

INTERPRETATION OF FREEZING POINT DEPRESSION IN MILK

MAY 2021

Summary

- This industry standard reviews the interpretation of freezing point depression in milk in view of the long-term decline in FPDs in Great Britain.
- It replaces the interpretation provided in British Standard (BS) 3095: Part 2: 1988, which has been withdrawn.
- The assumption underlying the revision of the FPD interpretation is that the sustained decline is a natural phenomenon and not a reflection of increasing inclusion of extraneous water in milk, i.e. the FPD of ‘genuine’ milk in GB has changed to a ‘new normal’.
- Dairy UK commissioned the University of Reading to conduct a detailed analysis of the most recent FPD data and 16 million records were included.
- Over the period from 2006 to 2020, the average FPD in Great Britain has fallen from 521 m°C to 516 m°C.
- New interpretation limits have been set in line with the approach followed in BS 3095: Part 2: 1988.
- The average and the standard deviation will be reviewed annually and Dairy UK members will consider the need to change the threshold for the year ahead in full consultation with relevant stakeholders.

From 6th May 2021, GB milk is considered acceptable if it records an FPD of 505 m°C or above.

Contents

SECTION 1: BACKGROUND	1
SECTION 2: RESEARCH IN 2020.....	2
SECTION 3: TRENDS IN FPD IN GREAT BRITAIN: 2006-2020.....	3
SECTION 4: NEW INTERPRETATION OF FREEZING POINT DEPRESSION IN MILK.....	4
SECTION 5: FURTHER INFORMATION	5
APPENDIX 1: Relevant material from an analysis of historic FPD data and discussion of options for revision of the FPD standard	7
APPENDIX 2: FREEZING POINT DEPRESSION MILK SCAN-CRYOSCOPE METHOD COMPARISON STUDY (February 2021)	13

SECTION 1: BACKGROUND

- Freezing point depression (FPD) is a required quality standard for milk. It is used as an indicator of possible presence of extraneous water to milk.
- The standard is set out in Regulation (EU) 1308/2013 Annex VII, part IV (now transposed in UK legislation). This regulation stipulates that: “*Drinking milk shall have a freezing point close to the average freezing point for raw milk recorded in the area of origin of the drinking milk collected*”.
- Thus far, British milk has been considered acceptable if it records an FPD of 509 m°C or above.
- This is based on a study conducted in 1977, the results of which are included in *British Standard 3095: Part 2: 1988*. That year-long study involved testing milk from 100 herds in England and Wales each quarter – it included testing authenticated milk¹ at evening and morning milkings, as well as farm bulk tank milk containing the yield of the two milkings.
- The average FPD of authenticated milk was 521.6 m°C². This was assumed to represent the natural variation of FPD in genuine milk containing no extraneous water.
- Action point thresholds were then set on the basis of the mean and the standard deviation as such²:

Table 1: Interpretation limits set in BS 3095: Part 2: 1988³

Mean	521.6	Action point thresholds	Interpretation
		≥514	Accepted as being free from extraneous water
m -2sd	514	513-509	FPD alone is insufficient justification to presume it contains extraneous water
m -3sd	509	508-504	Likely to contain extraneous water - repeat sampling strongly recommended
m -4sd	504	<504	Almost certainly contains extraneous water

- A contemporary survey of 48 herds was carried out in Scotland, the results of which were in general agreement with those of the larger survey (as reported in *British Standard 3095: Part 2: 1988*).
- On the basis of these results, the industry agreed on 509 m°C as a threshold to be used across the supply chain, whereby milk with FPD values falling below this would not be deemed acceptable.
- Since then, there has been a gradual downward trend in FPD in milk in Great Britain such that a review of the action point thresholds is justified.
- This document sets out the findings of that review and outlines new interpretation for FPD in milk to be applicable from 06th May 2021 and reviewed on a regular basis thereafter.

¹ Authenticated milk = samples of milk representative of a whole milking and taken under supervised conditions to ensure no pick-up of extraneous water (*BS 3095: Part 2: 1988*)

² Values have been converted from the Hortvet method to true Celsius (IDF method 108C)

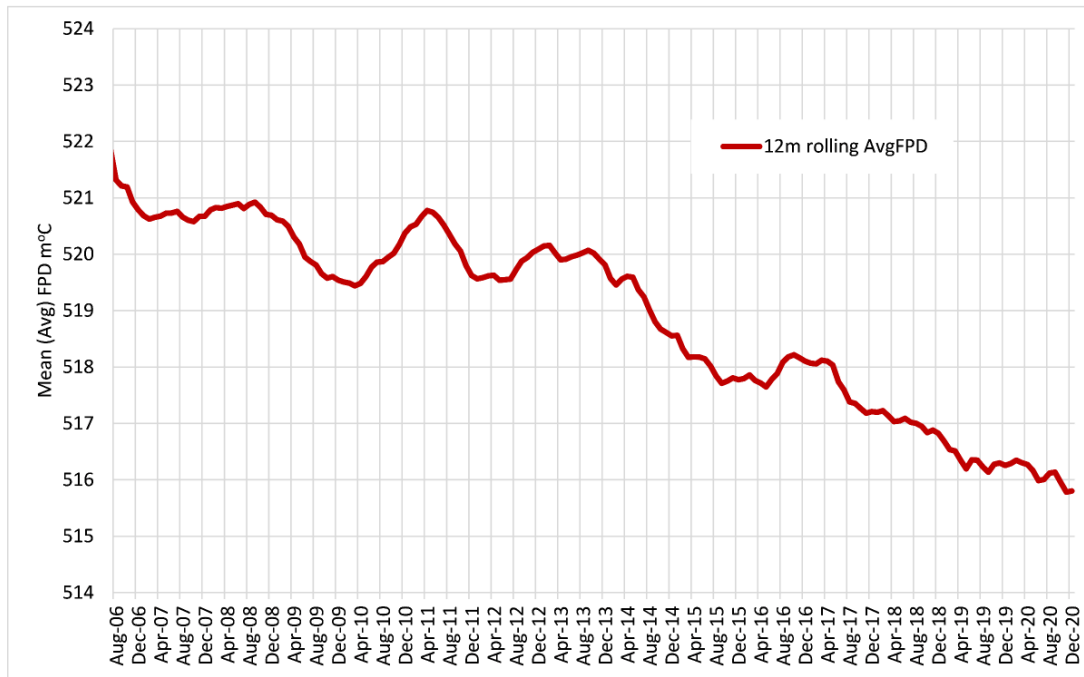
³ Interpretation is based on statistical analysis of FPD results obtained by Hortvet methods described in BS 3095: Section 1.1 and 1.2 and the method described in BS 1741 for determining titratable acidity

SECTION 2: RESEARCH IN 2020

- In 2020, the University of Reading was commissioned by Dairy UK to support the development of new interpretation limits based on the most recent FPD data from GB. This work was carried out with support from Arla Foods, Cotteswold Dairy and Lactalis.
- Data was provided by National Milk Laboratories (NML), the laboratory carrying out routine FPD testing of milk from virtually all farmers in Great Britain - 16 million records covering the period from August 2006 to December 2020 were assessed by the University of Reading to confirm the trend of FPD values over time.
- The reference analysis for freezing point is thermistor cryoscopy (ISO 5764: 2009), but in routine laboratory conditions it is determined by mid-infrared (mIR) technology (ISO 9622: 2013). NML conducted a study to confirm the comparability of FPD results secured by mIR and the cryoscope reference method and found that FPD test results secured by mIR can be used as a proxy for results secured through the reference method. Summarised results of the comparison study are in Appendix 2.
- While the quantity of data now available is far greater than the 1977 survey data, the quality is different in that authenticated milk samples are not available. While the 1977 samples could be taken as representative of genuine milk, the pool of routinely collected samples consists of both genuine milk and milk which, for a variety of reasons, may have extraneous water. The latter segment of the data can reasonably be expected to constitute a very small proportion of the samples but could sufficiently inflate the standard deviation.
- The 1977 study found a consistent but small difference in mean FPD between authenticated samples and farm bulk tank samples, from which it was concluded that “a small amount of extraneous water was present in the bulk tank samples”. It should therefore be considered that otherwise genuine milk may have a small amount of extraneous water as a result of processes between milking and point of collection from the bulk tank or tanker. This may apply to practically all the routine NML samples and could account for a slight difference in the mean FPD if it were possible to compare with authenticated samples.
- Routine samples will also include a small proportion that contain more than slight amounts of extraneous water, with much lower FPD. The effect on the mean FPD of these few samples will be small, but the effect of even a small proportion of low ‘outlier’ FPDs on the calculated standard deviation can be significant.
- Adjustments are therefore required to the approach used in *BS 3095: Part 2: 1988* in view of the fact that the standard deviation calculated from the routine NML data does not accurately reflect the distribution of 100% genuine milk.
- These adjustments are described in further detail in Section 4 of this document.

SECTION 3: TRENDS IN FPD IN GREAT BRITAIN: 2006-2020

- The trend seen in FPD results obtained from the testing of farm bulk tank samples in England, Scotland and Wales by NML is illustrated in the graph below:



Source: The University of Reading (NML data)

- Over the period 2006 to 2020, the average FPD has fallen from 521 m°C to 516 m°C.
- Trend analysis performed by the University of Reading indicates that:
 - Mean FPD has been decreasing by 0.41 m°C per year since January 2014.
 - The percentage of FPD values less than 509 m°C has been increasing by 0.56% per year since January 2017.
- There have been studies and surveys in the past that demonstrate that there are environmental, husbandry and cow-related factors associated with variations in FPD.
- The assumption underlying the revision of the FPD threshold has to be that the sustained decline in FPD is a natural phenomenon and not a reflection of increased presence of extraneous water in milk, i.e. the FPD of genuine milk in GB has changed to a 'new normal'. In practical terms, the quality of the milk will not change as a result of this standard - the values have been amended to reflect changes to the national average composition of milk and natural variance seen over the past 40 years, and the margin between herd average and the accept/reject level remains unchanged.
- Further information is available in Appendix 1.

SECTION 4: NEW INTERPRETATION OF FREEZING POINT DEPRESSION IN MILK

- Setting new interpretations of FPD in milk – as outlined below - has involved the following steps:
 - The removal of outliers to clean the data – this tackles the issue of inclusion of non-genuine milk samples in the routine NML data by excluding low outlier FPDs from the calculation of the mean and the standard deviation. When using data from 2020, the best fit between the main body of data and a normal distribution appears when data are excluded from ‘mean minus 2sd’.
 - An assessment of the 3-year rolling average for FPDs tested by the mIR method (i.e. covering data from 2018 to 2020) using the ‘cleaned’ data. Using a 3-year rolling average rather than a 12-month average reduces the sensitivity of the process to abnormal events and ensures that any changes to interpretation of FPD in milk reflect a sustained and well-established change. The differences between 12-, 24- and 36-month rolling statistics up to December 2020 are less than one thousandth of a degree, and do not affect the final threshold.
 - Interpretation limits are then determined in line with those set out in BS 3095: Part 2: 1988, as described in Section 1, Table 1. The new interpretation limits are set in **Table 2** below.
- Further information is included in Appendix 1, which provides extracts from the University of Reading study and which has been drafted by the author of that study.
- This new interpretation applies to Great Britain as data was sourced from this area - Northern Ireland is a geographically separate region with a different climate and it may not be appropriate to extend the conclusions of this new data across the UK. Northern Ireland processors may decide to adopt the standard using local data to check the comparability of the results.
- The rolling average and standard deviation covering the previous three years will be reviewed annually in January/February and Dairy UK members will consider the need to change the threshold for the year ahead with any adjustment applying from 01 April.
- Dairy UK will be responsible for overseeing the decision-making process and will secure input from relevant stakeholders.
- Any decision to change the standard will be made annually by the end of February and notification of the change will be through a statement shared through relevant industry channels.

From 6th May 2021, GB milk is considered acceptable if it records an FPD of 505 m°C or above.

Table 2: Interpretation limits applicable from 06 May 2021⁴

Mean	516.7		Interpretation
		≥509	Free of extraneous water
m -2sd	509	508-505	May contain extraneous water
m -3sd	505	504-501	Likely to contain extraneous water
m -4sd	501	<501	Assume presence of extraneous water

⁴ Interpretation applicable to results obtained by thermistor cryoscope method (ISO 5764:2009)

SECTION 5: FURTHER INFORMATION

- For more information regarding the background to this standard please contact Dairy UK at info@dairyuk.org

APPENDIX 1: Relevant material from an analysis of historic FPD data and discussion of options for revision of the FPD standard

This material is taken from the full report of the University of Reading study, written by Nick Taylor (n.m.taylor@reading.ac.uk), that presents data supporting revision of the threshold to be used for suspecting milk to contain extraneous water based on its FPD. Note that all FPD values discussed are expressed as m°C ('true Celsius': IDF method 108C) unless specifically indicated otherwise.

Freezing point depression (FPD) is a required quality standard for milk. Its intended use is as an indicator of possible presence of extraneous water in milk. FPD thresholds for 'acceptable' milk are set in relation to the average freezing point for raw milk recorded in the area of origin of the milk being collected. In setting the current standard (BS3095: Part 2: 1988) the British Standards Institution (BSI) used FPDs of 'authenticated' milk sampled from a survey of herds in England and Wales in 1977. The mean and standard deviation were assumed to represent the natural variation of FPD in genuine milk containing no extraneous water. Action point thresholds were set based on the mean minus 2, 3 and 4 standard deviations (sd), with mean minus 3sd considered the threshold below which suspicion of extraneous water is justifiable. The argument for this is rooted in the statistics of the normal distribution, which predict just a 0.14% probability that genuine 'authenticated' milk could have a FPD of less than mean minus 3sd.

In Great Britain (GB) a long term trend of reducing FPD has been observed and in the recent five to ten years there has also been a noticeable increase in the percentage of bulk milk samples failing initial FPD screening and requiring confirmatory cryoscopy. The increased incidence of low FPD has led to pressure for revision of the standard in GB to reflect current milk production conditions. The assumption underlying the call for revision of the threshold FPD is that the sustained decline is a natural phenomenon and not a reflection of increasing inclusion of extraneous water in milk, i.e. the FPD of '*genuine*' milk in GB has changed to a 'new normal'.

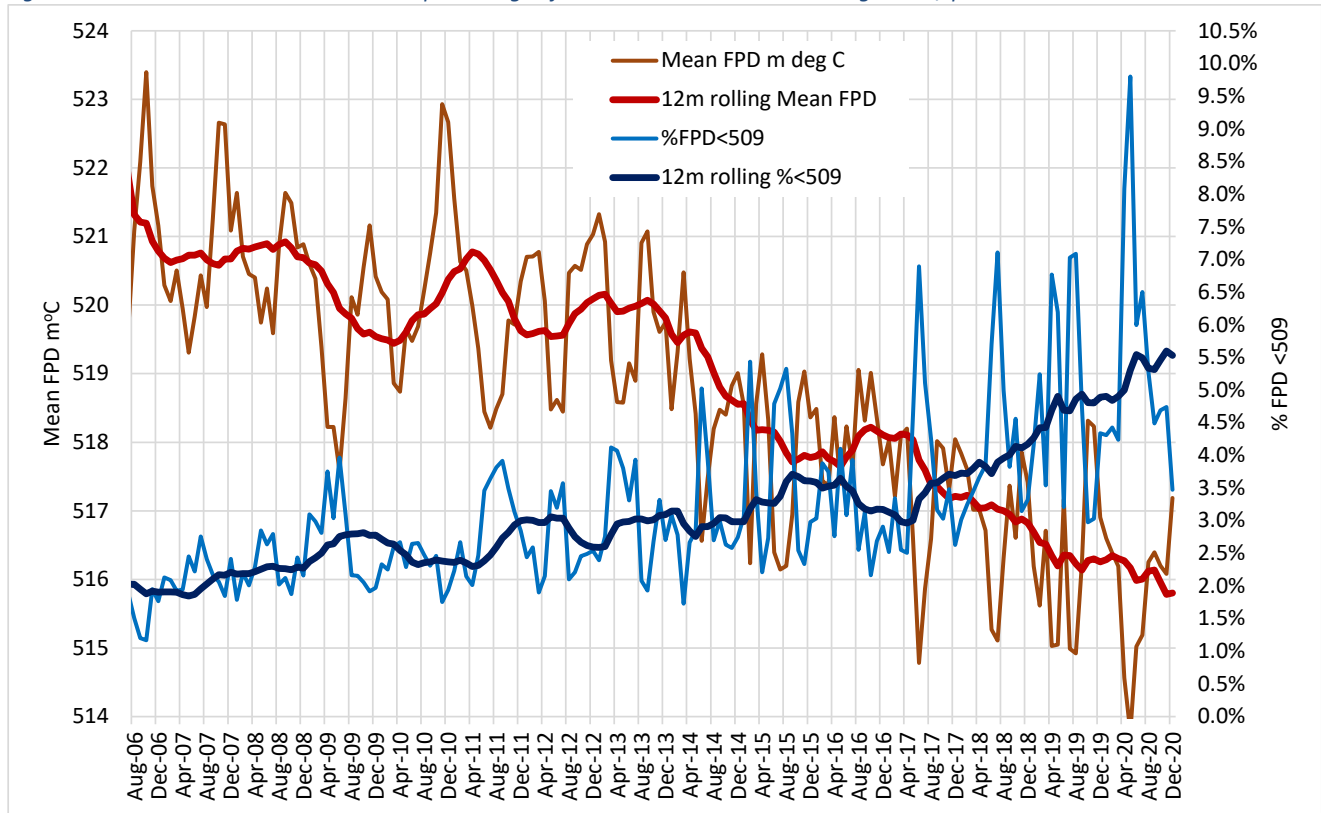
It seems somewhat bizarre that milk produced today is judged against standards established using a sample of 400 milk samples produced in 1977. The full University of Reading report presents in-depth descriptive analyses of the available FPD data for GB, with the aim of providing an objective and pragmatic basis for a revised standard, supported by the recent data. In recognition and acceptance that the milk being produced nowadays conforms to a 'new normal', with a lower mean FPD, a revision of the action point thresholds is needed. This revision must be made using up-to-date FPD data, which is available for routine milk samples collected and processed by National Milk Laboratories (NML) from virtually all milk producers in GB up to and including December 2020.

The data used were routine milk quality control samples processed by NML, the sole laboratory carrying out the required testing of milk from virtually all producers in GB. Routinely, the FPD of collected milk is assessed using a mid-infrared (MIR) estimation method (Combifoss). If the MIR 'screening' gives an estimated FPD of below 509 m°C the FPD is directly measured using cryoscopic instruments. The data analysed covers August 2006 onwards. This is when the full NML client list became established, covering almost all dairy producers in GB, and when the number of samples per month reached ~5,000. FPD values outside the Combifoss performance range (<450 m°C or >550 m°C) were excluded (<1% of samples from August 2006).

A seasonal pattern was observed in the data, with mean FPD lowest in June and July and highest through autumn and winter (September to February). As could be expected, the percentage of samples with FPD <509 follows an inverse pattern.

Figure 1 shows mean FPD value and percentage of FPD <509 each month, with 12 month (retrospective) rolling mean / percent lines to 'smooth out' the seasonal fluctuation.

Figure 1: Mean FPD value each month and percentage of FPD <509 with 12 month rolling mean / percent

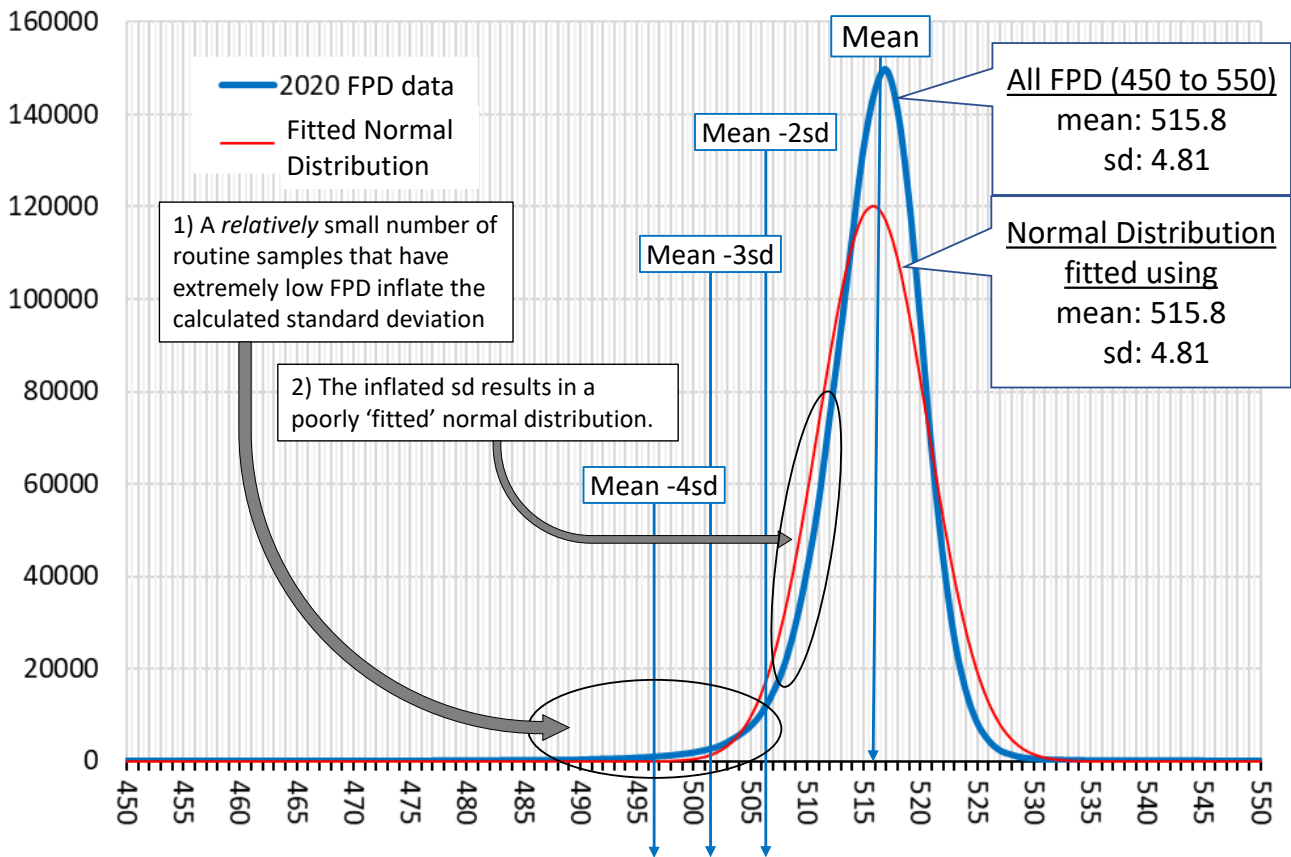


Trend analyses using linear regression indicated that:

- Mean FPD has been decreasing by 0.41 m°C per year since January 2014.
- The percentage of FPD values less than 509 has been increasing by 0.56% per year since January 2017.

The NML data are from routine milk collection samples as opposed to controlled 'authenticated milk' samples. In contrast to the 'authenticated' sample survey of 1977, these routinely collected NML samples represent a mixture of two populations: the majority genuine milk 'population' mixed with a small proportion of milk samples that have extraneous water. Inclusion of a small proportion of milk samples with extraneous water could be expected to 'fatten' the lower 'tail' of the distribution of FPDs in the NML data and inflate the standard deviation, as was indeed the case. Comparison of 'percentage points on a normal distribution' with actual percentiles in the real data distribution showed that the 'normal approximation' holds reasonably well down to close to the value 'mean-2sd'. However, below mean-2sd the real data deviate from a statistically normal distribution, with a 'fatter' and more stretched lower 'tail', as illustrated in Figure 2. This is because the routine NML samples inevitably contain some milk that contains extraneous water, leading to more extreme low FPD values than would fit a normal distribution.

Figure 2: Distribution of FPD of milk samples in range 450 to 550 (year 2020), with normal curve fitted using mean 515.8 and sd 4.81



In order to use routine NML data as a basis for revision of the action point thresholds adjustments are required to the approach used for the original BSI standard in view of the fact that standard deviation calculated from the routine NML data does not accurately reflect the distribution of 100% genuine milk.

Routine samples are likely to include a proportion of samples with extraneous water whereas the ‘authenticated’ milk samples of the 1977 survey were designed to exclude that possibility. Therefore, given that the intention is to summarise the distribution of FPD in ‘genuine’ milk, the proposed approach is that samples with FPD lower than an **agreed limit** should be excluded from the calculation when the mean and standard deviation are derived from routine samples. This is based on the understanding that the lower the FPD, the higher the probability that a sample contains extraneous water (is not ‘genuine’ milk). Table 1, using the routine NML data for 2020, shows that exclusion of low ‘outlier’ FPDs results in a substantial reduction in standard deviation with only a relatively small effect on the mean.

Table 1: Means and standard deviation calculated for FPD in the 2020 samples when different lower limits are set for exclusion of data

	n		Mean FPD	sd
Using all FPD 450 to 550	1,447,993		515.8	4.81
Additionally excluding*	n	% excluded[‡]	Mean FPD	sd
m-4sd (<497)	1,441,361	(0.5%)	515.9	4.40
m-3sd (<502)	1,433,456	(1.0%)	516.0	4.21
m-2sd (<507)	1,403,152	(3.1%)	516.3	3.89
m-~1.5sd (<509)	1,367,916	(5.5%)	516.5	3.68
m-1sd (<511)	1,297,629	(10.4%)	516.9	3.40

* ALL EXCLUDE FPD >550 (n=310 in 2020) and <450 (n=396 in 2020)

‡ compared with n=1,447,993

Setting revised thresholds using mean and sd of a subset of the routine FPD sample data, with a lower 'exclusion' limit on the data to be included in the calculation has the following logic:

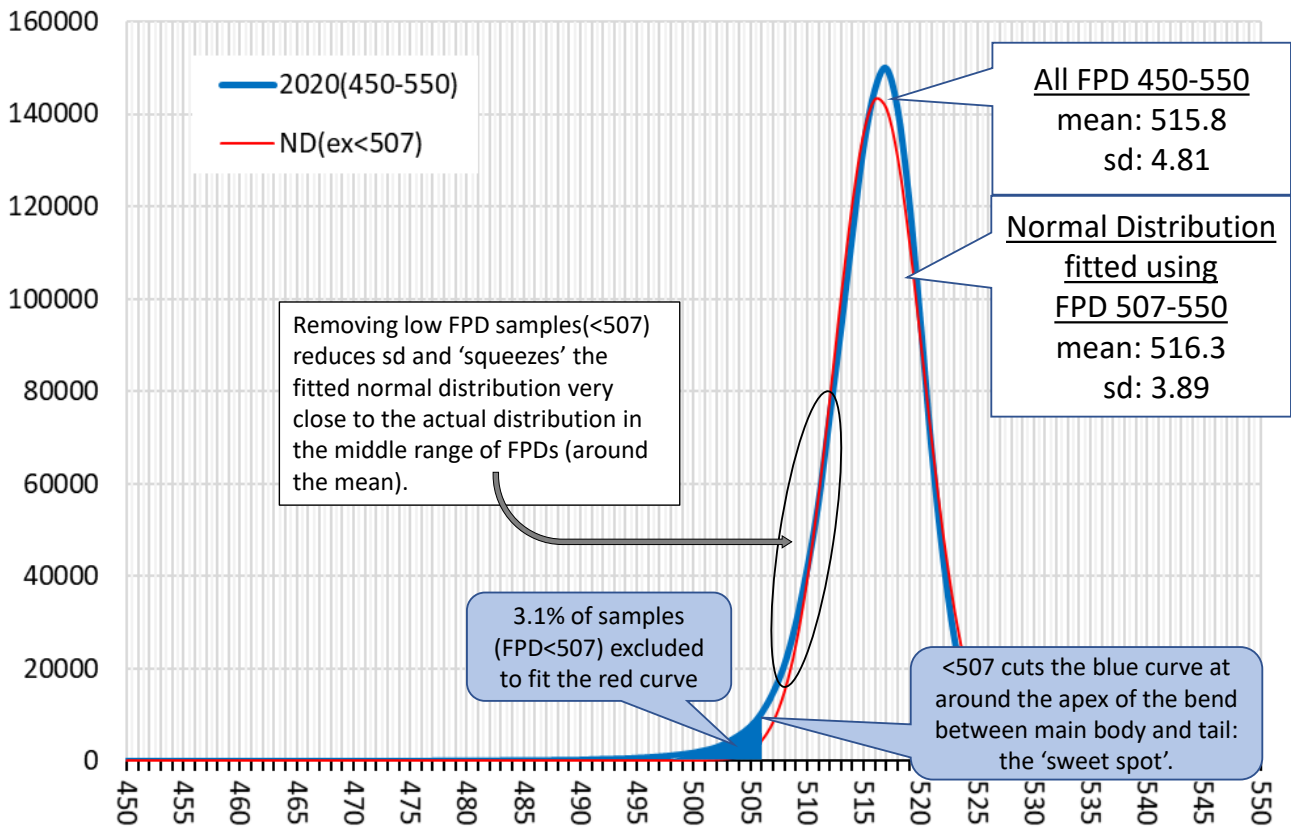
1. The routine samples include some 'non-genuine' milk with low FPD.
2. Therefore, 'clean' the data by excluding low FPDs (that have relatively high probability of containing extraneous water).
3. Use the cleaned dataset to give more accurate estimates of the mean and sd of FPD in genuine milk.
4. Action point thresholds can then be set based on these more accurate estimates using mean-2sd, mean-3sd and mean-4sd (as in BS 3095).

The lower limit for exclusion of FPD data is somewhat arbitrary, but could be set using multiples of sd below the mean (initially calculated on all routine samples). This will make future revision more transparent because if the mean FPD decreases, the lower exclusion limit, at [mean-'x'sd] will decrease accordingly. The only thing to decide is: *what is 'x'?*

The objective of the data cleaning exercise is to exclude as many as possible 'non-genuine' milk samples while at the same time **not** excluding genuine milk samples. Since the genuine and 'non-genuine' populations will likely have some overlap in FPD, this will always be a compromise. It is important to note that at this stage decisions are not being made about individual milk samples. This is a preliminary exercise to clean the routine data to achieve as accurate as possible estimates of the mean and sd of FPD in genuine milk. It does not matter that some genuine milk samples may be excluded and some non-genuine included. This should not have a serious impact on the mean and sd calculations. All samples will later be judged using the resulting thresholds.

As part of the University of Reading study, a graphic exploration was carried out to provide a visual impression of how different exclusion limits affect the mean and sd of FPD, and how the 'fit' of normal distribution changes. The best visual fit between the main body of data and a normal distribution appears when data are excluded for FPD <507 (~mean-2sd), as illustrated in Figure 3. In Figure 3, using the routine samples from the year 2020 as an example, the actual distribution of all FPD values within the Combifoss performance range (450 to 550) is shown as a blue curve. The red curve shows a normal distribution curve fitted using the mean and sd of the cleaned dataset. The blue shaded segment in the tail of the blue curve indicates the data that were excluded as 'outliers' when the exclusion limit was set at mean-2sd. The resulting 'cleaned' dataset now has a mean of 516.3 and sd of 3.89, and the action point threshold to be set at mean-3sd would be based on 504.6, which may be rounded up to 505.

Figure 3: Distribution of FPD of all routine milk samples within range 450 to 550, calendar year 2020, with normal curve fitted using data after exclusion of FPD <507 (~mean-2sd) for comparison



The thresholds could be set and monitored annually (and revised if considered necessary) based on the 12, 24 or 36 month rolling mean and standard deviation of FPD. There is currently nothing to choose between using 12, 24 or 36 month rolling statistics. Table 2 shows that using data from 12, 24 or 36 months up to December 2020 does not make much difference either to [mean-2sd] in the full datasets used to set the lower limits for exclusion of data, or to [mean-3sd] in the limited ('cleaned') datasets used to set the action point thresholds. **ALL would give 505 as the prime action point threshold (at mean-3sd, rounded up).**

The advantage of using multiple years' data to set, monitor and revise thresholds is that this reduces sensitivity to 'freak' years. The more years' data are used, the longer the delay is between real changes occurring in nature being reflected in any revised standard, but rolling more than one year's data together will help to ensure that any revisions made will reflect a sustained and well established change in nature. It can be argued that a sustained trend can hardly be recognised just by comparing two consecutive years, hence the use of three years' data could be recommended as the best option.

Table 2: Means and sd calculated for FPD in routine samples for 12, 24 and 36 months up to December 2020, when lower limits are set for exclusion of data, based on (rounded up) [mean-2sd] in the full datasets (450 to 550m°C)

Data used	n	% excluded [‡]	Mean FPD	sd	m-1sd	m-2sd	m-3sd	m-4sd
12 mth (all*)	1,447,993		515.8	4.81	511.0	506.2	501.4	496.6
24 mth (all*)	2,939,883		516.0	4.88	511.1	506.3	501.4	496.5
36 mth (all*)	4,450,904		516.3	4.92	511.4	506.5	501.5	496.6
12mth (>=507)	1,403,152	(3.1%)	516.3	3.89	512.4	508.5	504.6	500.7
24mth (>=507)	2,853,964	(2.9%)	516.5	4.00	512.5	508.5	504.5	500.5
36mth (>=507)	4,328,779	(2.7%)	516.7	4.06	512.7	508.6	504.5	500.5

*ALL EXCLUDE FPD >550 and <450

‡ compared with 'all*' for the relevant period

Assuming the use of 3 years' data (2018 to 2020), the steps to be taken to revise the FPD standard and thresholds are summarised below:

Step 1: Calculate mean and sd for all FPDs in range 450 to 550

Data used	n	Mean FPD	sd	m-1sd	m-2sd	m-3sd	m-4sd
All FPDs between 450 to 550 36 mth (2018, 2019, 2020)	4,450,904	516.3	4.92	511.4	506.5	501.5	496.6

Step 2: Use the value of mean-2sd, rounded up to next integer as the limit to exclude 'low outliers'. For the data being reviewed the mean-2sd is 506.5, so exclusion limit is 507 (exclude FPDs <507). This limited dataset is taken as consisting largely of 'genuine' milk.

Step 3: Re-calculate mean and sd in this 'limited' dataset.

Data used	n	Mean FPD	sd	m-1sd	m-2sd	m-3sd	m-4sd
FPDs >=507 to 550 36mth (2018, 2019, 2020)	4,328,779	516.7	4.06	512.7	508.6	504.5	500.5

Step 4: Use the values of mean-2sd, mean-3sd, mean-4sd, rounded up to next integer as the action point thresholds interpreted as BS 3095 (mean-3sd is 504.5, so threshold is 505).

	Action point thresholds	Interpretation (as BS 3095)
	>=509	"accepted free from extraneous water"
m-2sd	<509 (505 to 508)	"FPD alone is insufficient justification to presume it contains extraneous water"
m-3sd	<505 (501 to 504)	"likely that sample contains extraneous water"
m-4sd	<501	"sample almost certainly contains extraneous water"

APPENDIX 2: FREEZING POINT DEPRESSION MILK SCAN-CRYOSCOPE METHOD COMPARISON STUDY (February 2021)

This summary is taken from the full report on the MilkoScan-Cryoscope method comparison study, written by Andrew Dungey (andrewdu@nmrp.com), that presents data supporting the equivalence between the MilkoScan screening test method and the reference Cryoscope test method.

As part of the work to establish a revised industry standard threshold as an indicator of possible presence of extraneous water in milk based on Freezing Point Depression (FPD), a method comparison study was performed to demonstrate that FPD results from MilkoScan and Cryoscope instruments are comparable. The Foss MilkoScan instrument, as part of a Combifoss instrument set-up, is routinely used within NML laboratories to screen for FPD across a range of values. The MilkoScan method is used as a FPD screening method with the Cryoscope reference method being used to confirm screening results less than 509 m°C.

The Cryoscope reference method is based on *ISO 5764:2009 [IDF 108:2009] Milk — Determination of freezing point — Thermistor cryoscope method (Reference method)* The results are given in 'true' m°C.

The MilkoScan FPD test prediction model is developed with data from the entire mid-infrared spectrum plus an integrated conductivity sensor using multivariate techniques.

NML hold large amounts of data for FPD from the MilkoScan and Cryoscope instruments but any comparisons between the infrared/conductivity calibration model and reference method based on this data would be at less than 509 m°C. To aid in the comparison between test methods additional samples with a FPD of 509 m°C or greater were selected from storage based on their previous FPD results and the age of the sample being no more than 48 hours old. The criteria for the age of the sample were chosen to limit potential microbiological effects on FPD in relation to lactic acid/pH and to also fit in with current routine FPD screening practices.

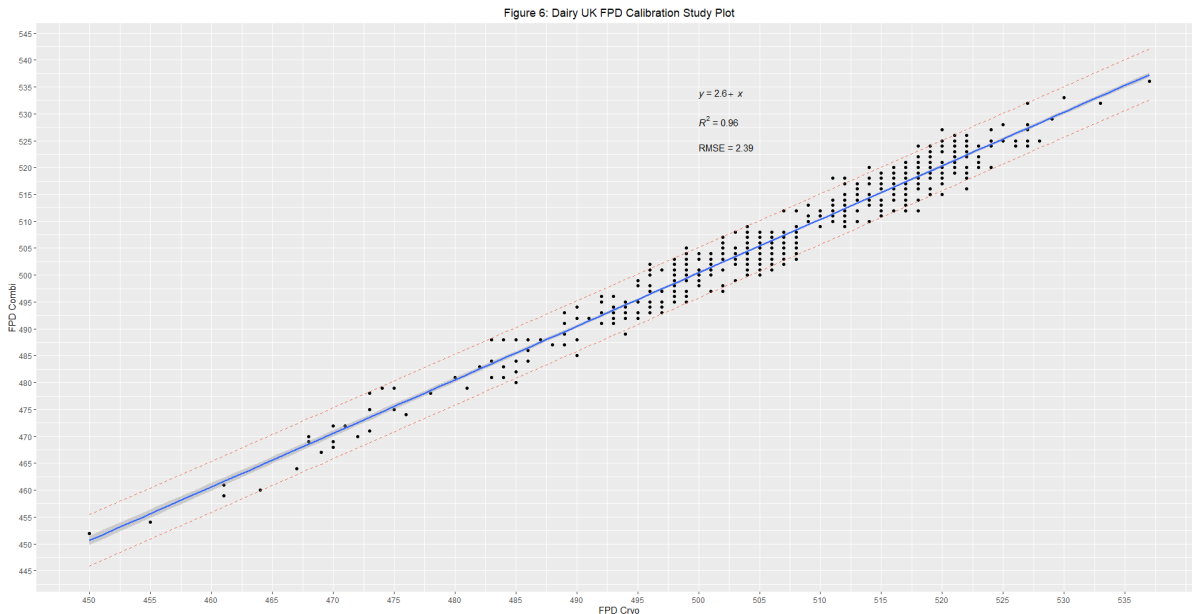
The geographic location of the selected samples was also taken into account and sampling without replacement was performed so that the same producer's samples weren't repeatedly selected. The Great Britain NUTS 2 statistical region data for each producer's samples were used in the selection algorithm.

A selection algorithm run from 21/11/2020 to 20/12/2020 involved stratified sampling where the selections were partitioned into suitable category levels ≤ 498 , 499-503, 504-508, 509-513, 514-518, 519-523, 524-528, ≥ 529 for each test site to enable a spread of data typically seen during routine FPD screening. A total of 719 usable samples were collected.

Once selected, samples with FPD screening results of greater than or equal to 509 m°C were tested on Cryoscope instruments at each laboratory site, replicating the routine FPD screening method where a single result from each method is used. The data from samples with FPD screening results less than 509 m°C already had Cryoscope test results from the previous night's testing so this data was used as found. The dataset was anonymized for milk purchaser and producer identification records.

As both laboratory sites used two Cryoscope and two Combifoss (MilkoScan part) instruments, the data from both sites were assessed to see if they could be combined into a single dataset for analysis by testing the difference between regression line slopes. An initial visual inspection of the regression plots followed by a comparison of slopes using an analysis of covariance showed that differences between slopes were not statistically significant and that the data were similar enough to combine into a single dataset.

A regression analysis was performed on the combined dataset. A plot of the data showing regression line and equation, confidence bounds, prediction bounds, R-squared and RMSE values was generated (see Figure 6 – taken from main report).



The regression model was assessed using the coefficient of determination, R^2 , and the residual standard error of the model.

The R^2 value represents the proportion of variation in 'FPD_Screening' that can be explained by the 'FPD_Cryoscope' data and will give some information about the goodness of fit of a model, it is a statistical measure of how well the regression predictions approximate the real data points. Values range between 0 and 1, with an R^2 of 1 indicating that the regression predictions perfectly fit the data. The value of 0.96 shows very good agreement between the regression predictions and the data.

The Residual standard error, RSE, is also used to assess the goodness of fit of a model. The RSE, corresponding to the prediction error, represents roughly the average difference between the observed outcome values and the predicted values by the model. The lower the RSE the better the model fits to the data.

In the dataset, for the predictor variable 'FPD_Cryoscope', the RSE = 2.392, meaning that the observed 'FPD_Screening' values deviate from the predicted values by approximately 2.392 units on average.

We can express the RSE as a percentage prediction error rate by dividing the RSE by the average value of the outcome variable. This corresponds to an error rate of $2.392/\text{mean}(\text{FPD_Screening}) = 2.392/508.4159 = 0.47\%$.

So, we can say that the MilkoScan FPD screening method accurately predicts the Cryoscope FPD reference method with a percentage error of 0.47% highlighting that the test methods compare well.