



# IWG BULLETIN

*IWG Bulletin Ref: 2000-6*

## Understanding the terms Accuracy, Precision & Bias

---

### Introduction

The terms **Accuracy**, **Precision** and **Bias** are important in understanding and interpreting objective wool measurement (both in laboratory and in field conditions) and when comparing one result to another.

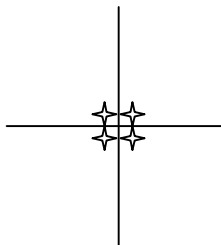
Small differences exist between instruments and operators even under ideal conditions for equipment, procedures and quality control.

- **Accuracy** is defined as how close you are on average to the reference value or “true” result.
- **Precision** is defined as the repeatability of a result and relates to how variable the results are compared to the average. This is often expressed as a 95% confidence limit or variance.
- **Bias** occurs when there is a consistent difference between the measured value and the “true” or reference value.

**Figure 1. The differences between Accuracy and Precision.**  
**Note:** The “true” result is the point where the two lines intersect

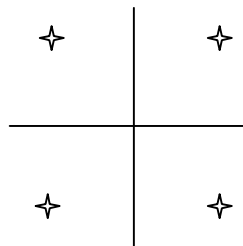
#### **Accurate AND Precise**

The four results are similar and close to the “true” result



#### **Accurate by NOT Precise**

The average of the 4 results is close to the “true” result, but individual results are scattered



### **EASTERN OFFICE**

Richard Manning  
Phone +61 2 6766 3077  
Fax +61 2 6766 6536  
Mobile 0419 466 014  
Email [rmanning@i-activewool.com.au](mailto:rmanning@i-activewool.com.au)  
132 North St  
Tamworth NSW 2340  
Australia

## **InteractiveWoolGroup**

ACN 082 549 835  
ABN 18 082 549 835

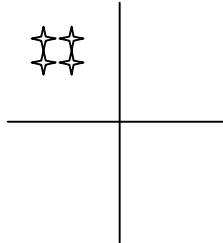
[www.i-activewool.com.au](http://www.i-activewool.com.au)

### **WESTERN OFFICE**

Bill Johnston  
Phone +61 8 9335 3495  
Fax +61 8 9430 5815  
Mobile 0407 986 105  
Email [wjohnston@i-activewool.com.au](mailto:wjohnston@i-activewool.com.au)  
4/65 Quarry St  
Fremantle WA 6160  
Australia

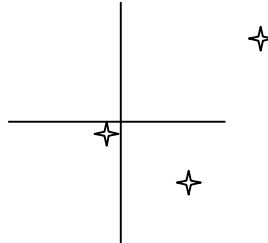
**Precise but NOT Accurate**

The average of the results is not close to the “true” result but the individual results are similar



**Inaccurate AND Imprecise**

The average of the results is not close to the “true” result and the individual results are scattered



**Accuracy**

True accuracy in wool measurement could be said to have three components

1. **Laboratory Accuracy** - Accuracy under the ideal conditions in a laboratory for the instruments, measurement procedures and quality control
2. **Practical Accuracy** – Accuracy achieved in practice by users with an instrument that is simple to operate, simple to check and to catch errors. The OFDA2000 has a high degree of practical accuracy.
3. **Sample source and identification** – Accuracy as it relates to sampling biases (see below), or mistyped or mishandled samples where the results can become meaningless. The OFDA2000 used for on-property wool testing minimises the chance of the latter errors, as the sheep or fleece is next to the instrument showing the result, and identification can easily be checked.

**The Precision of a Test Result and the 95% confidence limit**

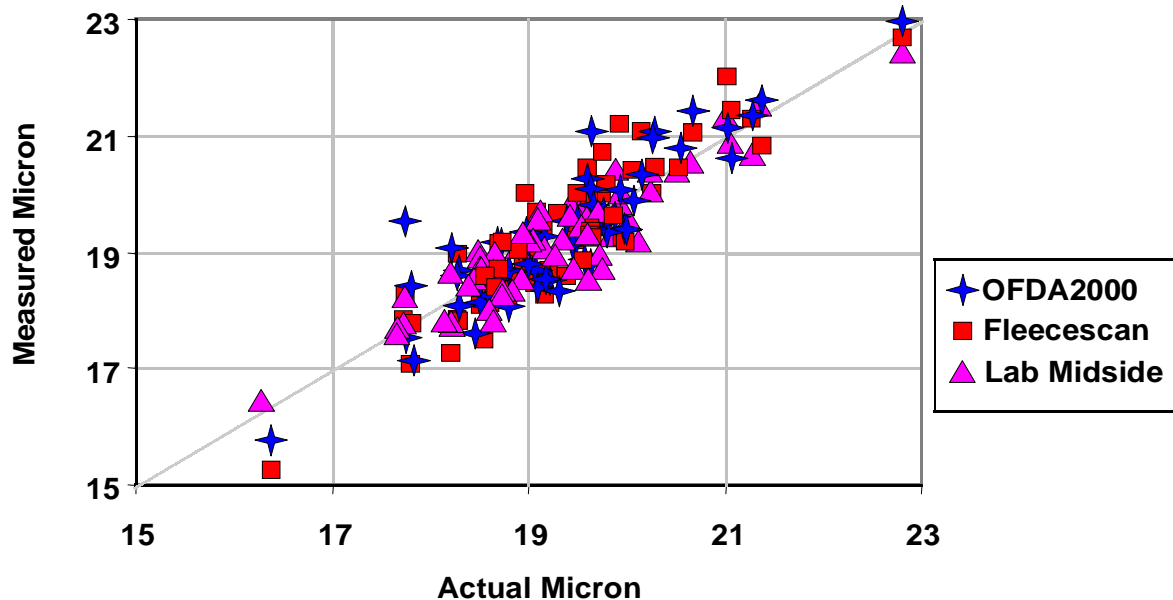
- The International Wool Textile Organisation (IWTO) test methods state that the precision of an individual wool test result should be expressed as a 95% confidence limit.
- The 95% confidence limit is the interval (+/-) on either side of the “true” test result within which you can expect 95 of 100 repeat measurements to lie.
- For example (as defined by IWTO test methods), the 95% confidence limit for 20 micron wool tested clean in the laboratory on OFDA100 or Laserscan instruments is +/- 0.4 microns for IWTO Test Certificates (i.e. repeat tests should fall between 19.6 microns and 20.4 microns, 95% of the time).
- Laboratories that test fleece samples (eg mid-side or hip-bone samples) for individual sheep have an average 95% confidence limit of +/- 1.1 microns for fibre diameter.
- The OFDA2000 has a 95% confidence limit of +/- 1.2 microns for fibre diameter at 20 µm.
- Therefore, the precision of the OFDA2000 measurement for mean fibre diameter (microns) of a single mid-side or hip-bone sample from a single animal should be as repeatable as a typical mid-side or hip-bone test carried out by a fleece testing laboratory as demonstrated in **Figure 2** on page 3.

---

*IWG would like to acknowledge Julia Harris of Innovative Wool Services, Longreach for the compilation of this IWG Bulletin*

Figure 2:

Source: Andrew Peterson, Department of Agriculture, WA

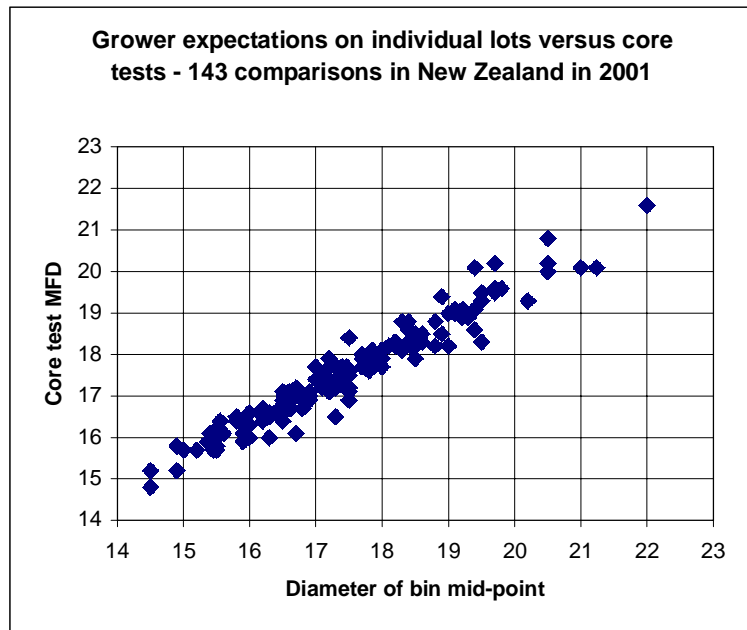


**What then is the “true” or “reference” value when testing individual sheep with OFDA2000?**

- In practice we must compare aggregated OFDA2000 results with IWTO core testing of complete lots of wool. (See figure 3 over – OFDA2000 vs Coretest).
- The accuracy of IWTO core testing is taken for granted, but in practice can be expected to be within +/- 0.2 micron in any one laboratory, and laboratory biases may drift slightly over time, due to calibration changes, etc.
- The published estimate of precision of MFD of a 5-bale line using OFDA2000 is approximately +/- 0.35 microns (which comes from the increased precision available when many measurements are combined).
- If no biases are present you can expect variation between core test and expected result within +/- 0.5 micron of combined results if every fleece is tested (due to the combined effects of the imprecision of the OFDA2000 result and of the core test result).
- If a small bias exists in either or both systems, it is possible to get overall differences of +/- 1.0 micron between expected results and core test result.
- Research has shown\* that fine bins core test coarser than predicted and coarse bins core test finer than predicted. This phenomenon occurs with ALL fleece testing systems. This bias can be as much as 0.5 microns. A factor that affects this is the variation in the fleece weights of the various fleeces put into each lot.

IWG would like to acknowledge Julia Harris of Innovative Wool Services, Longreach for the compilation of this IWG Bulletin

Figure 3:



**In summary, the key issues when comparing OFDA2000 results with test certificates**

1. **Imprecision / inaccuracy of certificates.** The accuracy of IWTO core testing is taken for granted, but in practice can be expected to be within +/- 0.2 micron in any one laboratory, and laboratory biases may drift slightly over time. The precision of the core test, due to random errors, is  $\pm 0.35$  micron.
2. **Sampling site bias** (mid-side v hipbone v whole fleece). A measurement of a mid-side or hip-bone sample from a fleece has been shown to be sufficiently accurate to rank sheep in a mob from the finest to the strongest, with the mid-side sample providing a slightly higher level of precision than the hip-bone. We expect that the average of the hip-bone site from a mob of sheep to be slightly stronger than the average of the mid-side and also slightly stronger than the average of the whole fleece.
3. **Effects of fleece weight/yield variation.** The MFD of all the fleeces in a bale is not simply the arithmetic average of the MFD of each of the individual fleeces. The relative weights of each fleece must be taken into account. Furthermore, fine fleeces contribute more to the overall mean than coarser fleeces of the same weight, simply because they contain more fibres. Each fleece going into the lot will also have a slightly different yield (ie greasy fleece weight vs clean fleece weight). \*\*
4. **Accuracy of shed record keeping for outclassed fleeces.** Remember that not all fleeces that are micron tested into a particular lot actually make it into that lot. For example, if you have decided that a lot of wool will contain fleeces that are tested finer than 17 microns, some fleeces matching this fibre diameter range may be outclassed into a tender, coloured or shorter line. A record of the number and micron of these fleeces should be kept so the correct average can be calculated for the line.

---

*IWG would like to acknowledge Julia Harris of Innovative Wool Services, Longreach for the compilation of this IWG Bulletin*

### Sampling Accuracy, Precision and Bias

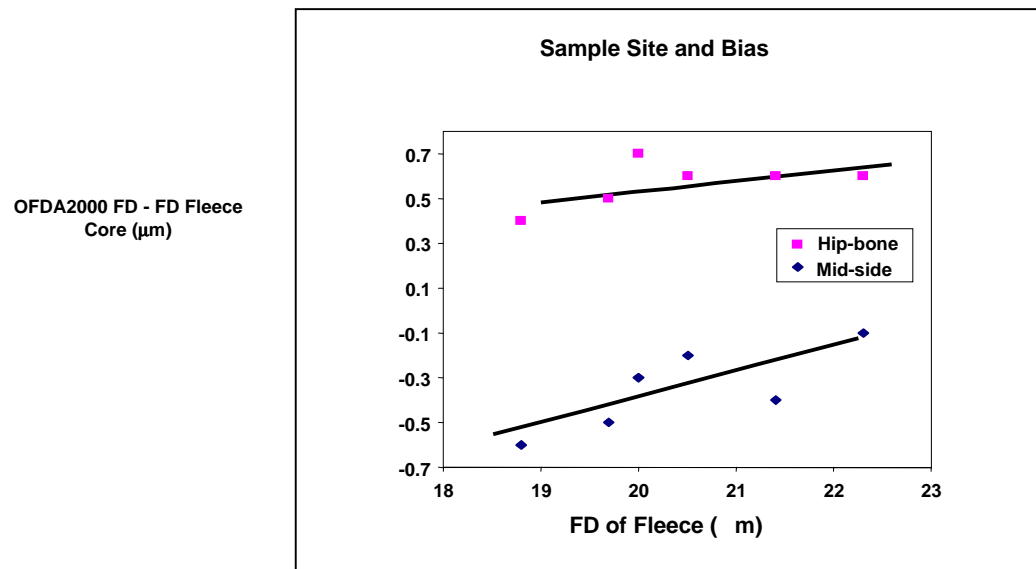
In all forms of wool testing, a measurement is generated from a **sample** of the fleece, bale or lot of wool being measured. The result from measuring the sample is used to **predict** the 'true' measurement of all of the fibres in the fleece, bale or lot of wool that would be obtained if it were possible to measure every single fibre.

Therefore, all methods of measurement involve errors generated in the sampling and preparation of samples. These are not mistakes, these are statistical variations.

In wool metrology, sampling and sample preparation can give rise to greater errors than the testing instrument itself. Therefore, care should be taken when locating the sample site on the sheep or fleece. Inconsistent sampling can affect the precision of the result, or create a bias.

Figure 4 shows some biases between sample sites and subsequent core test results determined in one trial, by way of example. The data in this figure should not, however, be taken as typical for all flocks.

**Figure 4:**

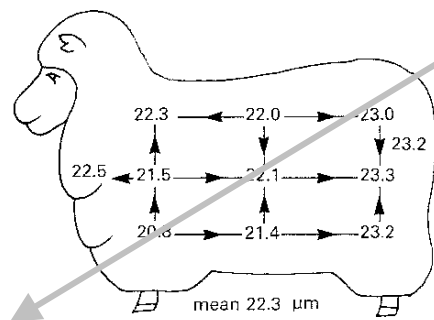


**Figure 4 Information courtesy Andrew Peterson, WA Department of Agriculture.**

**The two sites most commonly used to sample fleeces for OFDA2000 testing are:**

1. The **mid-side** site on the fleece is considered to be closest to that of the whole fleece fibre diameter (microns) This site is halfway down the side of the sheep (approx. a hand span on a hogget size sheep) and at the third rib from the end of the rib cage)
  - The **hip-bone** site is found at the points of the pelvic bone (off centre from the back bone) (often erroneously referred to as the 'pin' bone site) is on average slightly stronger than the average of the fleece, but this can vary between animals. The hip-bone site is easier to sample because of its location on the sheep.

**Figure 5** Shows an example of variation in fibre diameter (microns) across a sheep, and the location of the two most common sample sites. **Source: Andrew Peterson, Department of Agriculture, WA.**



**Conclusion**

In conclusion, a measurement of a mid-side or hip-bone sample from a fleece has been shown to be sufficiently accurate to rank sheep in a mob, from the finest to the strongest. The mid-side sample has provided a slightly higher level of precision than the hip-bone.

It is clear that it is possible to use a combined average of individual fleece measurements to predict the resultant core test of the lot of wool prepared from this information. Note the points discussed in this bulletin sheet that will affect the precision of your prediction.

\* Published data: Baxter, B. P. (2001). On-farm classing of animals & fleeces with the OFDA2000 Wool Technology and Sheep Breeding 49(2): 133-155; and Humphries, W., L. Vuckovic, W. A. Aspros, et al. (2001). Sirolan Fleecescan: Performance and Benefits of In-Shed Individual Fleece Testing. IWTO, CTF04, Shanghai

\*\* References: AWTA Ltd. Fact Sheet 013 (Classing using in-shed testing)

*IWG would like to acknowledge Julia Harris of Innovative Wool Services, Longreach for the compilation of this IWG Bulletin*