - a systematic review. Br J Clin Pharmacol. 2015; 80(3):381-8. 5. Divoll M, Greenblatt DJ, Ameer B, Abernethy DR. Effect of food on acetaminophen absorption in young and elderly subjects. J Clin Pharmacol. 1982; 22(11-12):571-6. 6. Forrest JA, Clements JA, Prescott LF. Clinical pharmacokinetics of paracetamol. Clin Pharmacokinet. 1982; 7(2):93-107. 7. Stillings M, Havlik I, Chetty M, Clinton C, Schall R, Moodley I, et al. Comparison of the pharmacokinetic profiles of soluble aspirin and solid paracetamol tablets in fed and fasted volunteers. Curr Med Res Opin. 2000; 16(2):115-24. 8. Grimm M, Scholz E, Koziolek M, Kühn JP, Weitschies W. Gastric water emptying under fed state clinical trial 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/356877/2022 Rev.1* 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	**Partly accepted.**

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		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-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	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
		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 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 C_{max} 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).	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		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-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 $AUC_{(0-inf)}$.	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 do 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 80–125% acceptance range. Only the point estimate should be within the 20% limits. It is agreed that the proposed approach is not able to preserve the type 1 error. But this is the approach 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.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		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 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	This approach is based on the present requirements of the guideline on the investigation of 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
		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 T_{max} as pivotal variable is only applicable in certain situations. Unless the rate of absorption is important with regard to for instance efficacy, statistical evaluation of T_{max} is not required.'* Accordingly, this shall be implemented in the revised guidance text. Another important aspect to consider is that a definition on the importance and acceptable difference in T_{max} is not possible at the active substance level. There are several types of immediate release formulations. This means that product-based granularity is required when assessing the relevancy of T_{max} differences and it is not reasonable to make a single general recommendation that is adequate for all types of immediate release paracetamol products. **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.	

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