Received Date : 01-Feb-2016

Revised Date : 12-May-2016

Accepted Date : 01-Jul-2016

Article type : Article

# Use of the oral sugar test in ponies when performed with or without prior fasting

E. J. Knowles*[1,2], P. A. Harris[3], J. Elliott[4] and N. J. Menzies-Gow[1]

[1]Department of Clinical Sciences and Services, Royal Veterinary College, London, UK; [2]Bell Equine Veterinary Clinic, Mereworth, Kent, UK; [3]Equine Studies Group, WALTHAM Centre for Pet Nutrition, Waltham on the Wolds, Leicestershire, UK; [4]Department of Comparative Biomedical Sciences, Royal Veterinary College, London, UK.

***Corresponding author email**: e.j.knowles@gmail.com

**Keywords:** horse; insulin; EMS; laminitis; endocrine

## Summary

**Background**: It is recommended that the Oral Sugar Test (OST) for insulin dysregulation (ID) is performed after an overnight fast but fasting is impractical in ponies kept solely at pasture. There are few data on OST repeatability and reliability in ponies.

**Objectives:** To report: 1) whether OST results obtained in the morning after an overnight fast or without fasting in the afternoon (FASTING/FED) can be used interchangeably, 2) $T_{max}$[insulin],

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/evj.12607

repeatability and reliability of insulin response to the OST when FASTING or FED, 3) dichotomous agreement (ID/normal) within a small sample when FASTING or FED.

**Study design:** Method comparison study.

**Methods:** OSTs were performed on 4 occasions in 10 adult native British ponies, twice FASTING and twice FED. Insulin concentrations were measured, by radioimmunoassay, at 0-120 minutes ($T_{0,30,60,75,90,120}$). Differences between FASTING and FED results were assessed using mixed effects models. Indices of repeatability and reliability were calculated; dichotomous agreement was reported using kappa statistics.

**Results:** Serum [insulin] was significantly ($p \leq 0.05$) higher at $T_{60}$-$T_{90}$ with prior fasting (estimated differences [95% confidence intervals]): $T_{60}$: 23.5 µIU/ml [8.7-38.4 µIU/ml], $T_{75}$: 27.1 µIU/ml [12.3-41.8 µIU/ml], $T_{90}$: 15.1 [0.36-29.9 µIU/ml]. $T_{max}$ [Insulin] most frequently occurred at $T_{30}$. At any single time point, within-subject coefficients of variation (CVs) were: FASTING: 40% and FED: 31%. The 95% limits for repeatability were FASTING: 29%-340%, FED: 41%-240%. Test reliabilities were FASTING: 0.70, FED: 0.67. For dichotomous interpretation similar results (kappa = 0.7) were obtained using cut-offs of [Insulin] >60 µIU/ml at $T_{60}$ or $T_{90}$ for FASTING and [Insulin] >51 µIU/ml at $T_{30}$ or $T_{60}$ for FED samples.

**Main limitations:** OSTs were performed on a small number of animals on one pasture during one season (spring).

**Conclusions:** Clinicians should beware of interpreting changes in absolute OST results due to poor repeatability. When stabling is unavailable, OSTs of ponies at pasture may yield similar dichotomous results without prior fasting.

**Introduction**

An oral sugar test (OST) is recommended for the identification of insulin dysregulation (ID) in horses at risk of laminitis [1,2]. Clinicians are advised to perform the OST in the morning after an overnight fast and to interpret insulin concentrations in one [1,3] or two [4] blood samples taken 60-90 minutes after administering oral corn sugar. Insulin concentrations >60 µIU/ml are considered to be consistent with ID [4].

Horses kept at pasture cannot always be easily fasted. In a recent survey, 35% of horses registered with UK veterinary practices were kept at pasture 24h/day [5], such animals may not have convenient access to stabling or bare paddocks. Moving animals to stabling for the purposes of ID testing could induce stress and may influence insulin sensitivity [6]. It is therefore unclear whether the OST can be applied to horses kept at pasture.

A previous study [7] tested for differences (but not agreement) between OST results in light breed horses when fasted and stabled or without fasting at pasture at several times of year and found no significant differences between insulin responses. Therefore, results obtained with or without fasting could potentially be used interchangeably subject to further analysis of the agreement between the methods (limits of agreement) [8].

There are few published data on the repeatability or reliability of the OST. One study [2] reported a high coefficient of variation (CV) (45%) for the insulin area under the curve ($AUC_{Insulin}$). Another reported a lower median CV (15.3%) but a wide range (0.3-78.5%) of individual horse CVs [9]. Repeatability describes the variation in repeat measurements under the same conditions (test-retest) [10], whilst reliability describes the measurement error relative to the inherent variability in the population. Reliability therefore indicates the ability of a diagnostic test to distinguish between members of a population [10].

Breed differences between ponies and horses are reported in the insulin response to the OST and the, similar, oral glucose test [11,12] suggesting that ponies may require a different sampling protocol. Further analysis of the OST in ponies rather than horses is therefore warranted.

In a group of pasture-kept adult ponies the present study aimed to determine whether OST results performed either in the morning after an overnight fast or without fasting in the afternoon can be used interchangeably; assess the $T_{max[ins]}$, repeatability and reliability of the insulin response to the OST with or without fasting (and for interchangeable use if appropriate); and report dichotomous agreement (ID or normal) within this sample with or without prior fasting.

**Materials and Methods**

*Animals*

The 10, clinically normal, adult British Native pony mares available from a research herd kept at pasture were used in the study (ages 10-22 years, weights 212-439 kg, body condition scores 4-7/9; further detail is provided in Supplementary Item 1).

*Study design*

The study was conducted in spring (April-May). On day 1, the ponies were weighed and randomly assigned to one of two groups of 5, FASTING and FED. At around 17.00 the FASTING group were brought from their grass paddock into a bare paddock and provided with haylage (to which they were accustomed to eating during the winter) in large ring feeders, estimated to last until 22.00-00.00. The FED group were left at pasture. On day 2 OSTs were performed (as described below) in both groups. Tests were started between 08.00 and 09.00 for the FASTING group and between 13.00

and 14.00 in the FED group. Following the test, ponies returned to grazing their usual grass paddock. Water was available *ad lib* throughout the experiment

One week later (days 7-8) the process was repeated but the two groups were crossed over and the OSTs were repeated. After a further week (days 14-15), the ponies were re-assigned randomly as FASTING or FED (n = 5) and the testing procedure repeated. After an additional week (days 21-22), these groups crossed over and the OSTs were repeated. At the conclusion of the study an OST had been performed on each pony 4 times, twice in the morning after fasting and twice in the afternoon without prior fasting.

*Oral Sugar Tests*

A 14 g catheter[a] was placed aseptically using local anaesthesia[b] into the jugular vein. Baseline blood samples were taken ($T_0$) and corn syrup[c] (0.15 ml/kg) was administered orally using a dosing syringe [2]. Further blood samples were collected at 30, 60, 75, 90 and 120 ($T_{30\text{-}120}$) minutes after oral dosing.

*Sample processing and analysis*

All samples were collected into clot activator blood collection tubes[d] and allowed to clot at ambient temperature. Serum was separated by centrifugation (10 minutes, 2000 g) and stored at -80°C prior to analysis. Insulin concentrations were measured by radioimmunoassay[e].

Basic validation of the radioimmunoassay[e] was performed as the previously validated radioimmunoassay [13] was no longer available. Briefly, intra and inter-assay CVs [14] were estimated using 7 (10-150 µIU/ml) and 15 (12-249 µIU/ml) samples respectively, each analysed two

or three times within or between assays. Dilutional parallelism was assessed by diluting each of 4 high endogenous samples (101-178 µIU/ml) to 50%, 25% and 12.5% with charcoal stripped equine serum (CSS) [13]. Spike recovery was assessed by spiking 6 samples of CSS with the kit manufacturer's human insulin standard (expected values 21-175 µIU/ml).

*Data analysis*

Continuous data were analysed as insulin concentrations taken at individual time points and as the insulin area under the curve ($AUC_{Insulin}$). $AUC_{Insulin}$ was calculated for each test using the trapezoidal sum method with the X axis (y = 0) as the baseline insulin concentration. It was assumed that the true underlying level of insulin dysregulation for each individual did not vary throughout the 22 days of the study. All analysis was performed using statistical[f] and graphical[g] software.

To determine whether insulin values obtained under FASTING and FED conditions agreed, and could therefore be used interchangeably, Bland Altman plots (using repeated measurements) were plotted of the differences between the methods against their mean [8]. Plots were examined for evidence of bias between the methods. As bias was detected, the limits of agreement were not quantified and linear mixed effects models were generated to characterise the differences between FASTING and FED samples. Sampling time (for single time points), sampling day, FASTING/FED and their interactions were initially included as fixed variables and removed according to statistical significance. Subject was included as a random variable. Insulin concentration was the outcome variable and an auto-regressive covariance structure (AR1) was used. Estimated marginal means were calculated from the final model and pairwise *post-hoc* comparisons were performed (without adjustment of confidence intervals for multiple comparisons/least significant difference). The normality of the distribution of the residuals was assessed to ensure normality by histogram.

$T_{max[ins]}$ was the sampling time at which the highest insulin concentration was recorded for each OST curve. To determine repeatability coefficients, Bland-Altman plots of the differences between test and retest values (measurement error) against their mean ('true value') were plotted to determine evidence of bias or any association between measurement error and means. Logarithmic transformations were performed if measurement error was proportional to the mean and used to calculate ratios for the repeatability coefficient [15]. The distribution of the differences (of transformed data) was checked to ensure normality using histograms. The CVs were calculated using the root mean square method [14]. Reliability coefficients (intra-class correlation coefficients) were calculated using two way ANOVA models applied to logarithmically transformed data [16].

To analyse dichotomous data, subjects were classified as normal (N) or as having insulin dysregulation (ID) such that ID was defined as [Insulin]>60 μIU/ml at $T_{60}$ or $T_{90}$ with prior fasting [4]. Agreement between the diagnosis on the first and second test for each subject is reported as the kappa statistic [16]. Agreement was also calculated for [Insulin]>60 μIU/ml at $T_{30}$ or $T_{60}$ based on the $T_{max[Ins]}$ data obtained in the present study.

The within-subject agreement (ID/normal) of single blood samples taken at $T_{60}$-$T_{90}$ and $T_{30}$-$T_{60}$ was calculated to indicate how commonly a single blood sample taken during the periods 60-90 minutes or 30-60 minutes produced the same result.

Finally, to report whether a dichotomous interpretation applied to FED subjects provided similar results to FASTING subjects, a sampling frame and cut-offs for insulin concentration under FED conditions were extrapolated. Two sampling times were chosen in order to include $T_{max[Ins\}}$ in most

cases and the cut-off was extrapolated based on the percentile of results separated using the recommended method ([Insulin]>60 µIU/ml at $T_{60}$ or $T_{90}$ [4].

**Results**

A complete data set was obtained for all ponies other than one sample for which the pony could not be caught (this OST curve was excluded from $T_{max[Ins]}$ analysis, and the subsequent sample ($T_{75}$) was used instead for dichotomous interpretation). OST results for each pony are shown in Supplementary Item 2.

Insulin assay validation gave estimated intra-assay and inter-assay CVs of 12 and 14% respectively. Mean (± s.d.) recovery for high endogenous sample dilutional parallelism was 90% ± 10%, observed: expected $r^2$ = 0.97. Spike recovery of human standard mean (± s.d.) recovery was 113% ± 5%, observed:expected $r^2$ = 1.

*Agreement between OST test results with and without fasting.*

Bland-Altman plots of mean test results for all time points under FASTING and FED conditions are shown in Figure 1. The plots provide evidence of changing bias between the methods. When mean values are low FED values are higher than FASTING values but when mean values are high FED values are lower. Methods to model the limits of such changing agreement are complex [17] and are not presented in light of the evidence of significant and changing bias. Instead differences between the methods were explored further using linear mixed effects models.

For single time points: the effects of FASTING/FED (p = 0.009), sampling time (P<0.001) and their interaction (p = 0.003) were significant. Estimated, statistically significant, differences FASTING-FED (95% confidence intervals) were: $T_{60}$: 23.5 µIU/ml (8.7-38.4 µIU/ml) p = 0.002, $T_{75}$: 27.1 µIU/ml (12.3-41.8 µIU/ml) p<0.001, $T_{90}$: 15.1 (0.36-29.9 µIU/ml) p = 0.045. $AUC_{insulin}$ values also differed between FASTING and FED conditions, estimated difference 1574 µIU/ml.min (577-2570 µIU/ml.min) p = 0.004. Estimated marginal means for FASTING and FED conditions are shown in Figure 2.

### $T_{max[ins]}$, *repeatability and reliability*

$T_{max[ins]}$, repeatability and reliability were calculated independently for each method FASTING and FED. $T_{30}$ was the most frequent $T_{max[ins]}$ (FASTING = 50%, FED = 60% of OST curves). The $T_{30}$ and $T_{60}$ samples included $T_{max[ins]}$ for 85% (FASTING) and 95% (FED) of OST curves.

Bland-Altman plots, of the differences between replicate measurements and their mean, are shown in Figure 3. There was evidence that measurement error was proportional to the mean; a logarithmic transformation was applied and improved the distribution of the differences [15]. The distribution of the differences between replicates (of transformed data) was approximately normal. Ratios, rather than absolute values, for repeatability coefficients are therefore reported (i.e. limits within which 95% of replicate samples are expected to fall). Repeatability coefficients were large with and without fasting and are shown in Table 1 with the coefficients of variation and reliability.

### *Dichotomous interpretation*

With insulin dysregulation (ID) defined as insulin >60 µIU/ml at either $T_{60}$ or $T_{90}$ in fasting subjects [4]. Four ponies were considered ID on both occasions, one was defined as ID on the first but not the second fasting test, five were considered normal on both occasions (kappa = 0.8). Due to the high

frequency of $T_{max[Ins]}$ at $T_{30}$ the same analysis was performed for insulin >60 µIU/ml at either $T_{30}$ or $T_{60}$. Six ponies were considered ID, 4 were not, agreement between the first and second tests was perfect (kappa = 1).

The within-subject agreement for single samples taken $T_{60}$-$T_{90}$ and $T_{30}$-$T_{60}$ were moderate (kappa = 0.55) and fair (kappa = 0.38) indicating that taking a single sample during either period often yielded different results in the same subject.

To determine if similar dichotomous results (ID or normal) were obtained without fasting, a cut off of [Insulin]>51 µIU/ml at $T_{30}$ and/or $T_{60}$ was extrapolated. Applying these criteria in FED subjects, four ponies were considered ID on both occasions, two were considered ID on the first but not the second test and 4 were considered normal on both occasions, (kappa = 0.62). When compared with sampling the same ponies with fasting ([Insulin]>60 µIU/ml at $T_{60}$ or $T_{90}$) there was good agreement between results (kappa = 0.7). Comparative results obtained using these cut-offs and conditions are shown in Table 2.

**Discussion**

Significant differences were found between the insulin response to the OST when performed with or without prior fasting. Fasting exacerbated the insulin response to oral sugar to an extent that is likely to affect clinical interpretation; for example the estimated difference (FASTING-FED) at $T_{75}$= 27.1 µIU/ml. Limits of agreement for FASTING and FED conditions were not calculated due to the magnitude and changing nature of the difference between methods. It is therefore inadvisable to use results interchangeably. However, results obtained under FED conditions had similar reliability (i.e. discriminatory power) to those obtained under FASTING conditions and produced similar

results, when interpreted on a dichotomous basis using alternative criteria. Therefore, when fasting is impractical, similar basic OST interpretation may be obtained without prior fasting as under standardised conditions. However, further validation of testing without fasting in a larger population is required.

In the current study, the most commonly recorded $T_{max\ [ins]}$ occurred at $T_{30}$ under both FASTING and FED conditions. Smith *et al.* [12] reported a shorter $T_{max\ [ins]}$ in ponies (mean = 60 minutes) than in horses (mean = 69 minutes), whilst Schuver *et al.* [2] reported maximum concentrations at 60 or 90 minutes in adult horses. The rapidity of the response in the present study suggests a role for incretins [18] but the cause of the apparent breed differences is unclear and warrants further investigation.

The vast majority of $T_{max\ [ins]}$ occurred at $T_{30}$ or $T_{60}$. Therefore if clinicians aim to include $T_{max\ [ins]}$ in ponies then earlier sampling times at $T_{30}$ and $T_{60}$ may be more appropriate than the recommended sampling window of 60-90 minutes that typically includes $T_{max\ [ins]}$ in horses [2]. Alternatively if the clinician seeks to detect a prolonged insulin response then later sampling times or calculation of the $AUC_{Insulin}$ may be required. The most clinically useful application and interpretation of the OST is yet to be determined. There are differences between the insulin response to oral and intravenous glucose [19], therefore an association with other measures of ID may be less important than an association with a clinical outcome such as laminitis. Published data have not yet specifically associated OST results with previous or future predisposition to laminitis.

The insulin response to the OST showed wide limits of agreement and high coefficients of variation at single time points and for $AUC_{Insulin}$ both with and without fasting. These findings indicate poor repeatability and are consistent with the high CV (45% for $AUC_{Insulin}$) reported previously [2]. Another study reported a lower median (15.3%) but wide range (0.3-78.5%) of individual horse CVs [9]. A mean CV at 90 minutes of 25% was reported for a similar test using dextrose powder [20] and a CV of 19% was reported for $AUC_{Insulin}$ when a Scandinavian syrup was used as the sugar challenge [21]. The reporting of a mean or median CV generated from individual CVs may produce a biased (typically lower) estimate, the results are therefore not comparable and this approach is discouraged [14,22].

Whilst the CV indicates a summary value for test-retest differences it does not indicate full extent by which two measurements of the same quantity may differ [10,16]. The repeatability coefficient is the maximum expected difference between test-retest values on 95% of occasions and is reported as a ratio when measurement error is proportional to the measurement value [15] as in the current study. Clinicians must exercise caution when interpreting test results for individual cases. Anecdotal reports suggest that some clinicians repeat oral sugar tests to assess the effect of dietary or pharmacological interventions on insulin regulation. The poor test repeatability shown in our study agrees with that published by Schuver *et al.* [2] and cautions against such an approach. Differences in OST results may simply be a reflection of the poor repeatability of this test and not due to the effects of veterinary/dietary/management interventions. Good levels of agreement between repeated tests were obtained using a simple dichotomous interpretation and may be a more appropriate use of the test in clinical practice.

The poor repeatability of the OST is perhaps unsurprising given the repeatability of other tests for insulin sensitivity in horses and other species. The CV for $AUC_{Insulin}$ in unfasted subjects in the present study (23% (14-29%)) is similar to that reported for estimates of insulin sensitivity derived from the frequently sampled intravenous glucose tolerance test 24% in horses [23]. In people a 75 g oral glucose tolerance test showed poor repeatability, particularly for insulin concentrations taken 2 hours post-challenge [24].

The estimated repeatability was worse when the OST was performed with rather than without fasting, but similar reliability occurred under both sets of conditions. This implies the test has similar discriminatory power with or without fasting and similar dichotomous results were obtained using alternative criteria for evaluation of samples without fasting. It is important to stress that the OST has not been optimised for use without fasting and the magnitude of the insulinaemic response was reduced without fasting. Higher doses of oral sugar could elicit a greater insulinaemic response and may improve test performance. The OST uses a small dose of dextrose derived digestible sugars (150 mg/kg [2]) compared with the oral glucose challenge test (1g/kg [12] or 1.5g/kg [11]). Higher doses of oral sugar are therefore likely to be well tolerated and require investigation.

An important source of measurement error is the radioimmunoassay used to measure insulin concentrations. Basic validation of the assay estimated good linearity but acceptable repeatability (intra-assay CV = 12%, inter-assay CV = 14%). All study samples were analysed in duplicate and the CV of duplicate measurements (2.6%) was good. A previously validated radioimmunoassay [13,25] has been discontinued. Alternative assays include an equine specific ELISA [13,25] and a chemiluminescent immunoassay [26]. Inter-assay CVs of 4-14% are reported for the chemiluminescent assay [26]; for the ELISA intra assay CVs of 8% and 11% and inter assay CVs of 7% and 9% are reported [13,25]. As discussed, the calculation of a mean CV may result in a biased,

lower, estimate [14], indeed In the present study mean CVs yielded lower values (data not shown) for the immunoassay validation that are consistent with previously reported values using the ELISA [13]. Clinicians should therefore be aware of the limitations of the available insulin assays when interpreting results. In particular, agreement between different assays is poor [13,25,26] and hampers comparison of results between different studies. Good recovery of an equine insulin standard or concordance with gas-chromatography-mass spectrometry has not been published for any currently available assays [25,26].

The value of taking post-challenge samples at more than one time point requires further investigation. In the present study, there was improved test repeatability when $AUC_{Insulin}$ was calculated rather than the use of single time points, however test reliability was similar. For dichotomous interpretation, within this small sample population, Kappa statistics indicated that single samples taken during the recommended sampling window $T_{60}$-$T_{90}$ yielded moderate agreement (i.e. different results would be obtained relatively frequently from single samples) but agreement was improved when two samples were taken 30 minutes apart.

This study has several limitations. The population was small and estimates of population parameters generated from these data are inherently imprecise as indicated by the confidence intervals. The kappa statistics report agreement in this sample population but cannot provide accurate estimates of true population parameters. Sample sizes of 100-200 are often recommended in the human literature for Bland-Altman analysis [27] but, for practical and ethical reasons, are rarely achieved using research animals in the equine literature. The repeated measures design maximised the data obtained from a limited number of research animals in line with the principles of the 3Rs. The small population was however considered to be representative of the clinical population in which the OST is employed in the UK however, extremely high insulin concentrations that are sometimes

encountered in practice [13] did not occur. Only one type of pasture was used and the study was only conducted at one time of year. The study was conducted in the spring as spring pasture was expected to emphasise any differences between FASTING and FED samples. It was assumed that each pony's insulin regulation did not change significantly over the 22 day study period such that each test replicate was measuring the same true value, an assumption that might be challenged. None of the ponies showed a visible change in body condition score during the study and all of the ponies were accustomed to their pasture diet before being enrolled in the study. Thus, large alterations in insulin regulation appear unlikely but cannot be ruled out. Any changes during the study period would have resulted in overestimates for the coefficients of variation and repeatability coefficient, however the similar CV reported by Schuver *et al.* [2] for tests performed on consecutive days suggests test spacing was not a significant factor. When conducting fasting tests, it was assumed rather than verified that the haylage had been consumed by midnight. Finally, test with fasting were performed in the morning whilst those without prior fasting were conducted in the afternoon. We cannot exclude the possibility that the different time of sampling may have influenced the results.

In conclusion, OST results obtained without prior fasting differed significantly from those obtained from fasted subjects and results should not be used interchangeably. The OST showed poor repeatability and absolute results should be interpreted with care. A simple dichotomous interpretation may be most appropriate in clinical practice. If stabling is unavailable, similar results may be obtained from ponies without fasting as under standardised conditions, however further test optimisation and validation for unfasted conditions are required. Taking two, rather than one, post challenge samples 30 minutes apart, perhaps at 30 and 60 minutes, is likely to yield more consistent results.

**Authors' declaration of interests.**

P. Harris is employed by WALTHAM.

**Ethical animal research**

The study was conducted under a UK Home Office project licence with ethical approval from the Royal Veterinary College Ethics and Welfare Committee.

**Sources of funding**

E. Knowles' PhD is funded by The Mellon Fund and WALTHAM. P. Harris is employed by WALTHAM. J. Elliott and N. Menzies-Gow are employed by The Royal Veterinary College.

**Acknowledgements**

The authors thank Olivia Morgan for her care of the ponies and assistance with sample collection.

**Authorship**

All authors contributed to study design, data analysis/interpretation, preparation and final approval of the manuscript. Data collection/study execution was by E. Knowles and N. Menzies-Gow.

**Figure legends**

**Fig 1:** Bland-Altman plots of agreement between FASTING and FED OST results for single time points (A) or $AUC_{insulin}$ (B).

**Fig 2:** Estimated marginal means (± 1.96 SEM) for OST insulin concentrations at single time points (A) and $AUC_{insulin}$ (B) for FASTING and FED conditions. Mean FASTING values that differ significantly ($p<0.05$) from the equivalent values without fasting are marked with an asterisk.

**Fig 3:** Bland-Altman plots showing the difference between replicated (test-retest) measures against their means for individual time points (A (FASTING) and B (FED)) and $AUC_{Insulin}$ (C (FASTING) and D (FED)).

**Supplementary Information**

**Supplementary Item 1:** Oral sugar test curves for all subjects.

**Supplementary Item 2:** Signalment, laminitis history and available metabolic data from all subjects.

**Table 1:** Coefficients of repeatability, variation and reliability for FASTING and FED OST results.

| | | Repeatability coefficient | Coefficient of Variation (95% confidence interval) | Reliability (95% confidence interval) |
|---|---|---|---|---|
| Single time points | FASTING | 29%-340% | 40% (33-46%) | 0.70 (0.55-0.81) |
| | FED | 41%-240% | 31% (25-36%) | 0.67 (0.50-0.79) |
| Area under the Curve | FASTING | 42%-240% | 29% (11-39%) | 0.70 (0.15-0.92) |
| | FED | 53%-189% | 23% (14-29%) | 0.69 (0.18-0.91) |

**Table 2:** OST test results (showing number of ponies in each category) showing Insulin Dysregulation (ID) or Normal (N) for FASTING (ID= Insulin>60 µIU/ml) at $T_{60}$ or$T_{90}$) or FED (ID= [Insulin] >51 µIU/ml at $T_{30}$ or $T_{60}$) samples.

| | **FED N/N** | **FED ID/N** | **FED ID/ID** |
|---|---|---|---|
| **FASTING N/N** | 4 | 1 | 0 |
| **FASTING ID/N** | 0 | 0 | 1 |
| **FASTING ID/ID** | 0 | 1 | 3 |

**Manufacturers' addresses**

[a]Intraflon, Vygon UK Ltd, Swindon, Wiltshire, UK.

[b]Intra-Epicaine, Dechra Veterinary Products, Shrewsbury, Shropshire, UK.

[c]Karo Light Corn Syrup, ACH Food Companies Inc, Cordova, Tennessee, USA.

[d]Venosafe, Terumo UK Ltd, Bagshot, Surrey, UK.

[e]Insulin CT I-125, MP Biomedical, Ilkirch, France.

[f]IBM SPSS Statistics 22, IBM UK, Portsmouth, Hampshire, UK.

[g]Graphpad Prism, Graphpad Software, La Jolla, California, USA.

**References**


1. Frank, N. (2011) Equine metabolic syndrome. *Vet. Clin. North Am. Equine Pract.* **27**, 73-92.
2. Schuver, A., Frank, N., Chameroy, K.A. and Elliott, S.B. (2014) Assessment of Insulin and Glucose Dynamics by Using an Oral Sugar Test in Horses. *J. Equine Vet. Sci.* **34**, 465-470.
3. Rendle, D. (2015) Proactive Approaches to Prevention of Laminitis. *Vet. Times Equine* **1**, 10-

15.

4. Frank, N. and Geor, R. (2014) Current best practice in clinical management of equine endocrine patients. *Equine Vet. Educ.* **26**, 6-9.

5. Wylie, C.E., Ireland, J.L., Collins, S.N., Verheyen, K.L.P. and Newton, J.R. (2013) Demographics and management practices of horses and ponies in Great Britain: A cross-sectional study. *Res. Vet. Sci.* **95**, 410-417.

6. Frank, N. and Tadros, E.M. (2014) Insulin dysregulation. *Equine Vet. J.* **46**, 103-112.

7. Banse, H.E. and McFarlane, D. (2014) Comparison of Three Methods for Evaluation of Equine Insulin Regulation in Horses of Varied Body Condition Score. *J. Equine Vet. Sci.* **34**, 742-748.

8. Bland, M.J. and Altman, D.G. (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **327**, 307-310.

9. Frank, N. and Walsh, D.M. (2015) Repeatability of the oral sugar test and glucagon like peptide-1 and high molecular weight adiponectin concentrations in horses with equine metabolic syndrome. *J. Vet. Intern. Med. 2* **29**, 1241.

10. Bartlett, J.W. and Frost, C. (2008) Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. *Ultrasound Obstet. Gynecol.* **31**, 466-475.

11. Bamford, N.J., Potter, S.J., Harris, P. A. and Bailey, S.R. (2014) Breed differences in insulin sensitivity and insulinemic responses to oral glucose in horses and ponies of moderate body condition score. *Domest. Anim. Endocrinol.* **47**, 101-107.

12. Smith, S., Harris, P.A. and Menzies-Gow, N.J.J. (2015) Comparison of the in-feed glucose test and the oral sugar test. *Equine Vet. J.* **48**, 224-227.

13. Borer-Weir, K.E., Bailey, S.R., Menzies-Gow, N.J., Harris, P.A. and Elliott, J. (2012) Evaluation of a commercially available radioimmunoassay and species-specific ELISAs for measurement of high concentrations of insulin in equine serum. *Am. J. Vet. Res.* **73**, 1596-1602.

14. Bland, M. How do I calculate a within-subject coefficient of variation. https://www-users.york.ac.uk/~mb55/meas/cv.htm. Accessed July 20, 2015.

15. Bland, J.M. and Altman, D.G. (1996) Measurement error proportional to the mean. *BMJ* **313**, 106.

16. Bland, M. (2015) Clinical Measurement. In: *An Introduction to Medical Statistics* 4th ed. Oxford University Press, Oxford, UK. pp 313-345.

17. Bland, J.M. and Altman, D.G. (1999) Measuring agreement in method comparison studies. *Stat. Methods Med. Res.* **8**, 135-160.

18. Kim, W. and Egan, J.M. (2008) The role of incretins in glucose homeostasis and diabetes treatment. *Pharmacol. Rev.* **60**, 470-512.

19. de Laat, M.A., McGree, J.M. and Sillence, M.N. (2015) Equine hyperinsulinemia: investigation of the enteroinsular axis during insulin dysregulation. *Am. J. Physiol. Endocrinol. Metab.* doi: 10.1152/ajpendo.00362.2015.

20. de Laat, M.A., Pollitt, C.C. and Sillence, M.N. (2014) Repeatability of an in-feed sugar test in ponies. *Equine Vet. J.* **46**, *Suppl.* **47**, 8.

21. Lindåse, S., Nostell, K., Askerfelt, I. and Bröjer, J. (2015) A modified oral sugar test for evaluation of insulin and glucose dynamics in horses. *Acta Vet. Scand.* **57**, O4.

22. Brightwell, B. and Dransfield, B. (2013) Coefficient of variation. In: *Avoiding and detecting statistical malpractice*, 1st ed., Eds: B. Brightwell and B. Dransfield, Influential points, unit 3. http://influentialpoints.com/hyperbook.htm.

23. Pratt, S.E., Geor, R.J. and McCutcheon, L.J. (2005) Repeatability of 2 methods for assessment of insulin sensitivity and glucose dynamics in horses. *J. Vet. Intern. Med.* **19**, 883-888.

24. Mooy, J.M., Grootenhuis, P.A., Vries, H. de, Kostense, P.J., Popp-Snijders, C., Bouter, L.M. and Heine, R.J. (1996) Intra-individual variation of glucose, specific insulin and proinsulin

concentrations measured by two oral glucose tolerance tests in a general Caucasian population: the Hoorn Study. *Diabetologia* **39**, 298-305.

25. Tinworth, K.D., Wynn, P.C., Boston, R.C., Harris, P. A., Sillence, M.N., Thevis, M., Thomas, A. and Noble, G.K. (2011) Evaluation of commercially available assays for the measurement of equine insulin. *Domest. Anim. Endocrinol.* **41**, 81-90.

26. Banse, H.E., Mccann, J., Yang, F., Wagg, C. and McFarlane, D. (2014) Comparison of two methods for measurement of equine insulin. *J. Vet. Diagn. Invest.* **28,** 527-530.

27 Bland, M. How can I decide the sample size for a study of agreeement between two methods of measurement? https://www-users.york.ac.uk/~mb55/meas/sizemeth.htm. Accessed 10th March 2016.