Performance Monitoring System for Post-SCT Chimerism

Aim

The scoring system is a rolling scheme that will identify unsatisfactory performance or persistent unsatisfactory performance of any participant. This is in order that UK NEQAS for Leucocyte Immunophenotyping (LI) can provide support and guidance where needed and ensure that the Genetics National Quality Assurance Advisory Panel (NQAAP) are informed as appropriate. Please note that each programme is scored independently. A linked document provides a hypothetical example to help describe the performance monitoring system used in this programme

Outline

Two post-stem cell transplantation (SCT) samples are issued at each trial distribution with varying levels of donor chimerism. There are a maximum of 5 trials per annum. Chimerism is measured using molecular techniques and is a quantitative measurement of **donor** engraftment (maximum 100% engraftment), which allows for the serial monitoring of patients post-transplant.

The scoring system is a quantitative approach for which participants are asked to produce a percentage **donor** engraftment result using their normal laboratory technique. Percentage donor chimerism results are to be reported to UKNEQAS LI as integers, in order to comply with technical recommendations (Clark et al., 2014).

A participant's submitted result for each post-SCT sample is then used in conjunction with the robust mean and robust standard deviation to calculate a z-score using the following formula:

$$z = (x - X)/\hat{\sigma}$$

where x is the result returned by the testing laboratory, X is the assigned value (robust mean) and $\hat{\sigma}$ is the standard deviation for proficiency assessment (robust SD).

The robust mean and robust SD are derived from participant data using Algorithm A (ISO 5725-5) that ensures that all data is included in the generation of the robust mean and robust SD but also minimizes the effect of outliers upon the final values. The robust mean and SD are calculated to 1 decimal place (d.p.).

Interpretation of z-scores in the context of this programme is as follows:

- A result between 2.5 and -2.5 will be classed as satisfactory.
- A result between 3.5 and 2.5 or -2.5 and -3.5 is seen as an 'action' result. This highlights a potential issue to the laboratory.
- A result above 3.5 or below -3.5 is considered to be a 'critical' result requiring immediate investigation by the laboratory.
- Two 'action' results in a period of any three consecutive samples will also result in a classification of 'critical'. Please note that whilst each 'action' will be considered in combination with any other 'actions' within rolling 3-sample periods, the same pair of 'action' scores will not be combined to result in a critical classification more than once (please see the linked document describing a hypothetical example of performance monitoring if further clarification of this is required).

Due to the nature of how z-scores are generated a positive z-score highlights a positive bias in a laboratory's results whereas a negative z-score shows a negative bias. As such, this adds value to the performance monitoring information provided to laboratories because the z-score immediately highlights to the participating centre if their result is above or below the expected consensus value. The normalisation of z-scores using the robust SD allows the direct comparison of performance across multiple samples/trials issues.

In addition to the z-score, methodological data featured on reports will be in the format of robust mean and robust SD. This will give participants the option to use the extra provided data to calculate additional 'in-house' z-scores based on machine types, methodologies etc and allow them to monitor if there are any 'in-house' technical biases. However, it is important to stress that the z-score issued by UK NEQAS for LI based on all methods will remain the only parameter that is used for performance monitoring.

Any laboratory who fails to return a result by the closing date will be regarded as an action for each sample. As such any laboratories that do not return results for both samples within a trial will be classified as critical. Please note, results should not be submitted if samples fail internal quality control measures. Repeat samples are available for all trials, if required. If following repeat sample(s) processing, results obtained still do not pass local internal QC please contact UK NEQAS LI. If results are submitted based on suboptimal results, they will be subjected to the same performance monitoring mechanisms as all other participants.

Unsatisfactory performance in this programme is defined as any occurrence of critical performance and this will be initially communicated to participants on their trial report. This will be followed up with an email and notification on the Participant Hub on each occurrence of unsatisfactory performance, highlighting that performance on the last sample(s) was out of consensus and offering support and guidance to assist in returning to satisfactory performance. The support and guidance offered will be tailored to the particular needs of the participant but may include the provision of repeat/additional samples, communications by email, telephone conversations or face to face communications.

If a participant amasses three critical performances within a three-trial period, their status is elevated to persistent unsatisfactory performance. A further email will then be issued, and, for UK participants, the Genetics National Quality Assurance Advisory Panel (NQAAP) informed.

Participant's results will be reviewed by the lead scientist and the participant may, at the discretion of the Director and Specialist Advisory Group chairperson, be referred to the Genetics NQAAP even if they have not met the criteria for Persistent Unsatisfactory Performance in any individual EQA.

As with all scoring systems it is important to note that the limits will be constantly reviewed to determine whether they are providing the information required. The Director of the programme retains the right to determine if an individual trial should not be scored.

Reference

Clark, J. R., Scott, S. D., Jack, A. L., Lee, H., Mason, J., Carter, G. I., … Barnett, D. (2014). Monitoring of chimerism following allogeneic haematopoietic stem cell transplantation (HSCT): Technical recommendations for the use of Short Tandem Repeat (STR) based techniques, on behalf of the United Kingdom National External Quality Assessment Service. *British journal of haematology*. doi:10.1111/bjh.13073