SECTION A – GENERAL INFORMATION

A.1. INTRODUCTION

A.1.1 For crops designated as 'minor' by the Plant Variety and Seeds Committee (PVSC) and the National List and Seeds Committee (NLSC), data review is by analysis of variance and comparison of candidates with control varieties, following standard statistical procedures and other approaches recommended by the Interdepartmental Statisticians Group (IDSG).

A.1.2 For other crops, data review is carried out by standard setting, a statistical approach developed by the IDSG for the NLSC. This compares candidates with varieties already on the National List.

SECTION B – DATA REVIEW, STANDARD SETTING AND DECISION MAKING

B.1. OVER-TRIALS MONITORING

This section applies only to characters agreed by the NLSC as being of major importance. It does not apply to minor crops.

B.1.1 Introduction

Variety means from trials harvested in the same year are monitored by examining a table of residuals for yield. A residual indicates the extent to which a variety in a particular trial performs better (+) or less well (-) than expected as judged by its overall performance. Expected performance is estimated from the variety mean over all trials and the mean of all varieties in the particular trial.

To make it easier to assess quickly their significance, residuals are standardised by dividing by their root mean square.

B.1.2 Data loss

Trials data may be lost for a number of reasons, e.g. mishaps, non-occurrence of a disease. In the case of major VCU characters, such as yield, it may not be possible to make a VCU decision if there are insufficient reliable data. Factors to be considered in determining that the data are adequate include:

- The precision of results from the reduced across-trials data should be no less than half the precision expected of the complete trial series.
- Data are present from at least one centre in both sowing years and from another centre (in the case of multi-centre trials) in one of the two years.

If a pattern in the residuals for a variety can be linked with a feature of the centres, e.g. water availability, then this must be reported to the National List and Seeds Committee, which may wish to consider this when making its decisions.

B.1.3 Validation of trials

While an individual residual reflects the performance of one variety in a particular trial, the group of residuals associated with a trial can tell something about the trial as a whole. The residuals are examined in a number of ways:

B.1.3.1 Variability ratio

The root mean square of the residuals for a trial, expressed as a ratio to the root mean square of the residuals in the rest of the table, is a measure of how variable varieties have been in a trial relative to their performance in other trials. The average ratio will be approximately 1.0 and a trial with above average variation will be indicated by a ratio greater than 1.0. The square of the ratio is distributed as a variance...
ratio (F) with (no of varieties-1) and (no of varieties-1) x (no of trials-2) degrees of freedom. For four trials the critical ratio, at the 1% probability level, depends on the number of varieties in the table, as follows:

<table>
<thead>
<tr>
<th>No. of varieties:</th>
<th>10</th>
<th>20</th>
<th>40</th>
<th>60</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD ratio criterion:</td>
<td>1.84</td>
<td>1.54</td>
<td>1.36</td>
<td>1.29</td>
<td>1.22</td>
</tr>
</tbody>
</table>

B.1.3.2 **Coefficient of correlation**

The coefficient of correlation between variety yields in the trial and the variety means over other trials is used to assess how closely the yields in the trial follow average yields from all other trials. From experience, correlations of r=0.50 or more are to be expected from yield trials of broadly similar background. Critical values for testing whether there is statistical evidence (p=0.01) that the “true value” of the correlation is less than 0.5 are derived from the confidence interval for the sample correlation. These critical values depend on the number of varieties:

<table>
<thead>
<tr>
<th>No. of varieties:</th>
<th>30</th>
<th>40</th>
<th>60</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum acceptable:</td>
<td>0.04</td>
<td>0.12</td>
<td>0.20</td>
<td>0.28</td>
</tr>
</tbody>
</table>

If either of checks B.1.3.1 and B.1.3.2 are not fulfilled, then the Data Review and Standard Setting Operator should confirm that the Trial Inspection and Technical Validation Operator and the Data Handling Operator are satisfied that the trial is providing useful information, with reference to the within-trials monitoring checks. Otherwise, the results for the trial should be considered for exclusion by the National List and Seeds Committee, by correspondence, before the final data analysis takes place.

B.1.4 **Checking variety consistency**

The Variability ratio for a variety can be used to identify greater than average consistency (or instability) in a variety's performance, in a manner similar to that used when validating trials.

B.1.5 **Non-yield characters**

The above checks may not be completely suitable for all characteristics, particularly those which are strongly influenced by trial conditions. In these cases, other checks may be carried out. For disease characteristics, the performance of the control varieties should be monitored. This should indicate whether there is adequate differentiation between susceptible and resistant varieties and how consistent is the disease pressure from trial to trial and year to year.

B.1.6 **Recording the decision to exclude**

The reason for excluding an aberrant trial at this stage is recorded by the Data Review and Standard Setting Operator.

**B.2. STEPS IN PRODUCING THE VCU STANDARDS AND SCORES FOR THE CANDIDATES**

Scores are calculated for each candidate variety by comparison with standards. A decision is made based on these scores. This does not apply to minor crops.

B.2.1 **Producing standards**

Superior and inferior standards are produced for each characteristic. There are three methods permissible. The choice of method should be agreed by the technical experts and approved by the NLSC in advance of decision-making.

B.2.1.1 **Method A: Standards are based on a statistical comparison with the mean of National List**

In this method, the superior and inferior standards are produced by making a statistical comparison with the mean of the National List varieties.
An appropriate over-year analysis is carried out, allowing estimation of National List and candidate variety means along with standard errors. The number of years of data included in the analysis will depend on the crop, with the aim of including as many of the National List as possible whilst ensuring that trial practices and conditions are reasonably consistent. Data may be included from other sources, such as recommended list trials, with NLSC approval. There is, statistically, benefit in including all varieties trialled, including former candidates, although means will not be required for all varieties. An appropriate linear mixed model should be used for the analysis, using REML.

The superior and inferior standards are calculated by adding and subtracting a critical difference to the mean of the National List varieties that have been tested. The critical difference is found by multiplying the one-sided student t-statistic with 20 degrees of freedom at 5% (this has a value of 1.725) by the square root of the mean of the variances for the differences between each of the candidates with the mean of the National List varieties.

This method will be used for most characteristics in most crops.

B 2.1.2 Method B: Standards are set by the technical experts based on their knowledge of the crop

Here the standards are set by the technical experts. This method can be used in cases where the relationship between the characteristic and the value to the grower can be established in advance. An example is for the _Globodera rostochiensis_ evaluation for potato where a single gene confers resistance and the characteristic score on a 1-9 scale can be interpreted in terms of presence and absence of this gene. For this method it is only necessary to carry out an analysis with the current two years of trials.

B 2.1.3 Method C: Standards based on a statistical comparison of the candidate with a fixed level set by the technical experts

This is similar to method A, except that comparisons are made with a fixed level rather than with the mean of the National List varieties. This level must be set in advance by the technical experts.

The analysis is based on the current two years of trials only. This will normally be carried out using a mixed model with REML. The superior and inferior standards are calculated by adding and subtracting a critical difference to the fixed level. The critical difference is found by multiplying the one-sided student t-statistic with 20 degrees of freedom at 5% (this has a value of 1.725) by the square root of the mean of the variances for the differences between each of the candidates with the fixed level.

This method may be used when method A is thought to be inappropriate due to the nature of the characteristic or for new characteristics where only a small proportion of National List varieties have been tested.

B.2.2 VCU variety categories

Where varieties are grouped according to their nature, then separate standards are calculated for each group. If the different groups are assessed under different conditions, such as separate trials or different harvest dates, or if there are substantial differences in behaviour between the groups then the calculation of standards should be based on completely separate analyses. This should be agreed in advance by the technical experts and a representative from the IDSG. Otherwise the data from each group can be combined in one mixed model analysis. The standards for each group are then calculated separately based on this single analysis.

B.2.3 Scoring of varietal performance based on the standards

The scoring scheme is used to judge the performance of a candidate variety in a characteristic. It will be used to balance the performances over the set of characteristics. The choice of scoring scheme should be agreed by the technical experts and approved by the NLSC in advance of decision-making.

Either a step-scoring scheme or a continuous-scoring scheme may be adopted for a characteristic. In the step scoring scheme, a set positive score is given to candidate varieties with performances that are equal to
or better than the superior standard and a negative score is given for performances that are worse than the inferior standard. Candidates with performances between the standards are given a score of zero. The continuous-scoring scheme differs by varying the score assigned to performances between the standards linearly according to the level of performance so as to interpolate the scores set at the standards. Variations on either scheme are permissible, given approval by the IDSG and NLSC. For example, negative scores for a characteristic might be converted to zero.

B.3. DECISION MAKING

B.3.1 Common Catalogue Directives

Decision making for National Listing is made in accordance with Common Catalogue Directives, which establish the need for a new variety to show satisfactory value for cultivation and use in relation to varieties already on the National List in the Member State. Recommendations are usually made by the technical experts with decisions made by the NLSC.

B.3.2 Decision making where standard setting is used

Where standard setting has been used and scores allocated, the scores can provide a basis for decision making. Technical experts are free to use rules as to how to use these scores but these must be agreed in advance and confirmed by the NLSC.

The simplest way of using the scores is to sum them over characteristics. Candidates with scores that are greater than a set threshold may be given a recommendation to accept. This threshold would normally be positive. Candidates with scores that are less than a set threshold may be given a recommendation to reject. This threshold would normally be zero. Other candidates between the two thresholds require further consideration, perhaps taking into account other aspects of performance.

Other more complex ways of using the scores are possible. These should be agreed by the NLSC under guidance from the IDSG. It is also possible to use the scores for individual characteristics as a basis for discussion without considering the total or setting down formal rules.
OPERATIONAL ASPECTS OF DECISION MAKING

Relative importance of each VCU characteristic

Except for minor crops, the characters are apportioned into one of the following levels of importance:

- Major
- Intermediate
- Disease resistance - scores for individual diseases are based on the relative importance of the disease and how difficult it is to breed resistance to it.
- “Other” –
  - characters may be in this section because:
    - They have only little or no importance for routine considerations
    - The data are not sufficiently robust for routine decisions to be made
    - They have a minimum standard for acceptance but no potential for consideration as improvements

Some characters may be considered as major weaknesses if below the weakness standard but only intermediate improvements if above the clear improvement standard. The converse may also be appropriate for some characters.

Disease resistance for an individual disease is generally considered to be of intermediate importance. However, where the resistance is very low (rating of 2) it is considered to be a major weakness. Where the disease rating is 1, this will result in failure of the variety for this characteristic alone. High levels of resistance to individual diseases may be awarded a higher score as well as a bonus score for an overall high level of resistance to all diseases tested.

Preliminary Recommendation

Preliminary recommendation by the technical experts is made on a quantitative basis according to the balance of superior and inferior characters and their relative importance. The following example with winter wheat indicates how the characters are balanced in a quantitative way:
<table>
<thead>
<tr>
<th>Character</th>
<th>Relative importance</th>
<th>Example Variety</th>
<th>Score for Character</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yield (Treated)</td>
<td>Major</td>
<td>Clear Weakness</td>
<td>-2</td>
</tr>
<tr>
<td>Yield (Untreated)</td>
<td>Intermediate</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>Resistance to Lodging (Treated)</td>
<td>Intermediate</td>
<td>Clear Weakness</td>
<td>-1</td>
</tr>
<tr>
<td>Resistance to Lodging (Untreated)</td>
<td>Intermediate</td>
<td>Clear Improvement</td>
<td>+1</td>
</tr>
<tr>
<td>Specific Weight</td>
<td>Intermediate</td>
<td>Clear Improvement</td>
<td>+1</td>
</tr>
<tr>
<td>Disease Resistance - Mildew</td>
<td>Disease scale</td>
<td>Clear Improvement</td>
<td>+ 0.5</td>
</tr>
<tr>
<td>Yellow Rust</td>
<td>as above</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>Brown Rust</td>
<td>as above</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>S. nodorum (winter wheat only)</td>
<td>as above</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>S. tritici (winter wheat only)</td>
<td>as above</td>
<td>Clear Improvement</td>
<td>+ 1.5</td>
</tr>
<tr>
<td>Eyespot (winter wheat only)</td>
<td>as above</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td><strong>BALANCE (Total)</strong></td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

A positive total score does not necessarily indicate a clear improvement.

There are three possible balances:

1. **Positive Balance:** Preliminary Recommendation is for acceptance.
2. **Negative Balance:** Preliminary Recommendation is for refusal.
3. **Zero Balance:** No Preliminary Recommendation.

The preliminary recommendation is based on a quantitative balance of characters with standards.

**REVIEW OF PRELIMINARY RECOMMENDATION**

The preliminary recommendation is then reviewed to consider if there are any reasons for overturning it. This review will involve consideration on a qualitative basis to ensure that other issues including the following have been taken into account:

- Performance in a character substantially beyond the superior/inferior standard, which would alter the assessment of the qualities of the variety.
- Performance beyond the superior/inferior standards in a number of characters which, although individually not exceptional, when considered together constitute a clear improvement / weakness beyond the qualitative balance of superior and inferior characters.
- A number of improvements / weaknesses (compared with the National List or category median) in individual characters which do not meet the standards needed to be considered superior or inferior characters, but when combined would amount to a clear improvement / weakness in terms of the qualities of the variety as a whole.
- Consideration of other characters for which no standards have been produced to check for superior or inferior performance.

Issued October 2008
• Consideration of regional performance.

• The applicant may have requested the recording of additional characters as detailed in the relevant Protocol and Procedures. The level of importance of these characters will be determined on an individual basis.