



anses

GUIDE MÉTHODOLOGIQUE & STATISTIQUE RELATIF AUX ESSAIS INTER-LABORATOIRES D'APTITUDE

Référence ANSES/PG/0029
Version d (2022) – Amendement juin 2024 (*)

(*) Amendement de la version 2022 : (i) mise à jour du S-score avec prise en compte de la publication Baudry et al, 2024 (Accreditation and Quality Assurance), et (ii) prise en compte de la version 2022 de la Norme NF ISO 13528.

Ce guide est à l'usage des laboratoires de l'Anses pour l'organisation d'essais inter-laboratoires d'aptitude (EILA), dans le cadre de leurs missions de laboratoires de référence nationaux et de l'Union européenne.

Il s'attache à recommander le type et le nombre d'entité soumise à essai, plans d'expérience et analyses statistiques pour (i) l'étude de l'homogénéité et de la stabilité des échantillons, et (ii) l'évaluation des performances des participants aux EILA.

En complément des recommandations expérimentales et statistiques, ce guide donne également des éléments de méthodologie pour la mise en œuvre d'un EILA dans le contexte spécifique des laboratoires de l'Anses et pour la rédaction des rapports d'EILA.

Sommaire

Introduction.....	6
1 Domaine d'application	7
2 Logiciels et outils de calcul.....	7
3 Prérequis à l'organisation d'EILA.....	8
3.1 Considérations générales	8
3.2 Informations sur la performance des méthodes d'analyse.....	9
3.2.1 Méthode(s) utilisée(s) par l'organisateur.....	10
3.2.2 Méthode(s) utilisée(s) par les laboratoires participants.....	10
4 Rappel du contexte pour le traitement des données	11
5 Contexte de l'EILA et plan expérimental	12
5.1 Description de la variable sur laquelle l'aptitude est évaluée.....	12
5.2 Classification de la variable.....	13
5.3 Distribution des résultats (Normale, Log-normale, Binomiale, ...)	13
5.4 Transformation des données avant la réalisation des traitements statistiques	14
5.5 Typologie et nombre des participants.....	14
5.6 Nature des échantillons.....	14
5.7 Nombre et type de résultats fournis par les participants	15
6 Homogénéité et stabilité des ESEA.....	15
6.1 Homogénéité des ESEA	15
6.1.1 Aspects généraux	15
6.1.2 Cas des EILA quantitatifs.....	17
6.1.3 Cas des EILA qualitatifs	20
6.2 Stabilité des ESEA.....	22
6.2.1 Aspects généraux	22
6.2.2 Cas des EILA quantitatifs.....	23
6.2.3 Cas des EILA qualitatifs	25

7	Prétraitement des résultats	26
7.1	Examen initial des informations et des données fournies par chaque participant	26
7.2	Vérification de la cohérence de l'EILA	27
7.2.1	Objectif	27
7.2.2	Méthode	27
7.2.3	Conclusion	28
7.3	Gestion des données manquantes	28
7.4	Prise en compte des données atypiques	28
8	Évaluation de la performance des laboratoires participants	29
8.1	Argumentaire pour l'utilisation d'une approche robuste	29
8.2	Détermination de la valeur assignée	30
8.2.1	EILA quantitatifs	30
8.2.2	EILA qualitatifs	30
8.3	Détermination de l'écart-type d'aptitude	31
8.3.1	EILA quantitatifs	31
8.3.1	EILA qualitatifs	31
8.4	Évaluation de la performance des participants	32
8.4.1	Estimation de la justesse	32
8.4.2	Estimation de la fidélité	41
8.4.3	Représentation graphique des performances	41
9	Vérification de la cohérence inter-EILA	42
10	Présentation des résultats	42
10.1	Rapport d'EILA	42
10.2	Autres données ou modalités de présentation des résultats	44
10.3	Suivi des actions correctives suite à l'EILA	44
11	Conclusions	44
	Annexe A : Homogénéité - : Calculs réalisés avec Prolab	45
A.1	Aspects généraux	45
A.2	Calculs selon "Norme NF ISO 13528:2015"	46
A.3	Calculs selon "Protocole harmonisé"	49
A.4	Rapports	49
A.5	Exemple de jeu de données	51

Annexe B : Homogénéité selon l'approche IUPAC.....	52
Annexe C : Homogénéité - Cas où le nombre d'ESEA par niveau est faible.....	53
Annexe D : Stabilité - Calculs réalisés avec Prolab.....	54
D.1 Aspects généraux.....	54
D.2 Calculs.....	54
D.3 Rapport.....	55
D.4 Exemple de jeu de données	57
Annexe E : Homogénéité et stabilité - Calculs réalisés avec QualiScore.....	58
E.1 Exemple de calculs d'homogénéité.....	58
E.2 Exemple de calculs de stabilité	58
Annexe F : Exemple de calcul des scores z et z' avec le logiciel Prolab.....	59
F.1 Valeur assignée et ET aptitude	59
F.2 Calcul des scores	62
F.3 Visualisation des résultats	64
F.4 Edition de résultats	66
F.4.1 Aspects généraux	66
F.4.2 score z.....	67
F.4.3 score z'.....	67
F.4.4 Exemple de jeu de données	69
Annexe G : Exemples de calcul du L-score (Prolab).....	70
G.1 Aspects généraux.....	70
G.2 Calculs.....	71
Annexe H : Principe et calcul du S-score	73
H.1 Principe	73
H.2 Calcul du S-score	74
Annexe I : Exemples de calcul du S-score pour les EILA qualitatifs	76
I.1 Exemple n°1 : panel composé des 4 types d'ESEA	76
I.2 Exemple n°2 : utilisation des ESEA2 non répliquées	78
Annexe J : Tableaux pour le calcul du S-score	81
Annexe K : Calcul du score z composé.....	84
Bibliographie.....	85

Introduction

Cette version du guide (avril 2022) représente une mise à jour générale de la version b (13/03/2017). Elle prend en compte les évolutions de la dernière version de la norme générale NF ISO 13528 (octobre 2015) portant sur les méthodes statistiques utilisées dans les essais inter-laboratoires d'aptitude (EILA), ainsi que le retour d'expérience des laboratoires de l'Agence qui organisent des EILA en utilisant la version précédente de ce guide.

Cette version a été préparée par le GT STAT EILA 2 :

- animé par Michel LAURENTIE, responsable de la Plate-forme nationale méthodologique d'Appuis Statistiques pour les activités de référence (PAS, ptf.pas@anses.fr),
- et comprenant :
 - les référents statistiques de la plateforme PAS dans les laboratoires : Christian BAUDRY et Stéphanie BOUGEARD (Laboratoire de Ploufragan-Plouzané-Niort), Aude CHABIRAND (Laboratoire de Santé des Végétaux), Maud MARSOT (Laboratoire de Santé Animale), Jean-Sébastien PY (Laboratoire d'hydrologie de Nancy), Jean-Luc SCHEREFFER (Laboratoire de la rage et de la faune sauvage de Nancy), Bertrand LOMBARD (ex-Laboratoire de sécurité des aliments),
 - un représentant de la Direction de l'appui au pilotage, de la qualité et de l'audit (DAPQA, Catherine DE MÈREDIEU),
 - un représentant de la Direction de la stratégie et des programmes (DSP, Barbara GOUGET puis Bertrand LOMBARD).

Dans un amendement au guide, la partie qualitative a été modifiée, utilisant le S-score développé par l'Anses et associé à l'outil de calcul Qualiscore a été mise à jour par Christian BAUDRY, Gaëlle JADÉ et Michel LAURENTIE à la suite de la publication dans *Accreditation and Quality Assurance* de l'article "S-score: a new score for binary qualitative proficiency testing schemes interpretable as the z-score" (février 2024) :

- Évaluation de l'homogénéité (6.1.3.2) et de la stabilité (6.2.3) des échantillons.
- Évaluation de la performance des participants (8.4.1.1.3.b).

Dans cet amendement, une mise à jour a été également effectuée pour tenir compte de la nouvelle version de la norme NF ISO 13528 (octobre 2022).

La mise à jour concernant l'évolution de la norme NF ISO 17043 sera faite ultérieurement.

1 Domaine d'application

Ce guide est destiné aux organisateurs d'EILA, des laboratoires de l'Anses, dans le contexte spécifique de leurs missions de référence (LNR, LRUE). Il donne principalement des lignes directrices pour, d'une part, mener et exploiter l'étude de l'homogénéité et de la stabilité des Entités Soumises à l'Essai d'Aptitude (ESEA), et d'autre part, concevoir le plan d'expérience et effectuer le traitement statistique des données nécessaire à l'évaluation de la performance des participants.

En complément des approches expérimentales et statistiques, ce guide fournit également des éléments méthodologiques pour la mise en œuvre d'un EILA et pour la rédaction des rapports d'EILA.

Certains éléments sont donnés sous forme de recommandation et d'autres sont apparus au GT nécessaires pour respecter les dispositions décrites dans les normes en vigueur, notamment les normes NF ISO 17043 pour les aspects généraux et NF ISO 13528 pour les aspects statistiques d'organisation des EILA.

Ce guide ne s'applique pas :

- ▶▶ aux essais inter-laboratoires permettant d'évaluer la performance des méthodes d'analyse (validation analytique), des kits, des réactifs, et des équipements issus des fournisseurs,
- ▶▶ aux essais inter-laboratoires permettant de caractériser des matériaux de référence.

Les termes généraux en matières de statistique peuvent être retrouvés dans la norme NF ISO 3435-1 (moyenne, variance, ...). Les termes plus spécifiques à ce guide sont définis dans le « *Glossaire relatif à la validation des méthodes d'analyses* » (code Ennov ANSES/FGE/0155 document uniquement disponible en interne).

2 Logiciels et outils de calcul

Le logiciel PROLAB (Quodata) a été sélectionné et mis à disposition des laboratoires de l'Anses pour réaliser l'ensemble des analyses statistiques concernant les EILA (homogénéité, stabilité, performance des participants). Il permet de mettre en œuvre les approches statistiques décrites dans la norme NF ISO 13528:2022 et dispose d'un module pour évaluer les performances d'un EILA à variable qualitative (L-score, S-score).

L'outil de calcul QualiScore (complément Excel ANSES) est également mis à disposition pour évaluer la performance des laboratoires sur la base d'EILA à variable qualitative. Il permet de calculer le S-score, nouveau score mis en œuvre dans le cadre du GT EILA et développé par l'unité Pathologie et Bien être des Ruminants (PBER) du laboratoire Ploufragan-Plouzané-Niort et la plateforme PAS.

Une notice d'utilisation de Prolab ainsi qu'une notice d'utilisation QualiScore (PPN/INS/0812) sont disponibles sous Ennov (*uniquement en interne*).

Le logiciel Prolab a été validé par Quodata et le jeu de données permettant de vérifier la validité des calculs est disponible auprès de la PAS. Cette vérification n'est pas nécessaire à chaque ouverture de session. Elle est réalisée par la plateforme PAS lors de la mise en production et lors de mises à jours de Prolab, ou au minimum une fois par an.

L'outil de calcul QualiScore a fait l'objet d'une validation par l'unité PBER. Toute mise à jour fera l'objet d'une validation et de mise à disposition avec un nouveau numéro de version.

Il est possible d'avoir recours à d'autres applications pour effectuer les statistiques relatives à l'homogénéité et la stabilité des ESEA, ainsi que la performance des laboratoires. Cependant, le recours à ces autres possibilités (autres logiciels, feuilles Excel, ...) nécessite que ces outils soient validés à la fois sur le plan statistique mais aussi sur le plan informatique, comme le précise les normes NF EN ISO/CEI 17043:2010 (article 4.7.1.1) et NF EN ISO/CEI 17025:2017 (paragraphe 7.11). Il est par ailleurs recommandé d'avoir un dossier de validation associé à ces outils, sauf dans le cas d'une conformité revendiquée vis-à-vis de la norme NF EN ISO/CEI 17043:2010 où ce dossier est obligatoire. Enfin, il est recommandé de consulter la plateforme PAS si des outils autres que PROLAB et QualiScore sont utilisés dans le cadre des EILA. La plateforme PAS peut notamment apporter un soutien pour la validation informatique et statistique de ces outils.

3 Prérequis à l'organisation d'EILA

3.1 Considérations générales

Avant toute mise en œuvre d'un EILA, diverses informations ou prérequis doivent être clairement connus, et en particulier :

- ▶▶ Quels objectifs sont attendus en fonction des domaines d'application des méthodes d'analyses et des seuils de décision (réglementaires, normatifs ...) ?
- ▶▶ Quelles sont les pratiques réalisées en routine et les résultats fournis ? En effet un EILA doit être le plus proche possible des conditions de routine. Toute autre démarche devrait être documentée dans le plan de campagne (évaluer le risque/impact de s'écarter des conditions de routine sur la fiabilité du résultat de l'EILA par rapport aux objectifs d'évaluation de compétence des participants).
- ▶▶ Quelles sont les matrices et les niveaux à envisager ?
- ▶▶ Quelles sont les performances des méthodes d'analyses (justesse, fidélité ...) disponibles dans un ou des dossiers de validation ou dans des documents de spécifications fournisseurs ? Ces informations devraient être connues aussi bien pour la ou les méthodes appliquées par l'organisateur que pour les méthodes utilisées par les participants. **En particulier, un EILA ne peut être organisé que si les performances des méthodes d'analyse utilisées dans la vérification de l'homogénéité et la stabilité des Entité Soumise à l'Essai d'Aptitude (ESEA) sont connues et documentées, c'est-à-dire, si ces méthodes d'analyse sont validées (cf. 3.1.1 de ce document).**
- ▶▶ Les outils informatiques développés en interne (feuille de calcul, ...) sont-ils validés ?

Ces informations permettent de bien concevoir le plan de campagne, dont les plans expérimentaux et les statistiques à appliquer, doivent être cohérents avec les objectifs annoncés.

Une règle majeure doit être respectée :
À un objectif d'EIL, un plan de campagne et un rapport spécifique !

Note : un EILA, orienté vers l'évaluation de l'aptitude des laboratoires, est différent d'un essai interlaboratoires de validation de méthode, orienté vers l'évaluation des performances de la méthode, et différent d'un essai interlaboratoires de transfert de méthode.

L'Agence dispose d'outils transversaux pour la mise en œuvre des EILA :

- ▶▶ LEILA : logiciel de gestion des EILA : inscription, instruction, dépôt du rapport,
- ▶▶ Support d'enquête de satisfaction auprès des participants aux EILA (Sphinx) – PR3/ANSES/6 (code Ennov ANSES/FGE/0146, document uniquement disponible en interne), et ses 4 formulaires de mise en œuvre associés, également disponibles sous Ennov Doc.
- ▶▶ Formulaire « EILA : support pour l'organisation et la rédaction du plan de campagne à l'usage des coordonnateurs » - PR3/ANSES/9 (Code Ennov ANSES/FGE/0281, document uniquement disponible en interne),
- ▶▶ Formulaire « Modèle de rapport d'EILA » - PR3/ANSES/9 (Code Ennov ANSES/FGE/0071, document uniquement disponible en interne),
- ▶▶ Formulaire « Modèle de rapport d'EILA » (Version anglaise) - ANSES/PR3/09/02-02 (Code Ennov ANSES/FGE/0072, document uniquement disponible en interne),
- ▶▶ Outils de calculs : Logiciel (Prolab) et QualiScore.

Une notice d'utilisation (PPN/INS/0812, document uniquement disponible en interne) décrit les fonctionnalités de l'outil QualiScore.

La plateforme PAS apporte un soutien méthodologique et transversal aux laboratoires de l'Agence : voir la Note d'Organisation disponible sous Ennov à la référence ANSES/NO/0035 (document uniquement disponible en interne), pour le détail des rôles et des missions de la plateforme PAS.

La suite de ce guide présente les informations nécessaires aux traitements statistiques et mentionne également des recommandations ou points de vigilance qui peuvent aider à concevoir le plan de campagne et mettre en œuvre l'EILA.

3.2 Informations sur la performance des méthodes d'analyse

Ce paragraphe concerne la performance des méthodes d'analyse utilisées et l'incertitude de mesure.

Ces informations devraient être connues, aussi bien pour la (ou les) méthode(s) appliquée(s) par l'organisateur que pour la (ou les) méthode(s) utilisée(s) par les participants.

3.2.1 Méthode(s) utilisée(s) par l'organisateur

Rappel de l'article 4.3.6 de la norme NF EN ISO/CEI 17043 :2010 :

Les organisateurs d'essais d'aptitude doivent s'assurer que les caractéristiques de performances des méthodes de laboratoire et du matériel utilisé pour confirmer la teneur, l'homogénéité et la stabilité des entités soumises à l'essai d'aptitude sont validées et maintenues de façon appropriée.

Dans le cas de la (ou des) méthode(s) appliquée(s) chez l'organisateur, ces informations sont utiles pour l'évaluation de l'homogénéité et de la stabilité des entités soumises à essai d'aptitude (ESEA) ainsi que pour la détermination de la valeur assignée, le cas échéant. Le Guide de validation des méthodes d'analyses (ANSES/PG/0069 [lien internet](#)) peut être utilisé pour caractériser les performances des méthodes ainsi que l'incertitude de mesure. Cette caractérisation doit être réalisée en amont de l'EILA.

Si une méthode utilisée est une méthode « fournisseur » (kit, méthode commercialisée en microbiologie des aliments...), la méthode est déclarée valide si elle fait l'objet d'une garantie statistique des performances (voir guide de validation) ou d'une certification par un organisme tiers (tel qu'AFNOR Certification ou MicroVal).

3.2.2 Méthode(s) utilisée(s) par les laboratoires participants

Rappel de l'article 4.5.2 de la norme NF EN ISO/CEI 17043:2010 :

Quand des participants sont autorisés à recourir à la méthode de leur choix, l'organisateur de l'essai d'aptitude doit :

- a) Avoir une politique et suivre une procédure relative à la comparaison des résultats obtenus par différentes méthodes d'essai ou de mesure**
 - b) Savoir, parmi les différentes méthodes d'essai ou de mesure de tout mesurande, celles qui sont techniquement équivalentes et prendre des mesures en conséquence pour évaluer les résultats des participants utilisant ces méthodes.**
-

Rappel de l'article 6.2.1 de la norme NF ISO 13528:2022 :

Lorsqu'on demande à tous les participants de consigner une valeur pour le même mesurande, il convient normalement que la valeur assignée pour les essais d'aptitude soit la même pour tous les participants. Cependant, lorsque des participants sont autorisés à choisir leur propre méthode de mesure, il est possible qu'une seule valeur assignée pour chaque analyte ou propriété ne soit pas appropriée pour tous les participants. Cela peut se produire, par exemple, lorsque des méthodes de mesure différentes fournissent des résultats qui ne sont pas comparables. Dans ce cas, l'organisateur d'essais d'aptitude peut utiliser une valeur assignée différente pour chaque méthode de mesure.

Les situations évoquées dans les deux rappels ci-dessus peuvent notamment se retrouver à l'Anses dans les domaines de la santé animale et de la santé végétale, pour lesquels de nombreux kits d'analyse commerciaux sont disponibles et ne font pas l'objet actuellement d'une certification (validation) par tierce partie. Cela peut aussi être le cas pour certains LR-UE. Certains organisateurs d'EILA peuvent donc se trouver confrontés à des résultats obtenus sur la base de méthodes (kits d'analyse) dont ils ne connaissent pas les performances, en dehors de celles revendiquées par les fabricants de ces kits.

Dans ce cas, il est recommandé, selon les possibilités ou les contextes :

- ▶▶ de limiter l'accès à l'EILA aux laboratoires utilisant les kits dont les performances ont pu être dûment contrôlées par le laboratoire de référence ;
- ▶▶ si cette limitation n'est pas possible, informer au préalable les laboratoires mettant en œuvre des kits non contrôlés que leurs résultats ne seront pas pris en compte pour établir la valeur assignée ;
- ▶▶ dans ce dernier cas, signaler à ces laboratoires que leur évaluation se fera avec la valeur assignée définie pour l'EILA (consensuelle sur les kits contrôlés, référence, ...) et qu'il pourrait en résulter que cette évaluation les pénalise le cas échéant ;
- ▶▶ de fixer, en relation avec le prescripteur de l'EILA, les valeurs seuils à partir desquels un participant pourra être déclaré apte, indépendamment de la méthode utilisée.

4 Rappel du contexte pour le traitement des données

Dans le cadre des EILA, des traitements informatiques, statistiques ou/et graphiques sont nécessaires à différentes étapes du processus :

1. Codification et randomisation des participants et des ESEA.
2. Vérification de l'homogénéité des ESEA envoyés aux participants.
3. Vérification de la stabilité des ESEA envoyés aux participants.
4. Examen initial des données fournies par chaque participant.
5. Vérification de la cohérence de l'EILA à partir des données issues de l'ensemble des participants.
6. Estimation de la valeur assignée.
7. Estimation de l'écart-type pour l'évaluation de l'aptitude (dans le cadre d'EILA quantitatif).
8. Évaluation des performances globales et des performances individuelles de chaque participant (sous forme de score).
9. Vérification éventuelle de la performance de l'EILA par rapport aux EILA précédents.

Au niveau normatif, il existe à ce jour essentiellement la norme d'application générale NF ISO 13528:2022 traitant de l'exploitation statistique des variables quantitatives continues ainsi que, dans le domaine de la microbiologie des aliments, la norme NF EN ISO 22117:2019 pour l'exploitation de variables quantitatives et qualitatives.

Au cours des réflexions menées au sein du groupe de travail et sur la base des retours d'expérience des EILA organisés au sein de l'Anses, la décision a été prise :

- » d'harmoniser les pratiques pour l'ensemble des laboratoires de l'Agence,
- » de préconiser les approches statistiques en accord avec les normes en vigueur, et plus particulièrement avec la norme NF ISO 13528:2022 d'application générale.

La suite de ce guide présente les choix retenus par le groupe de travail pour traiter les résultats à chacune des étapes et en fonction des types de variables, quantitatives ou qualitatives.

5 Contexte de l'EILA et plan expérimental

5.1 Description de la variable sur laquelle l'aptitude est évaluée

Ce chapitre concerne la nature du résultat rendu par les participants, obtenue avec la ou les méthode(s) d'analyse(s) utilisée(s) dans l'EILA et sur laquelle l'aptitude est évaluée (libellé - unité - expression).

La variable (c'est-à-dire le résultat ou le mesurande) doit être unique et clairement décrite.

Une attention particulière est à prévoir pour bien décrire le libellé, l'unité et le nombre de chiffre significatif de la variable.

Avant chaque EILA, il est impératif de bien définir la variable qui est exploitée, analysée et sur laquelle la décision d'aptitude est prononcée.

Il est nécessaire de bien distinguer les différences entre :

- » **Réponse ou signal**, c'est-à-dire la valeur provenant directement d'un équipement [la variable étant un signal ou une réponse instrumentale (par exemple, surface, hauteur de pic, PCR (Ct), densité optique)] ;
- » **Résultat**, c'est-à-dire la valeur que le participant donne dans ses rapports d'analyse et sur laquelle il est évalué (par exemple, teneur ($\mu\text{g}/\text{kg}$), analyte détecté/non détecté).

Les résultats peuvent être de deux origines :

- » **Valeur directe** : mesurée directement par un équipement ou lue par un opérateur (par exemple, comptage de parasites, présence/absence de l'analyte),
- » **Valeur dérivée** : calculée à partir d'une réponse ou d'un signal mesuré(e) par un équipement (par exemple via une courbe d'étalonnage).

Quelle que soit l'origine du résultat (valeur directe ou dérivée), les résultats doivent être obtenus avec des méthodes d'analyse validées, soit des méthodes maîtrisées avec une erreur systématique (justesse) et erreur aléatoire (fidélité) connue et acceptable.

Exemple 1 : la réponse instrumentale est la valeur de la densité optique (DO) (signal) d'une fluorescence mesurée.

Le résultat est : « détection d'entérotoxines de staphylocoques » à partir de la DO

Exemple 2 : la réponse instrumentale est une surface ou hauteur de pic mesurée.

Le résultat est : « concentration » calculée à partir de la réponse.

5.2 Classification de la variable

Différentes classes de variable existent :

- ▶ **Variable quantitative continue** : variable (résultats de mesure) dont les valeurs sont continues, par exemple : teneur, concentration, titre, ...
- ▶ **Variable quantitative discontinue (ou discrète)** : variable (résultats de mesure) dont les valeurs ne peuvent être que des nombres entiers, par exemple, dénombrements ou comptages
- ▶ **Variable qualitative** : variable (résultats de mesure) dont les valeurs sont classées en catégories ou modalités, par exemple :
 - « binaire » : cas où deux valeurs seulement sont possibles pour le caractère étudié et sont codées par exemple « 1 et 0 » ou « présence et absence », ou « détecté et non détecté »,
 - « qualitative ordinale » : cas où plus de deux valeurs sont possibles pour le caractère étudié (exemple : « présent, douteux, absent »).

Note : Comme précisé dans l'Annexe B/article B.3.2.2 de la norme NF EN ISO/CEI 17043:2010, une variable usuellement dénommée « semi-quantitative » est une variable qualitative ordinale.

Afin d'assurer une homogénéité pour l'ensemble des laboratoires de l'Anses, il est nécessaire de bien classer la variable en respectant les définitions ci-dessus.

5.3 Distribution des résultats (Normale, Log-normale, Binomiale, ...)

Pour les variables quantitatives, la connaissance de la loi de distribution (loi Normale généralement acceptée) peut être importante car la validité de certains calculs réalisés par la suite en dépend. À titre d'exemple, le calcul d'une moyenne ou d'un écart type n'a de sens que si la distribution des données suit une loi Normale.

Dans la mesure où le guide préconise par la suite l'utilisation des méthodes robustes (voir chapitre 8 paragraphe 8.1), la vérification de la « normalité » n'est pas requise (paragraphe 5.3.2. de la norme NF ISO 13528 :2022).

Si l'organisateur ne souhaite pas utiliser des méthodes robustes, alors la symétrie de la distribution doit être vérifiée, (en utilisant des outils tels que la droite de Henry (QQ plot), les graphiques de type « BOXPLOT » ou des tests d'hypothèses (Shapiro-Wilk, Anderson-Darling, Jarque-Bera).

La vérification de l'existence de plusieurs modes, par exemple liée à l'utilisation de méthodes différentes par les laboratoires, devrait être réalisée. L'approche définie dans la norme NF ISO 13528 :2022 (article 10.3, vérification de l'existence de plusieurs modes via le diagramme de densité par la méthode du noyau) est recommandée.

Ce point ne concerne pas les variables qualitatives.

5.4 Transformation des données avant la réalisation des traitements statistiques

Généralement, la transformation logarithme (par exemple \log_{10}) est utilisée pour rendre normale la distribution des données dans certains domaines. Elle est essentiellement pratiquée en microbiologie, notamment pour les dénombrements bactériens.

La transformation ne modifie pas la façon de traiter les résultats des EILA.

Après transformation, tous les calculs se font sur les données transformées, y compris pour déterminer les valeurs assignées, qu'elles soient fixées avant l'EILA ou calculées avec les résultats des participants.

Toute transformation doit être justifiée.

5.5 Typologie et nombre des participants

Dans le cas des **EILA quantitatifs**, si l'évaluation de la performance globale et individuelle des laboratoires s'effectue à partir de la valeur assignée et de l'écart-type d'aptitude déterminés à partir des valeurs des participants, un minimum de 13 laboratoires participants (p) est requis (Annexe D/article D.1.3.1 de la norme NF ISO 13528:2022).

Pour un nombre de laboratoires $p < 13$, une démarche statistique doit être justifiée au cas par cas avec un référent statistique ou la plateforme PAS, y compris pour des approches n'utilisant pas de valeurs issues des participants.

Si des sous-groupes de participants sont nécessaires (par exemple : effet méthode, regroupement des participants d'un pays...), un plan de campagne spécifique par sous-groupe doit être établi.

Les analyses en sous-groupes doivent respecter les mêmes critères statistiques qu'en l'absence de sous-groupe, en terme d'effectifs et de méthodes utilisées. Le nombre de participants par sous-groupe doit être suffisant (≥ 13) pour permettre les analyses statistiques robustes. Pour un nombre inférieur à 13, consulter le référent statistique ou la plateforme PAS.

Pour les **EILA qualitatifs**, le nombre de participants est précisé dans le paragraphe 8.4.1.1.3 de ce document ; il dépend du score utilisé.

Note : Dans le cas d'un seul participant d'autres méthodes de vérification de l'aptitude sont possibles, comme par exemple l'analyse en doublon ou la comparaison bilatérale (cf. Guide d'orientation relatif aux essais bilatéraux d'aptitude ANSES/PG/0143 *disponible sur le web*).

5.6 Nature des échantillons

Lors de la mise en place du plan expérimental, et quel que soit le type de variable (quantitative ou qualitative), deux types d'échantillons sont prévus :

- les échantillons nommés « **Entités Soumises à Essai d'Aptitude (ESEA)** », utilisés pour l'évaluation de l'aptitude,
- les échantillons nommés « **leures** » utilisés pour éviter les collusions.

Les **ESEA** doivent être choisies (matrice, niveau ...) conformément aux pratiques en routine. Ces ESEA peuvent aussi être de type « **blanc** » ou « **négatif** ».

Tous les participants d'un EILA doivent recevoir les mêmes ESEA (leures exceptés).

Il est recommandé de procéder à un traitement statistique des résultats d'EILA par type de matrice et par niveau.

Les « **leures** » ne sont pas utilisés pour les traitements statistiques appliqués pour l'évaluation de l'aptitude et servent à vérifier l'absence de collusion.

Note : Si possible et afin de permettre une meilleure homogénéité, il est préférable de préparer un « pool » ou une solution-mère globale. Les ESEA sont préparés à partir de ce « pool », plutôt que d'être préparés individuellement par supplémentation.

5.7 Nombre et type de résultats fournis par les participants

Dans le cas où l'on veut étudier, par matrice et par niveau, la dispersion des résultats par participant (donc évaluer l'aptitude des participants sur leur capacité à fournir des résultats peu dispersés ou fidèles), le plan de campagne doit décrire les conditions dans lesquelles les ESEA sont analysées. Pour étudier la justesse (par exemple z score, z' score pour des EILA quantitatifs) un résultat par laboratoire est suffisant. Pour évaluer la fidélité, un nombre minimal de 2 répétitions par matrice et par niveau est nécessaire. La fidélité est évaluée sous conditions de répétabilité.

6 Homogénéité et stabilité des ESEA

Il convient de choisir la méthode de vérification de l'homogénéité et stabilité des ESEA qui fournit la performance la plus adaptée pour détecter une éventuelle hétérogénéité et un éventuel défaut de stabilité de l'ESEA : par exemple, écart-type de répétabilité le plus faible possible pour les méthodes quantitatives ou taux de sensibilité/spécificité les plus élevés possibles pour les méthodes qualitatives.

6.1 Homogénéité des ESEA

6.1.1 Aspects généraux

Rappel de l'article 4.4.3.3 de la norme NF EN ISO/CEI 17043:2010 :

L'évaluation de l'homogénéité doit normalement être effectuée après avoir conditionné sous leur forme finale les entités soumises à l'essai d'aptitude et avant de les distribuer aux participants sauf si, par exemple, des études de stabilité indiquent qu'il convient de les stocker sous forme brute.

NOTE 1 : L'homogénéité peut être démontrée avant le conditionnement lorsqu'on peut raisonnablement s'attendre à ce que le conditionnement n'ait aucune influence.

NOTE 2 : Dans certains cas, il peut ne pas être possible d'exécuter l'essai d'homogénéité avant la distribution pour des raisons pratiques, techniques ou logistiques.

Même si la norme NF ISO 13528:2022 permet d'envoyer aux laboratoires participants des ESEA non homogènes, dans la mesure où l'écart-type d'aptitude tiendra compte de l'hétérogénéité, il est recommandé que l'étude d'homogénéité soit réalisée **avant l'envoi des ESEA** et après que ceux-ci aient été emballés dans leur forme finale (cf. Annexe B, article B.1.1.de la norme NF ISO 13528:2022).

En effet, les résultats obtenus dans cette étude d'homogénéité peuvent être utilisés pour l'étude de stabilité. De plus, compte-tenu des risques de recours et de la criticité du contexte des EILA organisés par les LNR/LRUE de l'Anses, il est préférable de mener cette étude avant envoi des ESEA, pour renforcer la fiabilité et la pertinence de l'EILA.

Note 1 : *le GT déconseille d'envoyer les ESEA sans vérification préalable de l'homogénéité. Un défaut d'homogénéité détectée a posteriori peut remettre en cause l'EILA.*

L'homogénéité doit être vérifiée :

- ▶▶ Pour chaque nouveau lot de production des ESEA, même si une étude d'homogénéité a été réalisée sur une campagne précédente ayant le même périmètre et même objectif.
- ▶▶ Pour chaque analyte ou famille d'analytes/chaque type de matrice/chaque niveau.

L'étude d'homogénéité doit être décrite dans le plan de campagne. La conduite à tenir en cas de non-homogénéité doit être définie au préalable de l'EILA dans le plan de campagne et cette information doit être fournie dans les instructions aux laboratoires participants.

Il est recommandé de ne pas faire de recherche de valeur(s) atypique(s)/aberrantes, avec élimination de ces valeurs (cf. 7.4). Si l'homogénéité n'est pas démontrée selon l'approche décrite à l'article B.2.2 (Annexe B) de la norme NF ISO 13528:2022, il est possible de démontrer que les ESEA sont suffisamment homogènes selon l'approche proposée par Thompson *et al*, (2006, International Union of Pure and Applied Chemistry (IUPAC)), décrite également dans l'article B.2.3.de la norme NF ISO 13528:2022. Il est également possible d'inclure la variabilité inter-ESEA dans l'écart-type d'aptitude, selon la formule B.3 indiquée dans l'article B.2.5.de la norme NF ISO 13528:2022.

Note 2 : *L'étude d'homogénéité n'a pas pour but de vérifier la variabilité (donc l'hétérogénéité) des kits d'analyse mais celle des ESEA.*

Note 3 : *L'étude d'homogénéité doit être effectuée avant l'étude de stabilité. Pour des matrices ayant une plage de stabilité réduite à quelques jours (par exemple 3-4 jours), les ESEA sont envoyées avant d'obtenir les résultats de l'étude d'homogénéité. Dans ce cas, il est recommandé de fixer le jour d'analyse par les participants et de réaliser l'étude d'homogénéité ce jour-là, les ESEA doivent être stockées dans des conditions équivalentes à ceux envoyés aux participants (cf. stabilité). Toujours dans ce cas, l'homogénéité et la stabilité sont confondues et les calculs statistiques sont ceux de l'homogénéité.*

Note 4 : *Indépendamment des aspects statistiques, une homogénéité acceptable peut également être justifiée par une bonne traçabilité de la préparation des ESEA, par exemple dans le cas d'une contamination ou supplémentation individuelle des ESEA, s'il n'est pas possible de vérifier l'homogénéité avant envoi. Il est possible également de vérifier l'homogénéité selon des*

connaissances acquises pour une combinaison équivalente analyte/matrice/niveaux, obtenues lors d'un EILA antérieur.

Note 5 : Dans le cadre des EILA réalisés en microbiologie alimentaire, l'article 6 de la norme NF EN ISO 22117:2019 peut être également suivi. Pour les EILA en santé animale, l'homogénéité peut être vérifiée en suivant l'Annexe B de la norme NF U47-400.

6.1.2 Cas des EILA quantitatifs

6.1.2.1 Plan d'expérience

Le plan d'expérience recommandé en première intention est le plan décrit dans la norme NF ISO 13528:2022 (Annexe B, article B.1.1).

Le nombre d'ESEA doit être au minimum de 10. Elles sont analysées chacune deux fois indépendamment, c'est-à-dire que chaque prise d'essai d'ESEA est analysée avec l'intégralité de la méthode d'analyse. La Figure 1 illustre ce principe d'indépendance.

Ces analyses sont conduites de préférence dans des conditions de répétabilité et dans un ordre aléatoire. Il est également recommandé de tracer l'ordre de production des ESEA, qui doit être lui aussi aléatoire. Cet ordre de production est utilisé pour l'analyse statistique des données.

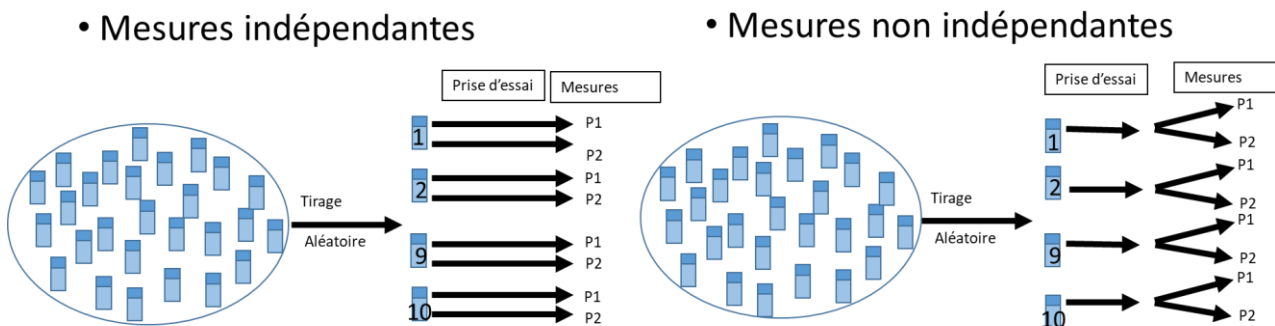


Figure 1: Illustration de la notion d'indépendance

Note : Il n'est pas recommandé de réduire le nombre d'ESEA à un nombre inférieur à 10, comme indiqué dans l'Annexe B, article B.1.1 de la norme NF ISO 13528:2022. S'il n'est pas possible de respecter ce nombre minimum de 10, le plan d'expérience doit être justifié (raison éthique par exemple).

Si l'analyse en double des ESEA n'est pas possible (contraintes techniques ne permettant pas de prélever deux prises d'essai), un minimum de 20 ESEA analysées une fois doit être prévu. S'il est envisagé d'utiliser moins de 20 ESEA, la démarche statistique doit être justifiée et approuvée par le référent statistique du laboratoire ou par la plateforme PAS.

Quel que soit le plan d'expérience retenu, les ESEA pour l'étude d'homogénéité sont prélevées au hasard parmi l'ensemble des ESEA préparées.

6.1.2.2 Évaluation de l'homogénéité

Il est recommandé de suivre les critères d'évaluation décrits dans l'Annexe B de la norme NF ISO 13528:2022 pour l'analyse statistique des données. Au préalable, deux vérifications sont nécessaires pour s'assurer que l'analyse statistique est possible. Ces deux vérifications, présentées ci-dessous, sont disponibles dans le logiciel Prolab (Annexe A).

■ Vérification 1

Avant de déterminer l'homogénéité des ESEA, effectuer la vérification suivante, décrite dans le rapport technique IUPAC (Thompson *et al*, 2006) et reprise dans l'article B.1.1 de la norme NF ISO 13528:2015:

L'écart type de répétabilité (s_r) de la méthode analytique doit être inférieur à **0,5 x l'écart type d'aptitude prédéfini** (σ_{PT})

$$s_r < 0,5 \sigma_{PT} \text{ OU } s_r/\sigma_{PT} < 0,5 \quad (1)$$

Si cette condition n'est pas vérifiée, les conclusions de l'EILA peuvent être remises en cause. Il est fortement recommandé de tenir compte de cette information selon l'approche « suffisamment homogène » décrite dans l'article B.2.3 de la norme NF ISO 13528:2022 et l'annexe B de ce document.

■ Vérification 2

- a) Conduire une étude de tendance sur l'**ordre de mesurage**, au moyen par exemple d'une régression linéaire. Si une tendance est décelée, identifier l'origine de la tendance (méthode, ESEA). Dans le cas où il s'agit des ESEA, interpréter les résultats avec précaution.
- b) De même conduire une étude de tendance sur l'**ordre de production** des ESEA (régression linéaire). Si une tendance est observée, se reporter au point b) de l'article B.2.1/Annexe B de la norme NF ISO 13528:2022.
- c) Effectuer la recherche d'une **différence entre les répétitions** au moyen du test de Cochran. Dans le cas d'une différence statistiquement significative, rechercher une explication et dans le cas d'une seule paire ou d'une donnée concernée, la paire ou la donnée concernée peut être supprimée.

Les ESEA sont déclarées suffisamment homogènes pour un niveau donné si la condition 1 est vérifiée et si :

« écart-type inter-ESEA » (s_s) calculé **≤ 0,3 x « écart type d'aptitude prédéfini »**

$$s_s \leq 0,3 \sigma_{PT} \text{ OU } s_s/\sigma_{PT} \leq 0,3 \quad (2)$$

Si les ESEA ne sont pas déclarées homogènes, il est recommandé d'étendre le critère de comparaison. Il est recommandé d'utiliser l'approche « suffisamment homogène » proposée dans l'article B.2.3 de la norme NF ISO 13528:2022, approche qui est décrite dans le rapport technique IUPAC et rappelée dans l'annexe B de ce document.

Note 1 : Quel que soit le type d'EILA, le facteur 0,3 signifie que la dispersion (ou variance) inter-ESEA ne contribue pas à plus de 10 % ($\approx 0,32$) de la dispersion (ou variance) d'aptitude. Ce facteur peut être augmenté sur justification. Par exemple, Thompson & Lowthian (1996) ont proposé un facteur de 0,4, soit une dispersion d'environ 16% ($\approx 0,42$).

Note 2 : Aucun appariement ne doit être effectué sur les 20 ESEA individuels pour revenir au plan expérimental décrit dans la norme NF ISO 13528:2022.

Dans le cas où il manque une valeur sur les 20, il est toléré de faire les calculs avec les 19 valeurs obtenues. S'il manque plus d'une valeur, il est recommandé de se référer à la plateforme PAS pour analyser les données.

Note 3 : Pour cette version du guide, le GT n'a pas défini un nombre optimum d'ESEA pour l'étude d'homogénéité, en fonction du nombre total d'ESEA prévus dans l'EILA : choix d'un nombre adapté et suffisamment représentatif. La plateforme PAS peut être saisie pour définir ce nombre optimal. Toutefois l'annexe C de ce document fournit des informations sur le nombre d'ESEA à tester quand le nombre d'ESEA produites est faible. Par ailleurs, le Guide ISO 35 propose dans l'article 7.4.1 des bonnes pratiques pour déterminer le nombre d'ESEA nécessaire à la vérification de l'homogénéité en fonction du nombre total d'ESEA produites.

Le critère d'évaluation de l'homogénéité dépend du choix de l'écart-type d'aptitude. Plusieurs possibilités existent, cf. 6.1.2.3.

6.1.2.3 Choix de l'écart-type d'aptitude σ_{PT}

Dans de nombreux EILA, l'écart-type d'aptitude provient des résultats des participants (valeur consensuelle) et il n'est pas connu au moment de conduire l'étude d'homogénéité.

Dans ce cas, plusieurs possibilités sont décrites dans l'article B.2.4/Annexe B de la norme NF ISO 13528:2022. L'écart-type d'aptitude peut :

- ▶ être obtenu par expérience acquise lors de campagnes précédentes,
- ▶ être calculé par une analyse de la variance (ANOVA), à un facteur (facteur : ESEA) en prenant une risque $\alpha=0,05$,
- ▶ correspondre à un écart-type déterminé par une méthode empirique applicable dans certains domaines analytiques (par exemple équation d'Horwitz modifié par Thompson pour les contaminants chimiques),
- ▶ correspondre à un écart-type déterminé par une expérience de fidélité de la méthode d'analyse utilisée pour l'étude d'homogénéité (*rappel : les méthodes doivent être validées avant de les appliquer*).

Dans le cas où la performance des laboratoires est évaluée en utilisant un écart-type d'aptitude déterminé à partir des valeurs des laboratoires participants (approche consensuelle) et que l'homogénéité a été évaluée avec un écart-type d'aptitude différent, il est conseillé de vérifier *a posteriori* l'homogénéité. Cette vérification peut s'effectuer :

- » par comparaison de l'écart-type d'aptitude utilisé pour évaluer l'homogénéité et l'écart-type d'aptitude consensuel pour l'évaluation de performance, via l'utilisation d'un test statistique de Fisher (comparaison de variances),
- » ou en vérifiant à nouveau l'homogénéité en utilisant l'écart-type d'aptitude consensuel.

Dans tous les cas, la vérification 1 (6.1.2.2) doit être conduite.

En Annexe A, un exemple d'analyse de l'homogénéité d'ESEA à l'aide du logiciel Prolab est fourni.

6.1.3 Cas des EILA qualitatifs

Le chapitre 11 de la norme NF ISO 13528:2022 introduit le plan d'expérience et l'analyse des EILA qualitatifs. En particulier l'article 11.2.3 introduit la notion de vérification de l'homogénéité pour ce type de variable.

Deux possibilités existent selon le type de variable étudiée : variable qualitative stricte (binaire et ordinale) ou possibilité d'utiliser une variable quantitative pour vérifier l'homogénéité.

6.1.3.1 Plan d'expérience

D'un point de vue expérimental, il est recommandé de suivre les préconisations de la norme NF ISO 13528:2022, avec un nombre minimum (n) de 10 ESEA, analysées chacune deux fois indépendamment, et de préférence dans des conditions de répétabilité.

Dans le cas où le nombre d'ESEA produites par niveau est faible, conséquence par exemple d'une faible disponibilité en matériaux de référence, d'un nombre de participants peu élevé ou de toute autre contrainte, il est recommandé de suivre le document ILAC « *Discussion Paper on Homogeneity and Stability Testing* » et du Guide ISO 35. L'annexe C de ce document décrit ces recommandations. Il est cependant recommandé de valider l'approche envisagée auprès du référent statistique de votre laboratoire ou de la plateforme PAS.

Si les analyses en double des ESEA ne sont pas possibles (contraintes techniques, essais destructifs), le nombre d'ESEA pour l'étude d'homogénéité peut être augmenté par exemple en prenant 20 ESEA, lorsque cela est possible et pertinent. Dans tous les cas, l'évaluation de l'homogénéité reste possible en adaptant les calculs, c'est-à-dire l'écart-type des résultats est utilisé comme écart-type inter ESEA (s_s). La stratégie doit être déterminée au cas par cas et doit être justifiée par une démarche statistique, avec si besoin, l'appui du référent statistique de votre laboratoire ou de la plateforme PAS.

Quel que soit le plan d'expérience retenu, il est nécessaire de prélever les ESEA au hasard parmi l'ensemble des ESEA préparées et les analyser de façon indépendante (cf. Figure 1).

6.1.3.2 Cas de vérification d'homogénéité uniquement avec une variable qualitative

L'homogénéité sera déclarée selon les spécifications décrites dans le plan de campagne. Par exemple pour l'ESEA dont la valeur assignée sera « Positive ou Négative », l'ESEA sera déclarée homogène si 100 % des résultats sont « Positifs ou Négatifs ».

Note 1 : Dans le cas d'un EILA où la performance des laboratoires est évaluée en utilisant le S-score, le critère d'homogénéité dépend du type d'ESEA utilisé. Pour les ESEA1, 100% de résultats correspondant à la valeur assignée « Positive ou Négative » sont attendus.

Note 2 : Pour les trois autres types d'ESEA (ESEA2, ESEA3 et ESEA4), c'est la loi binomiale $B(n, p)$ qui est recommandée pour vérifier l'homogénéité des ESEA avec :

- (n) : le nombre de répétitions correspondant à un nombre d'ESEA. Le protocole d'échantillonnage recommandé est celui utilisé en évaluation de l'homogénéité : soit 2 séries de $n=10$ ESEA en utilisant la loi binomiale $B(n=10, p)$, ou 1 série de 20 ESEA en utilisant la loi binomiale $B(n=20, p)$.
- (p) la probabilité de succès ($\geq 50\%$) correspondant à la probabilité d'obtenir un résultat binaire (« Positive ou Négative ») sur le nombre d'ESEA n . Cette probabilité peut être définie par un pourcentage (p) de résultats majoritaires (positifs ou négatifs) obtenu à partir de données historiques ou estimé à partir d'un échantillonnage d'ESEA représentatives des ESEA envoyées aux participants. Ce pourcentage de résultats majoritaires (p) permet de définir le nombre minimum et maximum de résultats (positifs ou négatifs) attendus pour vérifier l'homogénéité (loi binomiale au risque de 5 %).

Ce pourcentage (p) peut être modulé en fonction de critères métiers. Par exemple, dans le cas où (p) estimé est égal à 70% de résultats positifs, le critère d'acceptation statistique de l'homogénéité correspondant est de 5 à 9 résultats positifs sur chacune des 2 séries de 10 ESEA. Pour permettre de valider l'homogénéité des ESEA dans le cas où 100% de résultats positifs seraient observés, l'organisateur peut opter pour un pourcentage (p) proche, ou légèrement supérieur au (p) estimé qui devra être supérieur ou égal à 74,2% dans le cas de 2 séries de 10 ESEA (Annexe J, Tableau J.1 avec $g^2=10$).

Note 3 : Un S-score d'homogénéité pour les ESEA2, ESEA3 et ESEA4, nommé S_{HOM} , est disponible via l'outil de calcul Anses QualiScore (cf. notice d'utilisation). Les ESEA sont déclarées homogènes si $(1 \leq S_{HOM} < 2)$. Pour les ESEA2 non répliquées, cf. § 8.4.1.1.3 b), il est nécessaire de vérifier l'homogénéité pour chaque niveau.

Il est conseillé de définir la démarche statistique au cas par cas en fonction du score d'évaluation de la performance utilisée dans l'EILA, en liaison avec le référent statistique de votre laboratoire ou la plateforme PAS.

Dans le cas où les ESEA ne peuvent pas être déclarées homogènes, le GT n'a pas prévu de solution alternative. De nouveaux ESEA doivent être préparés.

Note 4 : Il est également possible, pour des EILA en microbiologie des aliments, d'utiliser l'approche décrite dans la norme NF EN ISO 22117:2019, qui fait référence à des publications spécifiques au domaine des dénombrements bactériens, telles que Fearn & Thompson (2001) et Heisterkamp et al. (1993).

6.1.3.3 Cas particulier : vérification de l'homogénéité avec une variable quantitative

Il est possible de procéder à une vérification de l'homogénéité en utilisant une variable quantitative intermédiaire, comme par exemple une mesure d'absorbance pour les tests ELISA, une mesure de nombre de cycles d'amplification pour une méthode de PCR temps réel ou de LAMP PCR, un comptage pour une analyse qualitative en bactériologie, etc. (article 11.2.3 norme NF ISO 13528 :2022).

À des niveaux proches du seuil de détection (LOD ou en-dessous du Niveau Exigible de Détectabilité (NED)) pour la variable qualitative finale, il n'est pas conseillé de vérifier l'homogénéité avec une variable quantitative intermédiaire, même si elle est fortement liée à la variable qualitative. En effet, dans la mesure où le résultat de l'analyse rendue au client est qualitatif (détecté/non détecté, conforme/non conforme, présent/absent ou formulation équivalente) et où les participants à l'EILA sont évalués sur la base de ces résultats qualitatifs, il est important de s'assurer de l'homogénéité selon la valeur assignée qualitative.

Lorsque l'utilisation d'une variable quantitative intermédiaire est possible et pertinente, cette variable est utilisée uniquement afin de sécuriser le processus décisionnel par le complément d'information qu'elle apporte. Dans ce cas, il est recommandé de procéder à l'évaluation de l'homogénéité avec cette variable quantitative intermédiaire selon le processus décrit dans l'article 6.1.2 de ce document.

Dans le cas d'une évaluation de l'homogénéité avec une variable qualitative finale et une variable quantitative intermédiaire, s'il y a une discordance de conclusion en utilisant les deux variables, la conclusion obtenue avec la variable qualitative est retenue, sauf exception justifiée par le coordonnateur de l'EILA.

Note : Dans le cas d'échantillons fortement positifs, s'il y a une incohérence entre l'homogénéité évaluée par une variable quantitative intermédiaire et l'interprétation par la variable qualitative (par exemple 1 échantillon sur 10 positif ayant une faible valeur intermédiaire, il est conseillé de rechercher les causes de cette incohérence et de faire une analyse de risques pour accepter les ESEA. Il est également possible de refaire une étude d'homogénéité en prenant de nouveaux ESEA.

6.2 Stabilité des ESEA

6.2.1 Aspects généraux

L'objectif de l'étude de stabilité est de garantir que le temps et les conditions de stockage et transport des ESEA (par exemple, congélation) n'ont pas d'impact sur la valeur assignée des ESEA.

D'une façon générale, deux échéances suffisent pour mener l'étude de stabilité : une première échéance au moment de l'étude d'homogénéité et une seconde choisie en adéquation avec la date-

butoir de retour des résultats par les participants. Dans ce cas, il faut appliquer l'annexe B de la norme NF ISO 13528 :2022 pour le plan expérimental et les calculs statistiques.

Note 1 : *La date-butoir du retour des résultats, qui correspond à la dernière échéance de l'étude de stabilité, doit être inscrite dans le plan de campagne et doit être communiquée aux participants, leur stipulant qu'au-delà de cette date, leurs résultats ne seront pas pris en compte. La traçabilité des analyses (date de réception, date de réalisation, etc..) est extrêmement importante dans ce cas pour l'acceptation des résultats.*

L'étude de stabilité est réalisée :

- ▶ soit sur le lot d'ESEA de production de l'EILA,
- ▶ soit selon des connaissances acquises pour une combinaison équivalente analyte/matrice/niveaux, obtenues lors d'un EILA antérieur.

Les ESEA doivent être analysées, pour la stabilité, dans des conditions équivalentes à celles des ESEA envoyées aux participants : par exemple, conditions d'emballage, de stockage, de transport, même période d'analyse que les participants après réception des ESEA, etc.

Note 2 : *Une étude préalable de stabilité peut être conduite pour définir la durée de l'EILA (délai maximal d'analyse autorisé), si besoin des échéances intermédiaires peuvent être définies et les conditions de stockage/transport optimales peuvent être précisées.*

Il n'est pas obligatoire de conduire l'étude de stabilité sur la totalité des combinaisons matrices/niveaux/analytes si des informations claires et justifiées provenant d'une étude de stabilité déjà faite au préalable, dans des conditions identiques, sont disponibles (cf. article B.4.1.1, NF ISO 13528:2015). Néanmoins, réaliser au minimum un suivi de la stabilité de la combinaison supposée la moins stable (matrice/niveau/analyte le plus sensible) (cf. article B4, NF ISO 13528:2015).

En revanche, si aucune information de stabilité n'existe au préalable, une étude de stabilité doit être conduite pour chaque combinaison analyte/matrice/niveau.

Il est recommandé de mentionner les conditions de transport des ESEA (par exemple la température et le délai) dans le plan de campagne.

Il est recommandé de décrire les conditions de stockage à réception des ESEA dans les instructions aux participants.

L'utilisation d'enregistreurs de température peut être prévue pour tracer la température pendant le transport.

6.2.2 Cas des EILA quantitatifs

Prendre au moins quatre ESEA (2 ESEA pour chacune des 2 échéances) pour l'étude de stabilité, analysées indépendamment en double et de façon aléatoire (cf. Figure 1). Ces ESEA sont

différentes des ESEA pour l'étude d'homogénéité et doivent être prises aléatoirement dans le panel d'ESEA. Si l'échéance d'homogénéité est utilisée comme échéance 0 de la stabilité, le nombre d'ESEA pour l'étude de stabilité peut être réduit à 2.

Si les analyses en double des ESEA ne sont pas possibles (contraintes techniques pour prélever deux prises d'essai par ESEA), il est possible de prendre un minimum de huit ESEA (4 par échéance) analysées en simple. Si l'échéance d'homogénéité est utilisée comme échéance 0, le nombre d'ESEA peut être réduit à 4. En cas d'impossibilité de respecter le nombre minimum d'ESEA, le plan expérimental doit être clairement justifié (raison éthique, disponibilité des ESEA par exemple). Au final, le nombre d'ESEA nécessaires à cette étude peut être déterminé au cas par cas à l'aide d'une démarche statistique et doit être justifiée avec l'aide du référent statistique du laboratoire ou de la plateforme PAS.

Prendre comme écart-type d'aptitude (formule de calcul B.17, article B.5.1., NF ISO 13528:2022) le même que celui retenu pour l'étude d'homogénéité. Comme pour l'homogénéité, il est recommandé de vérifier *a posteriori* la stabilité lorsque l'écart-type pour évaluer la performance est différent de l'écart-type pour déclarer les ESEA stables.

Les approches décrites dans l'article 6.1.2.3 de ce guide pour la vérification de l'écart-type d'aptitude sont applicables. La vérification utilisant le test de Fisher s'applique à la fois à l'homogénéité et à la stabilité.

Il peut être judicieux de prévoir des échéances intermédiaires en fonction de la durée de l'EILA. Dans ce cas, les calculs statistiques suivent également la norme NF ISO 13528:2022, l'échéance de référence étant toujours la première échéance correspondant aux résultats obtenus lors de l'étude de l'homogénéité (et non l'échéance intermédiaire (T-1)).

Déclarer les ESEA suffisamment stables si :

$$|\bar{y}_1 - \bar{y}_2| \leq 0,3\sigma_{pt} \quad (3)$$

Avec :

- \bar{y}_1 : moyennes des résultats obtenus à l'échéance 0.
- \bar{y}_2 : moyennes des résultats obtenus à l'échéance T.

Note 1 : Quel que soit le type d'EILA, le facteur 0,3 signifie que la dispersion (ou variance) inter-ESEA ne contribue pas à plus de 10 % ($\approx 0,3^2$) de la dispersion (ou variance) d'aptitude. Ce facteur peut être augmenté sur justification. Par exemple, Thompson & Lowthian (1996) ont proposé un facteur de 0,4, soit une dispersion d'environ 16% ($\approx 0,4^2$).

Dans le cas où l'ESEA ne peut pas être déclarée suffisamment stable, prendre en compte l'incertitude des résultats selon la formule B.18, article B.5.2 de la norme NF ISO 13528:2022 :

$$|\bar{y}_1 - \bar{y}_2| \leq 0,3\sigma_{pt} + 2\sqrt{u^2(\bar{y}_1) + u^2(\bar{y}_2)} \quad (4)$$

Où :

- $u(\bar{y}_1)$ est l'incertitude liée à la valeur \bar{y}_1
- $u(\bar{y}_2)$ est l'incertitude liée à la valeur \bar{y}_2

- le facteur 2 est un facteur d'élargissement pour l'incertitude élargie de la différence, fournissant une confiance d'environ 95 %, et le calcul de l'incertitude composée a intentionnellement supposé que \bar{y}_1 et \bar{y}_2 sont indépendants.

Il est également possible d'utiliser un test statistique d'hypothèse (cf. article B.5.4, norme NF ISO 13528:2022) de comparaison de deux moyennes (test t de Student en séries non appariées). À un niveau de confiance de 5 % et sous réserve d'avoir trois ESEA soumis à l'essai, la détection de l'instabilité est équivalente à celle proposée par l'équation 3.

D'autres approches sont possibles (cf. ISO Guide 35), comme par exemple modéliser la stabilité ou établir l'existence de tendance (décroissance ou croissance significative de la valeur assignée). Pour les mettre en œuvre, il est conseillé de consulter soit le référent statistique de votre laboratoire, soit la plateforme PAS.

En annexe D, un exemple d'analyse de la stabilité à l'aide du logiciel Prolab est fourni.

6.2.3 Cas des EILA qualitatifs

Prendre au moins quatre ESEA pour l'étude de stabilité (2 ESEA pour chacune des 2 échéances), analysées indépendamment en double et de façon aléatoire. Ces ESEA sont différentes des ESEA pour l'étude d'homogénéité et doivent être prises aléatoirement dans le panel d'échantillons. Si l'échéance d'homogénéité est utilisée comme échéance 0 de la stabilité, le nombre d'ESEA pour étudier la stabilité peut être réduit à 2.

Les analyses de tous les ESEA doivent être conformes aux résultats attendus (par exemple 100% de positifs ou 100 % de négatifs).

Note 1 : Dans le cas d'un EILA où la performance des laboratoires est évaluée avec le S-score, le critère de stabilité dépend du type d'ESEA utilisé. Pour les ESEA1, 100% de résultats correspondant à la valeur assignée « Positive ou Négative » sont attendus. Pour les trois autres types d'ESEA (ESEA2, ESEA3 et ESEA4), le nombre d'ESEA doit être augmenté dans les mêmes conditions d'effectifs que l'homogénéité : 10 ESEA analysées 2 fois ou 20 ESEA analysées 1 fois pour chacune des 2 échéances. La stabilité est évaluée en utilisant la loi binomiale avec un risque de 5%.

Note 2 : Un S-score de stabilité pour les ESEA2, ESEA3 et ESEA4, nommé S_{STAB} , est disponible via l'outil de calcul Anses QualiScore (cf. notice d'utilisation PPN/INS/0812). Les ESEA sont déclarées stables si ($1 \leq S_{STAB} < 2$). Pour les ESEA2 non répliquées, cf. § 8.4.1.1.3 b), il est nécessaire de vérifier la stabilité pour chaque niveau.

Il est conseillé de définir la démarche statistique au cas par cas en fonction du score d'évaluation de la performance utilisée dans l'EILA, en liaison avec le référent statistique de votre laboratoire ou la plateforme PAS.

Dans le cas où l'ESEA ne peut pas être déclarée suffisamment stable, cette ESEA est exclue de l'évaluation de la performance des laboratoires. Il est recommandé de préparer une nouvelle ESEA et de l'envoyer aux laboratoires en fixant une date butoir d'analyse. L'analyse doit être également réalisée par le laboratoire organisateur à cette date butoir. Si les valeurs trouvées par le laboratoire organisateur et les laboratoires participants sont similaires, l'ESEA peut être utilisée pour évaluer la performance des laboratoires.

Dans le cas où une variable quantitative intermédiaire est utilisée, suivre la démarche décrite dans le paragraphe 6.1.3.3.

Note 3 : *Indépendamment des aspects statistiques, la justification d'une stabilité acceptable doit également passer par une bonne traçabilité de la maîtrise des conditions de stockage et d'analyse. Pour toutes les échéances, les résultats doivent être obtenus avec la même méthode validée et le laboratoire qui a réalisé les analyses.*

Note 4 : *Pour des raisons éthiques ou de disponibilité, le nombre d'ESEA peut être diminué, un minimum de trois analyses en simple est cependant conseillé.*

7 Prétraitement des résultats

7.1 Examen initial des informations et des données fournies par chaque participant

Au préalable de l'analyse graphique et statistique des données, il est conseillé de vérifier que les informations et les résultats fournis par les participants sont en adéquation avec celles consignées dans le plan de campagne et dans les instructions fournies aux participants.

Par exemple :

- ▶ *Suivi de(s) méthode(s) d'analyse définie(s) dans le plan de campagne.*
- ▶ *Conformité des ESEA aux spécifications (transport, conservation...).*
- ▶ *Respect de l'unité de mesure.*
- ▶ *Respect du nombre de chiffres significatifs.*
- ▶ *Lisibilité des résultats.*
- ▶ *Respect du format du résultat.*

L'utilisation de « check-list » est fortement recommandée (cf. *PR3/ANSES//9 Code Ennov ANSES/FGE/0070, document uniquement disponible en interne*)).

Chacune des consignes définies dans les instructions aux participants est associée à des critères d'inclusion et d'exclusion.

Les critères de refus, ainsi que la conduite à tenir par l'organisateur d'EILA en cas de données non conformes, doivent être définis au préalable dans le plan de campagne. Dans ce cas et au vu des résultats bruts obtenus, l'organisateur a la possibilité d'exclure des participants. Ces dispositions doivent être connues à l'avance par les participants.

Ne pas utiliser de test statistique pour décider d'inclure ou d'exclure des résultats des participants.

La finalité première de cet examen initial est de ne pas prendre en compte des résultats non conformes aux instructions, afin de minimiser les risques de conclusion erronée sur l'aptitude des participants. À ce niveau, l'exclusion d'un laboratoire le rend « non évalué ».

Effectuer impérativement cet examen initial avant la vérification de la cohérence de l'EILA (paragraphe 7.2).

7.2 Vérification de la cohérence de l'EILA

7.2.1 Objectif

La vérification de la cohérence des résultats est un préalable indispensable à toute analyse statistique, même pour un EILA à variable qualitative.

Pour des EILA à variable quantitative, il convient de vérifier que les résultats suivent une distribution uni-modale et symétrique. Cette vérification permet de déceler des anomalies (par exemple, valeur atypique inhabituelle) ou des sources non prévues de variabilité (par exemple, distribution multimodale pouvant mettre en évidence des résultats issus de méthodes d'analyse différentes). Par ailleurs, si les données sont disponibles, il est recommandé de vérifier la cohérence des résultats par rapport aux campagnes précédentes.

7.2.2 Méthode

Cette vérification peut s'appuyer sur les graphiques suivants, qui fournissent une analyse descriptive:

- » variable **qualitative** : statistiques descriptives des résultats (par exemple, fréquence pourcentage d'erreur élevé)
- » variable **quantitative** : histogramme, « boxplot » (voir Figure 2) ou boîte à moustaches, diagramme de densité par la méthode du noyau (disponible sous Prolab)

Ces graphiques permettront de vérifier les points suivants :

- » adéquation des résultats bruts avec les objectifs décrits dans le plan de campagne,
- » cohérence dans la distribution des résultats,
- » présence d'éventuelles sous-populations de résultats,
- » présence de résultats atypiques,
- » présence de tendances ou des constats particuliers.

Les graphes de Mandel (« h » et « k ») ne sont pas recommandés pour ce type de vérification.

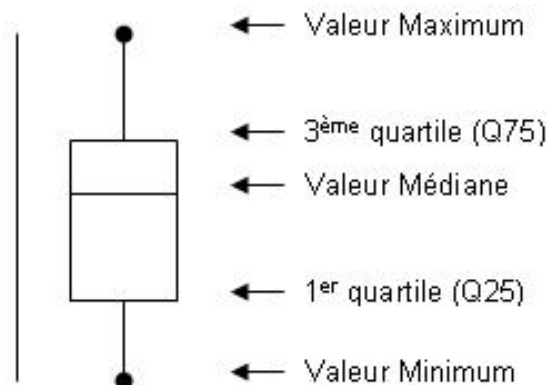


Figure 2 : Exemple de graphique « boxplot »/« diagramme à moustaches »

Médiane = 2^{ème} quartile (Q50), ce qui signifie que 50% des résultats ont une valeur supérieure à la médiane et 50% des résultats ont une valeur inférieure à la médiane

Certains graphiques (tels que boxplots) ne sont disponibles que dans des logiciels de statistique (par exemple R) ou via un complément pour Excel (« Analyse-it »).

7.2.3 Conclusion

Si des résultats inhabituels sont détectés, se référer au paragraphe 7.4. Si une incohérence est décelée (telle qu'une multi-modalité), une approche robuste par modalité peut être proposée si le nombre de participants par modalité est suffisant ($p \geq 13$).

7.3 Gestion des données manquantes

Dans le cas général, noter un résultat manquant quand le participant rend un résultat partiellement ou totalement « non évaluable », en tenant compte de la justification écrite apportée par le participant lors du rendu de ses résultats.

Si le coordonnateur souhaite évaluer tous les participants prévus, il doit alors décrire dans le plan de campagne et dans les instructions aux participants, les dispositions permettant un nouvel envoi des ESEA, en accord avec les résultats de l'étude de stabilité.

Respecter pour ces nouveaux envois les contraintes, telles que les conditions de répétabilité pour les analyses (si applicable, cf. 5.7) et la prévention de collusion (anonymat). Ces nouveaux envois peuvent être complets ou partiels ; le choix est lié au contexte et doit être fait en accord avec le référent statistique du laboratoire ou la plateforme PAS.

Note : *Il est de la responsabilité du coordonnateur de décider si un participant est déclaré partiellement apte, dans le cas où le participant n'a fourni qu'une partie des résultats évaluable. Définir impérativement ce point dans le plan de campagne et en informer chaque participant avant le lancement de l'EILA (instructions aux participants).*

7.4 Prise en compte des données atypiques

Plusieurs articles de la norme NF ISO 13528:2022 abordent la notion de « valeurs aberrantes ou atypiques » et en précisent les modalités de traitement.

Sur le principe, cette norme recommande de considérer que tout résultat transmis par le participant est de sa responsabilité et est exploitable par l'organisateur. Il n'y a donc pas de valeur dite atypique ou aberrante. De plus, le calcul, à partir des résultats des participants, de la valeur assignée et l'écart-type d'aptitude par des approches robustes permet de ne pas exclure des résultats des traitements statistiques.

Par conséquent, ne pas rechercher en utilisant des tests statistiques des valeurs atypiques ou aberrantes. Justifier toute valeur non retenue dans l'évaluation de l'aptitude par une raison autre que statistique, telle que :

- ✓ le cas où plusieurs méthodes d'analyse (techniques analytiques, milieux de culture, kits, ..) présentent des résultats divergents,
- ✓ le suivi de l'historique des performances d'un participant (étude des tendances par exemple),

- ✓ les données ne permettent pas d'appliquer le modèle statistique établi pour l'évaluation de la performance,
- ✓ les cas décrit en paragraphe 7.1.

Décrire dans les instructions fournies à chaque participant les cas de non-prise en compte de valeurs pour l'évaluation de l'aptitude.

Dans l'hypothèse, non recommandée, où il serait nécessaire d'évaluer la présence de données dite aberrantes, le responsable du traitement des données pourrait utiliser des tests statistiques classiques (Cochran, Grubbs, ...), en se référant si besoin au référent statistique du laboratoire ou à la plateforme PAS.

8 Évaluation de la performance des laboratoires participants

8.1 Argumentaire pour l'utilisation d'une approche robuste

D'un point de vue théorique, l'approche classique, qui utilise des estimateurs paramétriques (par exemple, moyenne arithmétique, variance...), est fondée sur l'hypothèse que les erreurs sont distribuées selon une loi Normale. Cette hypothèse est souvent acceptable, car les erreurs résultent de la somme de petites erreurs indépendantes. Au contraire, cette hypothèse n'est pas acceptable pour les résultats atypiques ou aberrants, qui résultent en général d'erreurs isolées importantes.

L'approche robuste présente plusieurs avantages, principalement :

- Elle s'accommode des résultats atypiques, au lieu de les rejeter.
- Il n'y a pas de consensus sur la procédure la plus appropriée pour rejeter les résultats atypiques. Par exemple, la norme NF ISO 5725-2:2020 fixe le niveau de risque pour rejeter les résultats aberrants par les tests de Grubbs ou de Cochran à 1%. Ce niveau de risque est différent selon d'autres sources (par exemple AOAC ou IUPAC).

Une représentation graphique type box plot est recommandée pour visualiser d'une part la répartition des données et d'autre part les données atypiques. Si le nombre de valeurs atypiques est élevé (par exemple plus de 20 % des valeurs), il est recommandé de prendre contact avec le référent statistique du laboratoire ou la plateforme PAS.

L'exclusion des valeurs erronées (c'est-à-dire des valeurs liées, par exemple, à des erreurs de transcription, de calcul, d'unités...) doit être justifiée. Elle repose sur l'identification de l'erreur et sur l'information fournie en amont aux laboratoires participants, par exemple via la fiche de présentation ou la fiche d'information aux participants (LEILA).

Utiliser seulement les statistiques robustes, et non les statistiques paramétriques, pour estimer la valeur assignée et l'écart-type d'aptitude à partir des données des participants (valeurs consensuelles). C'est également le choix de la norme NF ISO 13528 2022 (article 6.5.1). Ce guide reprend les estimateurs robustes définis dans cette norme, en particulier l'algorithme A (estimation d'une médiane pour la valeur assignée) ou de l'algorithme S pour des valeurs répétées (estimation d'un écart-type). Ces algorithmes sont disponibles dans le logiciel Prolab.

8.2 Détermination de la valeur assignée

8.2.1 EILA quantitatifs

Il est recommandé de suivre les approches décrites dans la norme NF ISO 13528 :2022, qui propose différentes possibilités :

Méthodologie	Rappel des articles de la norme NF ISO 13528:2022
Sans utiliser les résultats des participants à l'EILA : Par formulation	7.3
Sans utiliser les résultats des participants à l'EILA : Matériau de référence certifié	7.4
Sans utiliser les résultats des participants à l'EILA : Résultats provenant d'un seul laboratoire	7.5
Avec les résultats des participants à l'EILA : Valeurs consensuelles provenant de laboratoires experts	7.6
Avec les résultats des participants à l'EILA : Valeurs consensuelles des laboratoires participants	7.7

Mentionner clairement et justifier l'approche retenue dans le plan de campagne et dans le rapport d'EILA :

- ▶ Choisir le mode de détermination de la valeur assignée en accord avec le tableau précédent, c'est-à-dire avec la norme NF ISO 13528:2022.
- ▶ Si une des approches « valeurs consensuelles provenant de laboratoires experts » ou « valeurs consensuelles des laboratoires participants » est retenue, effectuer les calculs en utilisant des statistiques robustes, conformément à l'article 7.7 de la norme NF ISO 13528:2022.

8.2.2 EILA qualitatifs

Des informations sont fournies dans le chapitre 11 de la norme NF ISO 13528:2022, complétées par l'article B.2.4 de la norme NF EN ISO/CEI 17043 :2010 :

« ... En général, il est nécessaire que ces valeurs assignées soient déterminées par le jugement d'un expert ou par la fabrication. Dans certains cas, un organisateur d'essais d'aptitude peut utiliser une valeur faisant consensus, définie par un accord de pourcentage majoritaire prédéterminé de réponses (par exemple, 80 % ou plus). Toutefois, il convient que le pourcentage utilisé soit déterminé en se basant sur les objectifs du programme d'essais d'aptitude et sur le niveau de compétence et d'expérience des participants... ».

Si la variable peut se ramener à une variable quantitative (par exemple, transformation en logarithme d'un dénombrement/comptage, DO), la valeur assignée est définie selon les approches proposées dans le § 8.1.1.

8.3 Détermination de l'écart-type d'aptitude

8.3.1 EILA quantitatifs

Suivre les approches décrites dans la norme NF ISO 13528:2022, qui propose différentes possibilités :

Méthodologie	Article de la norme NF ISO 13528:2022
Détermination par perception d'experts	8.2
Détermination à partir de l'expérience acquise lors des campagnes précédentes d'un programme d'EILA	8.3
Détermination par l'utilisation d'un modèle général (exemple HORWITZ)	8.4
Utilisation des écarts-types de répétabilité et de reproductibilité obtenus lors d'une étude collaborative de fidélité de la méthode d'analyse	8.5
Détermination à partir de données obtenues lors d'une même campagne d'EILA	8.6

Mentionner clairement et justifier l'approche retenue dans le plan de campagne et dans le rapport EILA :

- » Choisir le mode de détermination de l'écart type d'aptitude en accord avec le tableau précédent, c'est-à-dire avec la norme NF ISO 13528:2022.
- » Si l'approche « **détermination à partir de données obtenues lors d'une même campagne d'EILA** » est retenue, effectuer les calculs selon les statistiques robustes.

8.3.2 EILA qualitatifs

Si la variable peut se ramener à une variable quantitative (par exemple, transformation en logarithme d'un dénombrement/comptage, DO), appliquer les approches proposées dans le § 8.3.1.

Sinon, il n'est pas possible d'estimer un écart-type d'aptitude. Cette impossibilité n'est pas gênante pour l'évaluation de l'aptitude des participants.

Toutefois, il peut être envisagé de déterminer l'aptitude des participants en appliquant les statistiques de performance proposées dans l'article 9.3 de la norme NF ISO 13528:2022 (différence : D ou différence en %).

8.4 Évaluation de la performance des participants

8.4.1 Estimation de la justesse

8.4.1.1 Cas d'un score unique

8.4.1.1.1 Aspects généraux

L'aptitude des participants se fait par le calcul d'un score lié à la justesse.

Il est recommandé de suivre les approches décrites dans la norme NF ISO 13528 :2022. D'autres approches, complémentaires à celles décrites dans cette norme, ne sont pas décrites dans ce guide. Clairement mentionner et justifier l'approche retenue dans le plan de campagne.

8.4.1.1.2 Cas des EILA quantitatifs

Il est recommandé d'utiliser le score z , **si l'incertitude de la valeur assignée est inférieure à 0,3 fois l'écart-type d'aptitude** (cf. article 9.2 de la norme NF ISO 13528:2022).

Calculer le score z_i pour chaque laboratoire par la formule suivante :

$$z_i = \frac{(x_i - x_{pt})}{\sigma_{pt}} \quad (5)$$

Où :

- ▶ x_i moyenne des résultats de chaque laboratoire i à un niveau donné
- ▶ x_{pt} valeur assignée (obtenue par la méthode robuste)
- ▶ σ_{pt} écart-type d'aptitude

Interpréter les valeurs du score z de la manière suivante (cf. norme NF EN ISO/CEI 17043:2010, article B.4.1.1 c) :

- ▶ si $|z| \leq 2,0$, les performances sont considérées comme « satisfaisantes », aucun signal n'est généré ;
- ▶ si $2,0 < |z| < 3,0$, les performances sont considérées comme « discutables », un signal d'avertissement est généré ;
- ▶ si $|z| \geq 3,0$, les performances sont considérées comme « insatisfaisantes », un signal d'action est généré.

Si l'incertitude de la valeur assignée est **supérieure ou égale** à 0,3 fois l'écart-type d'aptitude σ_{pt} , il est recommandé d'utiliser le score z' , selon la norme NF ISO 13528 :2022.

Calculer le score z'_i pour chaque laboratoire selon la formule :

$$z'_i = \frac{(x_i - x_{pt})}{\sqrt{\sigma_{pt}^2 + u^2(x_{pt})}} \quad (6)$$

où $u^2(x_{pt})$ est l'incertitude-type composée de la valeur assignée x_{pt} .

Interpréter le score z' comme le score z .

Un exemple de scores z et z' calculé avec le logiciel Prolab est fourni en annexe F.

8.4.1.1.3 Cas des EILA qualitatifs

Se référer au chapitre 11 de la norme NF ISO 13528 :2022 et/ou de l'article 8.4 de la norme NF EN ISO 22117 :2019.

Comme indiqué dans ces deux normes, un **score** doit être utilisé pour évaluer la performance des laboratoires.

Depuis quelques années, deux scores de performance pour les EILA qualitatifs ont fait l'objet de publications : L-score (Uhlig *et al*, 2015) et a-score (Beavis *et al*, 2019). Le GT déconseille l'utilisation du a-score : le risque est de déclarer trop de laboratoires insatisfaisants.

Le GT a décidé de développer un nouveau score, plus adapté aux problématiques des EILA de l'Anses. Ce score appelé S-score (Baudry *et al*, 2024) a été développé plus particulièrement au travers d'une collaboration entre l'unité Pathologie et Bien-être des ruminants (PBER) du laboratoire de Ploufragan-Plouzané-Niort et la plateforme PAS.

a) L-score

Le principe en est le suivant : le score est défini sur la base d'un modèle logit prenant en compte le niveau de compétence des laboratoires (LCL) et le niveau de difficulté des tâches (LDT) constituées par les ESEA.

Ce score est calculé à partir des paramètres LCL et LDT estimés via la méthode du maximum de vraisemblance du modèle suivant:

$$\ln\left(\frac{p}{1-p}\right) = LCL - LDT \quad (7)$$

où p : probabilité d'un résultat correct

Un minimum de 10 laboratoires participants et 5 ESEA avec des résultats corrects et incorrects sont nécessaires.

L'interprétation du L-score proposée par les auteurs est la suivante :

- ▶ Si L-score < -2,0, les performances sont considérées comme « inférieures à la compétence moyenne »
- ▶ Si $-2,0 \leq \text{L-score} \leq 2,0$, les performances sont considérées comme « égales à la compétence moyenne »
- ▶ Si L-score > 2,0, les performances sont considérées comme « supérieures à la compétence moyenne ».

En cas d'utilisation du L-score, le GT recommande l'interprétation suivante :

- ▶ Seules les performances inférieures à la compétence moyenne (L-score < -2,0) sont considérées comme « insatisfaisantes »
- ▶ Pour des L-scores ≥ -2 , les performances sont considérées comme « satisfaisantes ».

Les calculs du L-score sont réalisables en utilisant le logiciel Prolab et un exemple est présenté en annexe G.

b) S-score

b.1) Principe

Ce score a fait l'objet d'une publication dans *Accreditation and Quality Assurance* (Baudry *et al*, 2024). Il s'applique à des ESEA ayant une Valeur Assignée (VA) binaire (ex : Positif versus Négatif, Détecté versus Non Détecté, Présence versus Absence). Il peut également être utilisé pour des ESEA à réponse qualitative à plus de 2 modalités en procédant à des regroupements (ex : Positif et Douteux versus Négatif, Positif versus Douteux et Négatif). Un résultat participant est incorrect s'il est différent de la VA.

Quatre types d'ESEA classés dans deux catégories de résultats ont été définis (Tableau 1) :

- ▶ Les ESEA à **résultat correct exigé** :
 - ESEA1 : ESEA individuelles.
 - ESEA2 : ESEA groupées.
- ▶ Les ESEA à **résultat correct attendu** :
 - ESEA3 : ESEA individuelles.
 - ESEA4 : ESEA individuelles.

Tableau 1 : Quatre types d'ESEA classés dans deux catégories de résultats pour évaluer la performance des laboratoires avec le S-score

Type d'ESEA	Catégorie de résultats				
	Résultats corrects exigés		Résultats corrects attendus		
	ESEA1	ESEA2		ESEA3	ESEA4
Nature des ESEA	Individuelle(s)	Répliquées	Non Répliquées	Individuelle(s)	Individuelle(s)
Détection	Facile	Difficile	Difficile/Facile ^(a)	Difficile	Difficile
Traitement des données	Individuel	Groupé		Individuel	Individuel
Valeur Assignée (VA)	Organisateur	Consensus des laboratoires participants (LC)		Organisateur	LC
Modalité binaire de la VA	Modalité définie par l'organisateur	Modalité majoritaire observée	Modalité majoritaire identique vérifiée sur chaque ESEA2	Modalité définie par l'organisateur	Modalité majoritaire statistiquement testée
Contraintes sur les participants et les ESEA	Aucune	Au moins 5 laboratoires et au moins un groupe de 6 ESEA2		Aucune	Au moins 8 laboratoires
Résultat correct défini selon	VA	Densité de probabilité binomiale		VA	LC

VA Valeur assignée, LC Consensus des laboratoires participants, (a) La détection peut être facile ou difficile ou les deux.

Au moins un des 2 types d'ESEA à résultat correct exigé doit être inclus dans le panel :

- ▶▶ **ESEA1** : ESEA pour lesquelles il n'est pas attendu d'erreur, dont la valeur assignée (VA) est définie par l'organisateur selon des critères « métiers » et dont le résultat correct est exigé individuellement pour chacune d'entre elles. Il s'agit d'ESEA dont la VA comporte peu de variabilité. La performance attendue sur l'ensemble de ESEA1 est 100% de résultats corrects.

Par exemple : échantillons fortement contaminés ou négatifs.

- ▶▶ **ESEA2** : ces ESEA peuvent être de 2 natures : répliquées ou non répliquées.

Un groupe constitué d'au moins 6 ESEA2 est nécessaire pour évaluer la performance des participants. Pour les ESEA2 répliquées, la VA est définie par la modalité binaire majoritaire observée ($\geq 50\%$) sur l'ensemble des ESEA2. Dans le cas des ESEA2 non répliquées, la même modalité majoritaire ($> 50\%$) doit être vérifiée pour chaque ESEA2.

Dans les deux cas, le pourcentage de résultats majoritaires est utilisé comme probabilité de succès de la loi binomiale.

Par exemple :

- ESEA2 répliquées : échantillons contaminés ou dilués à un même niveau qui peuvent être proche ou égal à la limite de détection (LD).
- ESEA2 non répliquées : échantillons contenant chacun une cible analytique différente mais analysés de façon groupée, par exemple différents sérotypes de *Salmonella* ou différentes familles d'analytes à des teneurs différentes.

Le panel peut être complété par des ESEA à résultat correct attendu :

- ▶▶ **ESEA3** : ESEA pour lesquelles la VA est définie par l'organisateur selon des critères « métiers ». Il s'agit d'ESEA pour lesquelles la variabilité est supérieure à la variabilité des ESEA1 tout en restant acceptable.
- ▶▶ **ESEA4** : ESEA pour lesquelles la VA est une valeur consensuelle, définie sur la base des résultats des laboratoires participants par la modalité statistiquement majoritaire ($> 50\%$). Il s'agit d'ESEA pour lesquelles la variabilité est supérieure à la variabilité des ESEA1 tout en restant acceptable.
- ▶▶ Les **ESEA3** et **ESEA4** sont interprétées (comme les ESEA1) de façon individuelle. Il peut s'agir d'échantillons pour lesquels la détection/non détection n'est pas exigée. La performance attendue sur les ESEA3 et les ESEA4 est 100% de résultats corrects.

Le nombre de laboratoires requis dépend de la constitution du panel et du type d'ESEA utilisé dans le panel :

- ▶▶ ESEA1 : pas de nombre minimal requis.
- ▶▶ ESEA2 : un minimum de 5 laboratoires et 6 ESEA2 (soit 30 résultats) est requis.
- ▶▶ ESEA3 : pas de nombre minimal requis.
- ▶▶ ESEA4 : un minimum de 8 laboratoires est requis.

Le principe et le mode de calcul du S-score sont présentés dans l'annexe H.

b.2) Interprétation du S-score

Elle est comparable au z score décrit dans la norme ISO/IEC 17043 : 2010 [1].

En cas de résultat incorrect, les ESEA à résultat correct exigé (ESEA1 et ESEA2) sont associées à une performance insatisfaisante et les ESEA à résultat correct attendu (ESEA3 et ESEA4) sont associées à une performance discutable.

Note : Pour les ESEA2, la performance est évaluée statistiquement (loi binomiale) par un nombre minimum et maximum de résultats corrects exigés au risque 1% (performance insatisfaisante) et attendus au risque 5% (performance discutable).

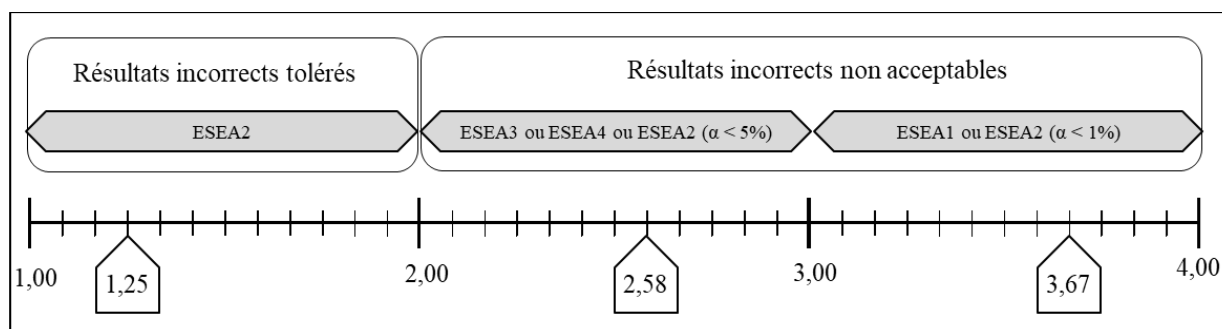
La valeur calculée du S-score est utilisée pour évaluer la compétence du laboratoire conformément aux règles suivantes :

S-score	Interprétation
$1 \leq \text{S-score} < 2$	Satisfaisante
$2 \leq \text{S-score} < 3$	Discutable
$3 \leq \text{S-score} \leq 4$	Insatisfaisante

- ▶ La partie entière (*Integer Part*, IP) 1, 2 ou 3 du S-score indique respectivement une performance satisfaisante, discutable ou insatisfaisante :
 - IP = 1 : aucun résultat incorrect pour les ESEA1, ESEA3 ou ESEA4 et un nombre acceptable de résultats corrects pour le groupe ESEA2 conformément au tableau J.1 ($\alpha = 0,05$) en Annexe J.
 - IP = 2 : aucun résultat incorrect pour les ESEA1 et au moins un résultat incorrect pour les ESEA3 ou ESEA4 et/ou un nombre acceptable de résultats corrects pour le groupe ESEA2 selon le tableau J.2 ($\alpha = 0,01$), mais pas selon le tableau J.1 ($\alpha = 0,05$) en Annexe J.
 - IP = 3 : au moins un résultat incorrect pour les ESEA1 ou un nombre inacceptable de résultats corrects pour le groupe ESEA2 selon le tableau J.2 ($\alpha = 0,01$) en Annexe J.
 - IP = 4 : est la valeur maximale du S-score qui indique que tous les résultats sont incorrects.
- ▶ La partie décimale (*Decimal Part*, DP) correspond au pourcentage de résultats incorrects en tenant compte de toutes les ESEA dont la VA est connue.

La figure 1 illustre la lecture du S-Score pour différentes valeurs sélectionnées.

Figure 1 : Lecture du S-Score. Illustration avec trois valeurs de S-score extrait du tableau I1 de l'annexe I.



b.3) Mise en œuvre

Pour la mise en œuvre du S-score, les recommandations suivantes sont préconisées :

► ESEA1

Critères de choix

- 100% de résultats (Positif ou Négatif) obtenus avec la méthode utilisée
- En santé animale, niveau exigible de détectabilité (NED) égal à 100%
- En fonction du principe de la méthode d'analyse, signal intermédiaire quantitatif de l'ESEA1 supérieur ou égal à 3 écart-types de la valeur du seuil analytique (cut-off de la méthode)

Utilisation du score ESEA1

En fonction du nombre de participants, on peut utiliser la valeur du score ESEA1 (S_{ESEA1}) comme indicateur d'alerte. Par exemple, pour un nombre de participants $n > 50$, un score ESEA1 supérieur à 3,2 peut constituer une alerte avec plus de 20% de laboratoires ayant une performance insatisfaisante. Pour un nombre de participants $n \leq 50$, $n \leq 30$ et $n \leq 10$, les valeurs 3,3, 3,4 et 3,5 correspondant respectivement à 30%, 40% et 50% de laboratoires ayant une performance insatisfaisante peuvent également être utilisés comme seuil d'alerte.

► ESEA2

Critères de choix

- ESEA2 répliquées
 - Pourcentage de résultats (Positif ou Négatif) supérieur ou égal à 50% et qui peut être proche du seuil analytique de la méthode
 - En santé animale, niveau exigible de détectabilité (NED) inférieur à 100%

- En fonction du principe de la méthode d'analyse, signal intermédiaire quantitatif des ESEA2 inférieur à 3 écart-types de la valeur du seuil analytique.
- ESEA2 non répliquées
 - Pourcentage de résultats strictement supérieur à 50% et si possible pas trop proche du seuil analytique pour les ESEA2 ayant le niveau le plus faible. Exemple 6 ESEA2 : 2 ESEA2 + 3 ESEA2 + 1 ESEA2 avec des niveaux estimés respectivement à 60%, 75% et 100% de résultats positifs.

Utilisation du score ESEA2

Pour les **ESEA2 répliquées**, plusieurs scores ESEA2 (S_{ESEA2}) supérieurs à 2 ou au moins un score ESEA2 supérieur à 3 constituent une éventuelle alerte. Une analyse de résultats est nécessaire pour expliquer cette hétérogénéité de résultats observés sur des ESEA2 répliquées dont l'homogénéité a été au préalable vérifiée. Dans le cas où le consensus participants est faible, proche de 50%, cette hétérogénéité peut s'expliquer par une absence de consensus des laboratoires participants : cas par exemple où 50% des laboratoires indiqueraient une majorité de résultats négatifs et les 50% restants, une majorité de résultats positifs.

Composition du panel

- Il est conseillé dans un panel contenant des ESEA2 de compléter le panel avec au minimum une ESEA1 afin de pallier les cas où :
 - Le nombre de participants serait inférieur à 5.
 - La modalité majoritaire des ESEA2 non répliquées ne serait pas la même individuellement pour chaque ESEA2 (i.e. non toutes majoritaires négatives ou non toutes majoritaires positives).
- Dans le cas où il est probable que le nombre de participants soit inférieur à 5, (par exemple traitement par méthode d'analyse (réactif de diagnostic), il est possible de prévoir dans le plan de campagne des dispositions alternatives : analyser les ESEA2 en ESEA3 et ne pas prendre en compte les ESEA2 dans le calcul du S-score.

Calcul du S-score

Dans le cas d'utilisation de plusieurs méthodes d'analyse et de variabilité intra-méthode, notamment pour des ESEA proches de la limite de détection, il est fortement recommandé de calculer un S-score pour chaque méthode. Si cela n'est pas possible pour des raisons d'effectifs (moins de 5 laboratoires), un regroupement approprié de certaines méthodes est possible. En cas d'impossibilité de réaliser des analyses par méthode, une analyse a posteriori des résultats sera nécessaire pour vérifier qu'il n'a pas de conséquence sur l'évaluation de la performance des participants.

» ESEA3 ou ESEA4

Critères de choix

- Résultat attendu qui génère des performances discutables dans le cas de résultat incorrect.
- Pourcentage de résultats obtenus pas trop proche du seuil analytique de la méthode.

Utilisation des scores ESEA3 et ESEA4

La valeur du score ESEA3 (S_{ESEA3}) peut être utilisée selon les règles appliquées aux score ESEA1 mais en utilisant la partie entière 2. Ainsi, pour un nombre de participants $n > 50$, un score ESEA3 supérieur à 2,2 peut constituer une alerte avec plus de 20% de laboratoires ayant une performance discutable. Pour un nombre de participants $n \leq 50$, $n \leq 30$ et $n \leq 10$, les valeurs 2,3, 2,4 et 2,5 correspondant respectivement à 30%, 40% et 50% de laboratoires ayant une performance discutable, peuvent également être préconisées comme seuil d'alerte.

La valeur du score ESEA4 (S_{ESEA4}) peut être utilisée selon les règles appliquées aux score ESEA3. Ainsi pour un nombre de participants $n > 50$, un score ESEA4 supérieur à 2,2 peut constituer une alerte avec plus de 20% de laboratoires ayant une performance discutable. Pour un nombre de participants $30 < n \leq 50$, la valeur 2,3 correspondant à 30% des laboratoires ayant une performance discutable peut également être préconisée comme seuil d'alerte.

Composition du panel

Il est conseillé de limiter le nombre d'ESEA3 et ESEA4 proches des limites de détection (favorisées les ESEA2 répliquées) afin de ne pas générer un nombre de laboratoires ayant une performance discutable trop important.

L'annexe I donne des exemples d'application du S-score calculé avec l'application QualiScore.

c) Autres paramètres

D'autres paramètres peuvent être calculés et utilisés en parallèle d'un score, à condition qu'ils soient justifiés et que des limites d'acceptabilité aient été définies (cf. note 1), comme par exemple les taux suivants :

- » Taux de spécificité.
- » Taux de sensibilité.
- » Taux d'exactitude.

Ces termes sont définis dans le glossaire (*Glossaire relatif à la validation des méthodes d'analyses* » (ANSES/FGE/0155 [lien internet](#)).

Le **taux de spécificité** r_{SP} est calculé par la formule :

$$r_{SP} = \frac{n_-}{E_{(n_{-tot})}} \times 100 \quad (8)$$

où n_- est le nombre de résultats négatifs trouvés et $E_{(n_{-tot})}$ est le nombre total de résultats négatifs attendus.

Le **taux de sensibilité** r_{SE} est calculé par la formule :

$$r_{SE} = \frac{n_+}{E_{(n_{+tot})}} \times 100 \quad (9)$$

ou n_+ est le nombre de résultats positifs trouvés et $E_{(n_{+tot})}$ est le nombre total de résultats positifs attendus.

Le **taux d'exactitude** r_{AC} est calculé par la formule :

$$r_{AC}(\%) = \frac{n_- + n_+}{n_{tot}} \times 100 \quad (10)$$

ou n_{tot} est le nombre total de résultats .

Note 1 : Ces paramètres **peuvent être utilisés pour évaluer la performance, sous réserve de définir des limites d'acceptation** soit justifiées par l'expérience, soit correspondant à des limites réglementaires ou à des valeurs issues d'autres sources, selon le domaine concerné. Par exemple un laboratoire est déclaré satisfaisant s'il obtient une spécificité $\geq 95\%$.

Note 2 : Le nombre d'ESEA doit être suffisant pour évaluer le taux de sensibilité ou le taux de spécificité. Le nombre d'ESEA peut être fixé en suivant les règles décrites dans l'article 8.4 de la norme NF EN ISO 22117:2019, prenant en compte à la fois le nombre de niveaux et le nombre d'ESEA.

Note 3 : Les taux de spécificité, de sensibilité et d'exactitude peuvent être obtenus directement avec le S-score.

8.4.1.2 Cas d'un score global (performance multicritères)

Une approche graphique de combinaison de scores de performance sur plusieurs campagnes est présentée dans la norme NF ISO 13528:2022, chapitre 10 Méthodes graphiques de description des scores de performance.

Il est possible de calculer également des **scores composés** : se reporter à l'article 9.9 Scores composés de performance de la norme NF ISO 13528:2022.

Certains EILA impliquent la mesure de plusieurs variables simultanément par les participants, (par exemple : méthodes multirésidus pour les pesticides). Si l'évaluation individuelle pour chaque variable séparément reste la base pour déterminer l'aptitude, il peut être aussi pertinent d'avoir une évaluation globale des performances d'un laboratoire sur l'ensemble des variables mesurées dans le cadre de l'EILA. Dans ce cas, un score z global combinant les scores z individuels pour chaque variable peut être utilisé.

Il faut cependant noter que l'usage des scores z composés n'est souvent pas recommandé : cf. note de l'article 9.9.5 de la norme NF ISO 13528:2022. Les principales limites évoquées sont la difficulté d'interprétation d'un score global et le fait que le calcul d'un score moyen ou cumulé peut masquer un score individuel élevé, si d'autres scores sont bas ou opposés.

Plusieurs méthodes ont été proposées pour calculer des scores z composés (Senyuva *et al*, 2006, Medina-Pastor *et al*, 2010a et 2010b).

Le GT recommande d'utiliser la somme de scores z pondérés. Cette méthode est décrite dans le protocole général des LRUE pour les résidus de pesticides (*General protocol for EU proficiency tests for pesticides residues in food and feed*, 2013). Il faut noter que, dans ce contexte, cette méthode est appliquée uniquement aux laboratoires ayant détecté un grand nombre de substances présentes dans les ESEA (par exemple au moins 90%) et ne présentant pas de faux positifs.

L'utilisation d'un système de pondération permet en effet d'associer un poids plus important aux scores élevés et ainsi d'éviter qu'un grand nombre de scores z « acceptables » compense trop fortement des scores « inacceptables ».

Le GT recommande deux méthodes pour le calcul d'un score z composé. L'annexe F présente en détail les modalités de calculs.

Il est recommandé de consulter la plateforme PAS lorsqu'il est nécessaire de calculer un score composé.

8.4.2 Estimation de la fidélité

Si nécessaire, la variabilité des résultats fournis par les participants peut être estimée graphiquement par le graphe k de Mandel, conformément à la norme NF ISO 5725-2:2020 et l'article 10.4.3 de la norme NF ISO 13528:2022. Le logiciel Prolab offre l'opportunité de calculer le paramètre k de Mandel.

L'article 10.6 de la norme NF ISO 13528 :2022 propose également d'autres types de graphiques pour étudier la répétabilité.

8.4.3 Représentation graphique des performances

Il est recommandé de prévoir une représentation graphique des performances des laboratoires participants.

Justifier clairement le type de graphique utilisé et le mentionner dans le plan de campagne.

Exemples de représentation graphique, selon la norme NF ISO 13528:2022 :

- Graphe représentant la justesse des laboratoires participants : graphe sur les scores z et z' , ou autres scores.
- Graphe sur la fidélité des laboratoires participants : graphe k de Mandel.

Un exemple de graphes pour des scores z et z' obtenu avec le logiciel Prolab est fourni en annexe F.

9 Vérification de la cohérence inter-EILA

La norme NF ISO 13528:2022 décrit, dans le chapitre 10, plusieurs approches graphiques pour suivre l'aptitude d'un laboratoire sur plusieurs EILA.

Trois approches sont proposées : la carte de contrôle de Shewhart, la carte de contrôle de CUSUM et la représentation graphique du biais normalisé.

Ces trois approches sont utilisées :

- ▶ pour le score z ou score z' ,
- ▶ pour suivre uniquement l'évolution cumulée du score z ou z' d'un laboratoire.

Ces approches peuvent être également utilisées pour suivre globalement dans le temps les performances du système analytique (équipement, opérateur, méthode, etc.) au sein d'un laboratoire (suivi intra-laboratoire). L'application de ces approches sur des résultats issus d'ESEA de type contrôle qualité interne (CQI) ne relève pas de l'évaluation d'aptitude des laboratoires mais permet :

- ▶ d'accepter les résultats du jour au sein du laboratoire,
- ▶ et de suivre l'évolution du système analytique via des analyses de tendance.

En aucun cas :

- ▶ les informations issues du suivi des performances du système analytique ne peuvent pas être utilisées pour évaluer l'aptitude des laboratoires,
- ▶ et le suivi des performances des laboratoires sur plusieurs EILA ne peut pas être utilisé pour l'évaluation des performances des méthodes d'analyse (ou des kits).

10 Présentation des résultats

10.1 Rapport d'EILA

Le rapport d'EILA doit être conforme aux exigences de l'article 4.8 de la norme NF EN ISO/CEI 17043:2010 : en particulier, celle d'émettre un rapport unique, qui comporte les données relatives aux résultats de tous les participants ainsi qu'une indication de leur performance, de manière anonyme. Le rapport doit rappeler tous les critères d'évaluation des participants, y compris lorsque des laboratoires sont exclus (laboratoire non évalué), ou qu'il y a des données manquantes (laboratoire non évaluable).

Transmettre ce rapport aux laboratoires participants (dépôt sur LEILA ou envoi). Lorsqu'il doit être également adressé à un tiers (par exemple, le prescripteur), avertir les participants au préalable, lors de leur inscription, de cet envoi. Si une levée de l'anonymat des participants accompagne cet envoi, informer également au préalable les participants.

Mentionner dans le rapport les règles utilisées : variable et unité (mesurande) pour lesquelles les calculs ont été faits, gestion des données manquantes et atypiques, méthodes de calculs utilisées (méthode robuste ou non), etc.

Le rapport peut comporter des résumés et des annexes contenant les données brutes, des résultats intermédiaires ou toute information explicitant les calculs ou les procédures utilisés. Une synthèse générale en début de rapport est conseillée. Des données complémentaires à l'EILA, non prises en compte dans l'évaluation de la performance, peuvent être fournies à **titre informatif** en annexe, en identifiant clairement le statut de ces annexes, par exemple, par une mention du type « Données complémentaires non prises en compte dans l'évaluation de la performance ».

Si l'EILA doit être invalidé, présenter dans le rapport un argumentaire justifiant l'annulation de l'EILA, les conséquences et les solutions envisagées. Il est également recommandé de décrire les dispositions d'annulation dans le plan de campagne.

Etablir des règles documentaires pour la gestion des rapports : identification, historique des versions, circuit de signatures, références bibliographiques, liens avec d'autres documents (tels que rapport précédent), méthode d'analyse, etc.

Il est recommandé de veiller à harmoniser la terminologie et les symboles utilisés. Par exemple, ne pas confondre u et U , σ et $\hat{\sigma}$, \bar{x} (moyenne arithmétique) et x^* (moyenne robuste), etc. Utiliser la terminologie et les symboles de la norme NF ISO 13528:2022. Définir tout nouveau symbole introduit.

Il est important d'être vigilant au rendu des résultats, en particulier au nombre de chiffres significatifs affecté aux résultats : valeur assignée, écart-type d'aptitude, scores, ... Pour cela, respecter le chapitre 5 de la norme NF ISO 13528:2022.

Présenter et justifier tout écart au plan de campagne dans le rapport d'EILA.

Mentionner le résultat de l'évaluation de l'homogénéité et de la stabilité dans le rapport. Toutefois, la norme NF EN ISO/CEI 17043:2010 n'impose pas un compte-rendu détaillé de cette évaluation, qui peut faire l'objet d'un rapport distinct, dans le cas où un participant souhaiterait en disposer. Cette approche, qui reste au choix du coordonnateur, a l'avantage de simplifier la rédaction du rapport d'EILA et de le recentrer sur l'objectif principal, c'est-à-dire l'évaluation de l'aptitude des participants.

Apporter une attention particulière au processus de génération du rapport, notamment au bon fonctionnement des importations de tableaux ou de graphiques : insertion d'image, « copier/coller ».

Valider tout outil informatique utilisé pour générer des rapports.

Protéger le rapport sous format PDF pour éviter toute modification par un tiers et préciser que la reproduction du rapport n'est autorisée que sous sa forme intégrale.

Dans le cas d'envoi des rapports sous forme dématérialisée, il est recommandé d'établir une convention de preuve avant la réalisation de l'envoi et *a minima*, d'assurer la traçabilité des envois.

Par ailleurs, indiquer clairement la ou les restrictions concernant l'usage des données du rapport : cf. article 4.8.4 de la norme NF EN ISO/CEI 17043:2010 sur la politique d'utilisation des rapports.

Note : Une trame de rapport est proposée en version française et anglaise (voir les documents et références en chapitre 3.1 de ce guide).

10.2 Autres données ou modalités de présentation des résultats

Il peut être utile de fournir aux participants leurs résultats individuels, sous forme par exemple de feuille de résultats individuels, notamment pour faciliter l'accès des participants à leurs résultats. Cependant, une feuille de résultats individuels ne peut pas se substituer au rapport d'EILA. Cette feuille peut être soit constitutive du rapport d'EILA (elles sont alors anonymes), soit adressée individuellement à chaque participant, au choix du coordonnateur. Déposer dans LEILA seulement le rapport d'EILA.

Dans le cas des EILA organisés par les laboratoires Anses, généralement non séquentiels et de durée relativement courte, il n'est pas recommandé de préparer des rapports d'EILA ne prenant en compte qu'une partie des laboratoires participants. La production de rapports intermédiaires est proscrite à l'Anses.

10.3 Suivi des actions correctives suite à l'EILA

Conformément aux missions de LNR/LRUE des laboratoires de l'Agence, le rapport d'EILA anonymisé doit être adressé dans le cas général au prescripteur, accompagné de la décodification des participants sur demande.

Le laboratoire de référence s'assure de l'efficacité des mesures correctives prises par les laboratoires ayant obtenu un (ou des) résultat(s) non satisfaisant(s) par tout moyen qu'il juge adapté et efficient. Selon la demande du prescripteur, le suivi des mesures correctives lui est adressé, selon le moyen ou le support convenu.

Note : La gestion des écarts est réalisée selon les recommandations du guide ANSES/PG/0142 (document uniquement disponible en interne).

11 Conclusions

Ce guide a pour principale vocation d'harmoniser les règles et les outils nécessaires à une bonne organisation expérimentale des EILA et utilisation des statistiques dans le champ des EILA organisés par les laboratoires de l'Agence. Des outils de calcul sont disponibles pour analyser les données :

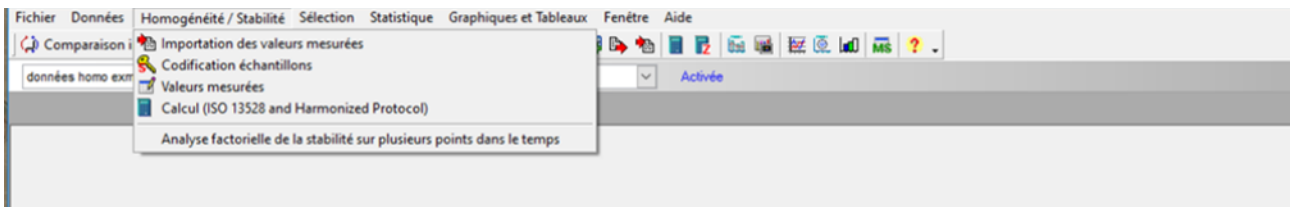
- Pour les EILA quantitatifs, le logiciel Prolab répond à l'ensemble des analyses statistiques nécessaires pour organiser un EILA de ce type (homogénéité, stabilité, évaluation de la performance) et répond aux vérifications demandées dans la norme NF ISO 13528:2022.
- Pour les EILA qualitatifs, le L-score peut être calculé via Prolab et le S-score à partir d'un complément Excel (QualiScore) mis à disposition des utilisateurs.

Annexe A : Homogénéité - : Calculs réalisés avec Prolab

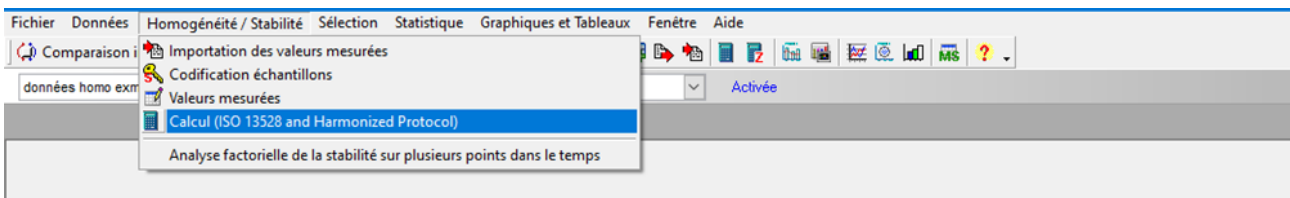
A.1 Aspects généraux

Le module "**Homogénéité/Stabilité**" dans Prolab permet d'ouvrir plusieurs sous-modules d'importation des données, de codification des échantillons, de vérification des valeurs mesurées et de conduite des calculs.

Dans un premier temps il faut créer une base de données correspondante à l'EILA. La création de la base de données et l'importation des données sont décrites dans la notice d'utilisation de Prolab disponible sous ENNOV.

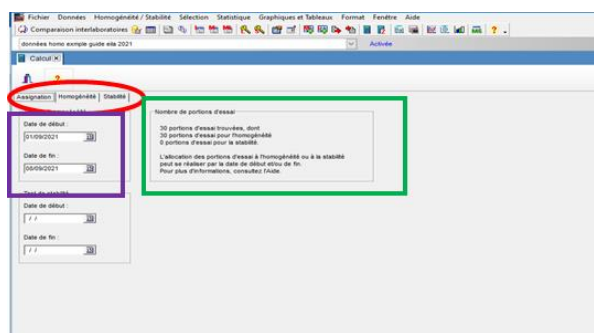


Pour effectuer les calculs, activer dans le menu déroulant le sous-module "**Calcul (ISO 13528 and Harmonized Protocol)**".



La fenêtre qui s'ouvre compte 3 onglets (ellipse rouge) : **Assignment**, **Homogénéité**, **Stabilité**.

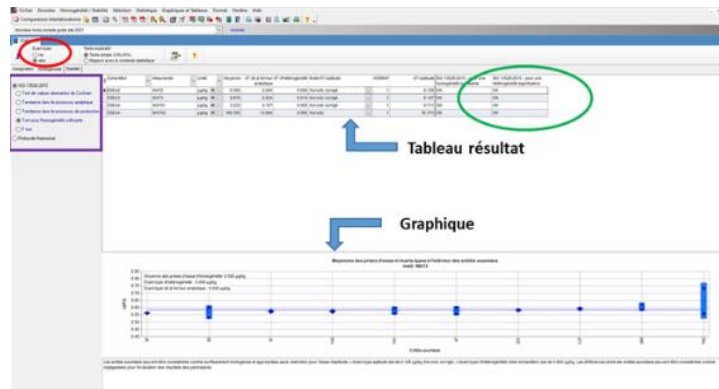
L'onglet "**Assignment**" permet de définir les dates d'homogénéité (rectangle violet). Dans cet onglet il est également possible de vérifier le nombre d'ESEA utilisées pour l'homogénéité et la stabilité (rectangle vert).



Activer les calculs en cliquant sur l'onglet "Homogénéité". Les calculs peuvent se faire selon la norme NF ISO 13528:2015/Annexe B ou selon le protocole harmonisé IUPAC.

A.2 Calculs selon "Norme NF ISO 13528:2015"

La fenêtre de résultats se présente ainsi.



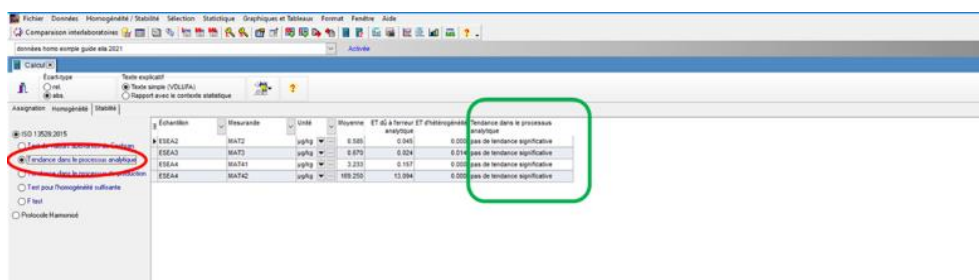
Un tableau des résultats et des conclusions (ellipse verte) sur l'homogénéité est fourni ainsi qu'un graphique.

Les écart-types (ellipse rouge) peuvent être exprimés en valeur relative ou en valeur absolue.

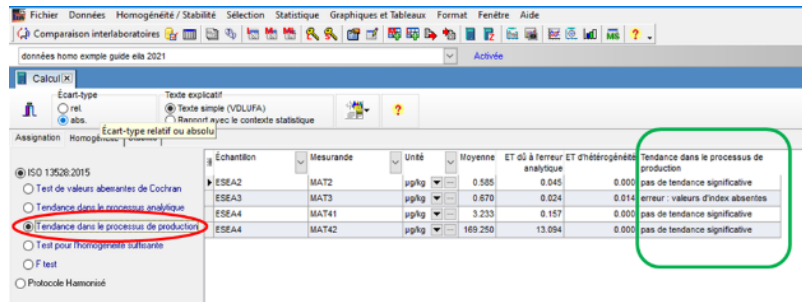
La visualisation des tableaux de résultats (rectangle violet) peut se faire pour 5 types d'analyse :

- Test de valeurs aberrantes de Cochran.
- Tendance dans le processus analytique.
- Tendance dans le processus de production.
- Test pour l'homogénéité suffisante.
- Test F.

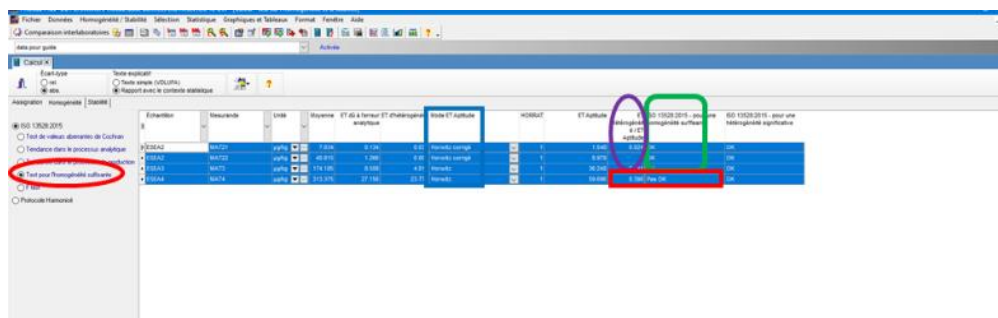
Les résultats concernant les **tendances dans le processus analytique** (ellipse rouge) sont affichés dans la colonne "**Tendance dans le processus analytique**" (encadrés verts). L'absence de tendance est traduite par "pas de tendance significative".



La conclusion sur les **tendances dans le processus de production** (ellipse rouge) est affichée dans la colonne "**Tendance dans le processus de production**" (encadré vert). Cette colonne indique la présence ou l'absence de tendance significative. Si aucun index n'est indiqué et que l'analyse est demandée, l'indication suivante sera fournie "Erreur : valeurs d'index absentes".



L'évaluation de l'homogénéité est affichée en cliquant sur "**Test pour l'homogénéité suffisante**" (ellipse rouge). La conclusion est obtenue dans la colonne "**ISO 13528:2015 – Pour une homogénéité suffisante**" (rectangle vert).



La conclusion se traduit par "**OK**" lorsqu'il est conclu à l'homogénéité (rectangle vert) et "**pas OK**" (rectangle rouge) quand la condition suivante n'est pas respectée : $Ss/\sigma_{pt} \leq 0.3$ avec Ss l'écart-type inter échantillons (appelé ET hétérogénéité dans Prolab). À noter que la valeur du rapport est affichée dans la colonne "**ET Hétérogénéité/ET Aptitude**" (ellipse violette).

Vérifier également que les calculs de l'écart-type d'aptitude sont correctement sélectionnés (encadré bleu).

Quatre mode de calcul sont possibles : :

- "**ET aptitude de la comparaison interlaboratoire**" lorsqu'une ET d'aptitude consensuel est calculé et que l'homogénéité est vérifiée *a posteriori*,
- "**Saisie Manuelle**" pour prendre en compte un ET d'aptitude défini *a priori* (perception, ET aptitude issu d'EILA(s) antérieurs),
- "**Horwitz**" qui correspond à la formule classique d'Horwitz,
- "**Horwitz corrigé**" qui correspond à la formule de Thompson lorsque la concentration massique c est $< 1.2 \times 10^{-7}$ (il est fixé par défaut à 22 %).

Sous le graphique, une synthèse est fournie sur les conclusions obtenues.

Moyennes des prises d'essai et écarts-types à l'intérieur des entités soumises ESEA2 / mat21

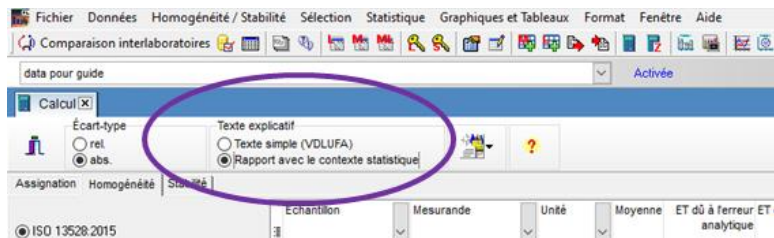
Pour le test d'homogénéité, 10 des entités soumises pour échantillon ESEA2 ont été sélectionnées de façon aléatoire, et le mesurande mat21 a été déterminé 2 fois. La moyenne sur les 10 entités soumises est de 7.034 µg/kg. L'écart-type intra-échantillon s(analytical) (+précision analytique) est de 0.134 µg/kg, et l'écart-type d'hétérogénéité (inter-échantillon) est de 0.036 µg/kg.

F test
Selon le test F, l'écart-type d'hétérogénéité (inter-échantillon) ne diffère pas significativement de 0 (au niveau 5 %). Les entités soumises peuvent donc être considérées comme suffisamment homogènes, selon ce critère.

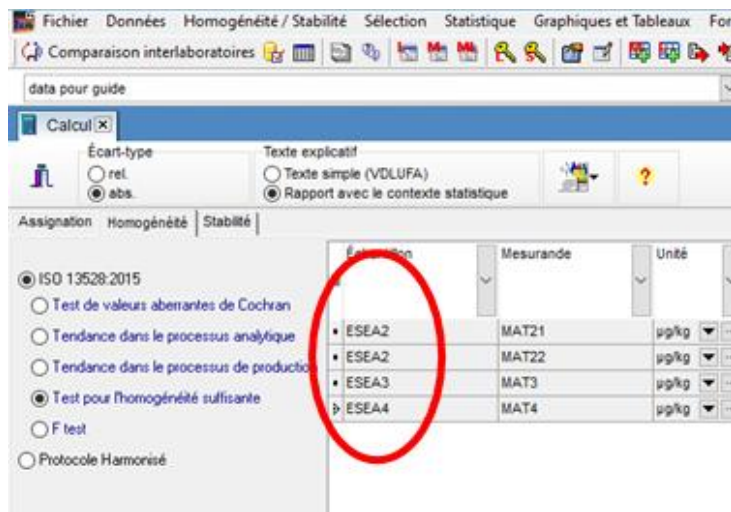
ISO 13528:2015 - Test d'homogénéité adéquate
Selon la ISO 13528:2015, l'écart-type d'hétérogénéité (inter-échantillon), caractérisant la variabilité entre les entités soumises, ne doit pas dépasser 30 % de l'écart-type aptitude. L'écart-type d'hétérogénéité est inférieur à 30 % de l'écart-type aptitude de 1.548 µg/kg (Horwitz corrigé). Les entités soumises peuvent donc être considérées comme suffisamment homogènes, selon la ISO 13528:2015.

ISO 13528:2015 - Test d'hétérogénéité significative
Aucune hétérogénéité significative n'a pu être identifiée. Les entités soumises peuvent donc être considérées comme suffisamment homogènes.

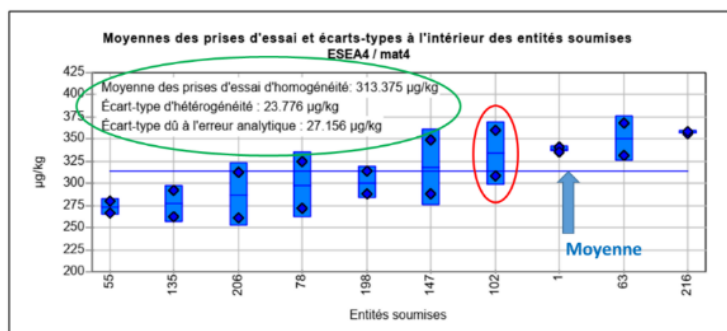
Pour obtenir le plus d'informations possible dans cette synthèse, il faut sélectionner dans "Texte explicatif", "Rapport avec le contexte statistique" (ellipse violette). Sinon, le choix "Texte simple (VDLUFA)" propose une conclusion concise sur l'homogénéité.



Les graphiques sont édités par ESEA en sélectionnant l'ESEA (ellipse rouge).



Sur le graphique, les points et les répétitions sont fournis (ellipse rouge) ainsi que le résumé des paramètres statistiques (ellipse verte). La ligne bleue matérialise la valeur moyenne générale obtenue avec l'ensemble des prises d'essai.

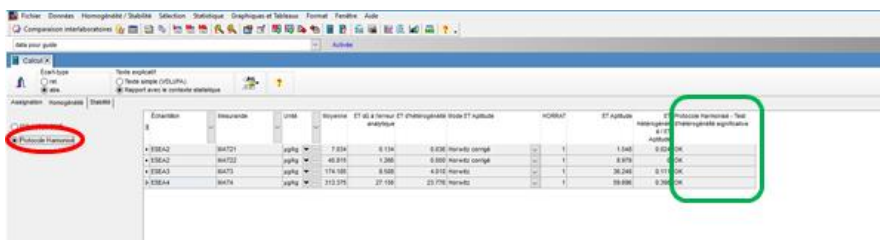


À noter que ces calculs sont strictement identiques entre la version 2015 et la version 2022 de la norme NF ISO 13528.

A.3 Calculs selon "Protocole harmonisé"

Il est possible de s'assurer d'une homogénéité suffisante en appliquant le protocole harmonisé décrit dans le rapport technique IUPAC. Cette approche est présentée en annexe B.

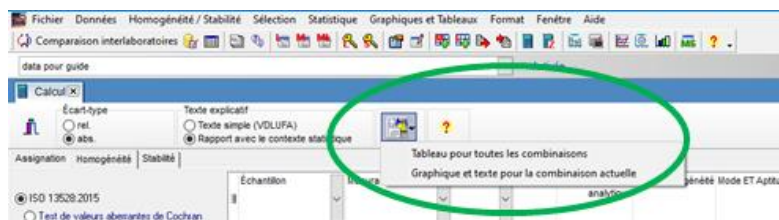
Dans Prolab en sélectionnant "Protocole harmonisé" (ellipse rouge), les calculs sont effectués et la fenêtre suivante est obtenue :



Les conclusions sont affichées dans la colonne "Protocole harmonisé – Test d'homogénéité significative" (rectangle vert). "OK" correspond à une déclaration d'homogénéité suffisante selon IUPAC. "Pas OK" rejette la déclaration d'homogénéité suffisante.

A.4 Rapports

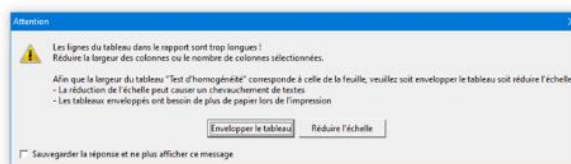
Il est possible de visualiser, d'imprimer et de sauvegarder les rapports en cliquant sur le bouton



Une fenêtre s'ouvre alors, et il est possible de voir :

- soit un tableau avec toutes les ESEA (combinaisons) ;
- soit le graphique et le texte correspondant pour une ESEA choisie.

Dans le cas où le tableau contient un nombre de colonnes ou un nombre de lignes trop élevé, l'écran suivant apparaît :



Si « envelopper le tableau » est sélectionné, le tableau est réparti sur plusieurs lignes :

Test d'homogénéité

Après des valeurs statistiques et résultats

Echantillon	Mesurande	Unité	Moyenne	ET de à l'erreur analytique	ET d'hétérogénéité	Mode ET Aptitude	HORRAT	ET Aptitude
ESEA2	MAT21	µg/kg	7.034	0.134	0.096	Horra FZ corrigé	1	1.548
ESEA2	MAT22	µg/kg	40.815	1.286	0.000	Horra FZ corrigé	1	8.979
ESEA3	MAT3	µg/kg	174.185	8.508	4.010	Horra FZ	1	36.248
ESEA4	MAT4	µg/kg	313.315	27.158	23.776	Horra FZ	1	59.898

Echantillon	Mesurande	E1 Hétérogénéité / E1 Aptitude	ISO 15628:2015 - pour une homogénéité satisfaisante
ESEA2	MAT21	0.028 OK	
ESEA2	MAT22	0.000 OK	
ESEA3	MAT3	0.111 OK	
ESEA4	MAT4	0.388 Pas OK	

Echantillon	Mesurande	ISO 15628:2015 - pour une hétérogénéité significative
ESEA2	MAT21	OK
ESEA3	MAT3	OK
ESEA3	MAT3	OK
ESEA4	MAT4	OK

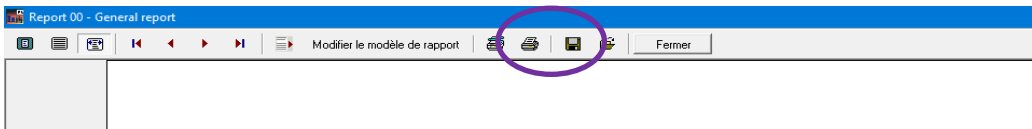
Si « Réduire l'échelle » est sélectionné, le tableau est dimensionné pour la page :

Test d'homogénéité

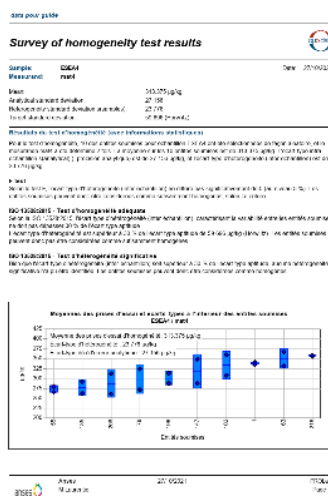
Après des valeurs statistiques et résultats

Echantill	Mesuran	Unité	Moye	ET de à l'erreur d'hétérogénéité	ET Mode ET Aptitude	HORR	ET Aptitude	ET Hétérogénéité ISO 15628:2015 - pour une Aptitude homogénéité satisfaisante	ISO 15628:2015 pour une hétérogénéité significative
ESEA2	MAT21	µg/kg	7.034	0.134	0.096	Horra FZ corrigé	1	0.021 OK	OK
ESEA2	MAT22	µg/kg	40.815	1.286	0.000	Horra FZ corrigé	1	0.000 OK	OK
ESEA3	MAT3	µg/kg	174.1	8.508	4.010	Horra FZ	1	0.111 OK	OK
ESEA4	MAT4	µg/kg	313.3	27.158	23.776	Horra FZ	1	0.388 Pas OK	OK

Une fois ce choix effectué, vous pouvez sauvegarder  en format PDF le tableau ou l'imprimer :



Si vous sélectionnez la deuxième option « Le graphique et le texte correspondant pour une ESEA choisie », vous obtiendrez :



A.5 Exemple de jeu de données

Les données fournies sont les données qui ont permis de réaliser l'ensemble des calculs décrits :

code eila	échantillon	mesurande	homogeneite/Stabilite	date	code entité soumise	valeur mesurée 1	valeur mesurée 2	Unité	facteur Horwitz	Index de Production	Index de mesure VM 1	Index de mesure VM 2
guide	ESEA2	MAT21	homogeneite	01/10/2018	511		7.23	6.900 µg/kg			13	11
guide	ESEA2	MAT21	homogeneite	01/10/2018	922		7.000	7.230 µg/kg			3	9
guide	ESEA2	MAT21	homogeneite	01/10/2018	801		6.960	6.860 µg/kg			8	7
guide	ESEA2	MAT21	homogeneite	01/10/2018	471		7.340	7.020 µg/kg			6	15
guide	ESEA2	MAT21	homogeneite	01/10/2018	1751		6.950	6.950 µg/kg			14	1
guide	ESEA2	MAT21	homogeneite	01/10/2018	1151		6.900	7.080 µg/kg			18	12
guide	ESEA2	MAT21	homogeneite	01/10/2018	1421		7.010	7.080 µg/kg			5	16
guide	ESEA2	MAT21	homogeneite	01/10/2018	951		7.180	7.130 µg/kg			20	2
guide	ESEA2	MAT21	homogeneite	01/10/2018	701		7.120	6.990 µg/kg			17	4
guide	ESEA2	MAT21	homogeneite	01/10/2018	1951		6.960	6.800 µg/kg			10	19
guide	ESEA2	MAT22	homogeneite	01/10/2018	512	40.700	40.200 µg/kg			13	11	
guide	ESEA2	MAT22	homogeneite	01/10/2018	922	40.000	40.700 µg/kg			3	9	
guide	ESEA2	MAT22	homogeneite	01/10/2018	802	41.400	39.800 µg/kg			8	7	
guide	ESEA2	MAT22	homogeneite	01/10/2018	472	43.800	42.100 µg/kg			6	15	
guide	ESEA2	MAT22	homogeneite	01/10/2018	1752	41.000	41.700 µg/kg			14	1	
guide	ESEA2	MAT22	homogeneite	01/10/2018	1152	41.000	40.400 µg/kg			18	12	
guide	ESEA2	MAT22	homogeneite	01/10/2018	1422	39.300	40.800 µg/kg			5	16	
guide	ESEA2	MAT22	homogeneite	01/10/2018	952	42.100	38.100 µg/kg			20	2	
guide	ESEA2	MAT22	homogeneite	01/10/2018	702	42.300	39.700 µg/kg			17	4	
guide	ESEA2	MAT22	homogeneite	01/10/2018	1952	40.600	40.600 µg/kg			10	19	
guide	ESEA3	MAT3	homogeneite	04/10/2018	232	162.60	151.30 µg/kg			5	3	
guide	ESEA3	MAT3	homogeneite	04/10/2018	97	173.60	175.00 µg/kg			12	7	
guide	ESEA3	MAT3	homogeneite	04/10/2018	123	174.80	179.40 µg/kg			10	16	
guide	ESEA3	MAT3	homogeneite	04/10/2018	90	168.80	172.00 µg/kg			11	17	
guide	ESEA3	MAT3	homogeneite	04/10/2018	203	167.30	180.80 µg/kg			9	19	
guide	ESEA3	MAT3	homogeneite	04/10/2018	98	168.40	191.00 µg/kg			13	1	
guide	ESEA3	MAT3	homogeneite	04/10/2018	146	177.80	184.30 µg/kg			15	2	
guide	ESEA3	MAT3	homogeneite	04/10/2018	26	185.30	184.30 µg/kg			20	4	
guide	ESEA3	MAT3	homogeneite	04/10/2018	211	171.40	168.50 µg/kg			6	14	
guide	ESEA3	MAT3	homogeneite	04/10/2018	189	172.10	189.50 µg/kg			8	18	
guide	ESEA4	MAT4	homogeneite	02/10/2018	55	266.3000	279.6000 µg/kg			1	11	
guide	ESEA4	MAT4	homogeneite	02/10/2018	102	307.8000	359.4000 µg/kg			4	17	
guide	ESEA4	MAT4	homogeneite	02/10/2018	63	368.1000	331.6000 µg/kg			2	10	
guide	ESEA4	MAT4	homogeneite	02/10/2018	147	348.5000	287.7000 µg/kg			3	18	
guide	ESEA4	MAT4	homogeneite	02/10/2018	1	340.9000	336.2000 µg/kg			5	16	
guide	ESEA4	MAT4	homogeneite	02/10/2018	206	261.5000	312.6000 µg/kg			19	6	
guide	ESEA4	MAT4	homogeneite	02/10/2018	198	313.4000	288.1000 µg/kg			12	15	
guide	ESEA4	MAT4	homogeneite	02/10/2018	216	356.4000	359.1000 µg/kg			7	20	
guide	ESEA4	MAT4	homogeneite	02/10/2018	135	291.7000	262.3000 µg/kg			13	8	
guide	ESEA4	MAT4	homogeneite	02/10/2018	78	324.3000	272.0000 µg/kg			9	14	

Annexe B : Homogénéité selon l'approche IUPAC

Lorsqu'il n'est pas possible de conclure à l'homogénéité des ESEA selon l'approche classique décrite dans l'article B.2.2 de la norme NF ISO 13528:2022, il est possible d'étendre le critère de vérification.

En effet, il est possible que l'erreur d'échantillonnage et la répétabilité de la méthode d'analyse ne permettent pas de conclure sur l'homogénéité des ESEA. Pour pallier ce problème, il est possible de prendre en compte la variabilité supplémentaire liée à la méthode d'analyse, via l'approche décrite dans le rapport technique IUPAC et référencée dans l'article B.2.3 de la norme NF ISO 13528:2022 (Thompson *et al*, 2006).

Cette approche consiste à vérifier, pour déclarer les ESEA suffisamment homogènes, que :

$$S_s \leq \sqrt{c}$$

Où :

- S_s est l'écart-type inter-échantillons et,
- c est une valeur calculée par l'équation suivante :

$$c = F_1 \sigma_{allow}^2 + F_2 S_w^2$$

Où :

- $\sigma_{allow}^2 = (0.3\sigma_{pt})^2$ (le carré de la condition d'homogénéité)
- S_w^2 est l'écart-type intra-échantillons
- F_1 et F_2 sont des quantiles statistiques, normalisés, dépendant du nombre d'ESEA et qui sont analysés dans des **conditions de répétabilité (n=2)**. Le tableau suivant précise les valeurs de F_1 et F_2 . Utiliser les valeurs de F_1 et F_2 en gras pour 10 ESEA analysés 2 fois.

Tableau B.1 : Facteurs F_1 et F_2 dans des conditions de répétabilité (n=2)

g	20	19	18	17	16	15	14	13	12	11	10	9	8	7
F_1	1.59	1.60	1.62	1.64	1.67	1.69	1.72	1.75	1.79	1.83	1.88	1.94	2.01	2.10
F_2	0.57	0.59	0.62	0.64	0.68	0.71	0.75	0.80	0.86	0.93	1.01	1.11	1.25	1.43

Dans le cas où le nombre de répétitions est > 2 , calculer les quantiles F_1 et F_2 . Pour cela utiliser le mode de calcul indiqué dans l'article B.2.3 de la norme NF ISO 13528:2022, ou consulter la plateforme PAS.

Cette approche est disponible dans Prolab (voir annexe A).

À noter que ces calculs sont strictement identiques entre la version 2015 et la version 2022 de la norme NF ISO 13528.

Annexe C : Homogénéité - Cas où le nombre d'ESEA par niveau est faible

Étant donné les particularités des EILA qualitatifs organisés à l'Anses (nombre d'ESEA constituant le panel souvent plus élevé que pour les EILA quantitatifs pour permettre une évaluation valide, faible disponibilité en matériaux de référence, etc.) et lorsque le nombre d'ESEA produites par lot de production est inférieur à 100, il est possible de diminuer le nombre d'ESEA pour le test d'homogénéité de certains lots de production d'échantillons, selon les recommandations du document « *ILAC Discussion Paper on Homogeneity and Stability Testing* » et du FD ISO Guide 35, en particulier l'article 7.4.1.2 concernant les études d'homogénéité menées sur des propriétés qualitatives et l'article 7.4.1.3 qui concernent les petits lots de production (<100).

Déterminer le nombre d'échantillons à tester selon le tableau C.1, après une analyse de risque pertinente justifiant cette réduction : par exemple, expérience de l'organisateur sur le type d'échantillons et d'analyte, et/ou historique d'organisation d'EILA portant sur le même type d'échantillons et analyte, préparés selon les mêmes procédures.

Tableau C.1 : Détermination du nombre d'échantillons à tester selon le nombre d'échantillons préparés par lot de production

Nombre d'échantillons préparés par lot de production	Nombre d'échantillons à tester
≤ 19	3
20-39	4
40-49	5
50-59	6
60-69	7
70-79	8
80-95	9
>95	10

Vérifier les lots considérés comme critiques avec 10 ESEA analysés deux fois, quel que soit la taille du lot de production.

Annexe D : Stabilité - Calculs réalisés avec Prolab

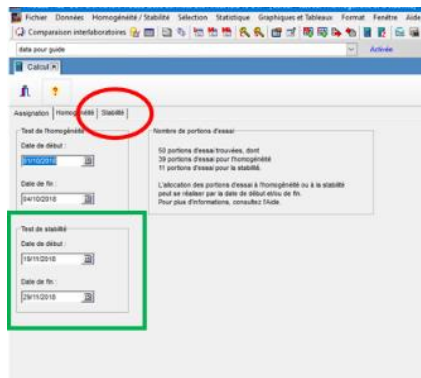
D.1 Aspects généraux

Le module "**Homogénéité/Stabilité**" dans Prolab permet d'ouvrir plusieurs sous-modules d'importation des données, de codification des échantillons, de vérification des valeurs mesurées et de conduite des calculs.

Créer une base de données correspondant à l'EILA. La création de la base de données et l'importation des données sont décrites dans la notice d'utilisation de Prolab disponible sous ENNOV.

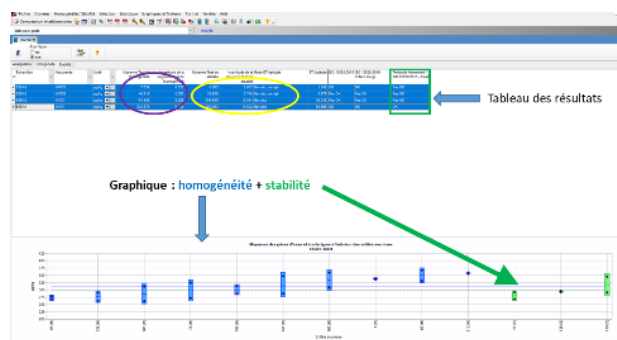
D.2 Calculs

Une fois que les données de stabilité ont été importées, entrer les dates de début et de fin d'étude de la stabilité (encadré vert). Puis sélectionner « stabilité » (ellipse rouge) :



Dans la fenêtre d'affichage des calculs de stabilité apparaissent :

- un tableau de résultats ;
- un graphique regroupant les valeurs obtenues pour l'homogénéité (points bleus) et la stabilité (points verts) ;
- un résumé avec les conclusions sur la stabilité (en bas de la fenêtre) :



Le tableau de résultats contient diverses colonnes (encadré vert, agrandi ci-dessous) incluant la moyenne des valeurs obtenues pour l'homogénéité (ellipse violette) et la stabilité (ellipse jaune) ainsi que les incertitudes de ces moyennes. Le mode de calcul de l'écart-type d'aptitude est fourni, ainsi que sa valeur. La vérification de la stabilité se fait selon trois procédures :

- Annexe B de la norme NF ISO 13528:2015 (paragraphe B.5.1) (approche classique).
- Ajout de l'incertitude de la moyenne d'homogénéité et de stabilité (norme NF ISO 13528:2015, paragraphe B.5.2.c) (critère élargi).
- L'approche décrite dans le rapport technique IUPAC (Thompson *et al*, 2006) utilisant un test t et reprise dans le paragraphe B.5.4 de la norme NF ISO 13528:2015 :

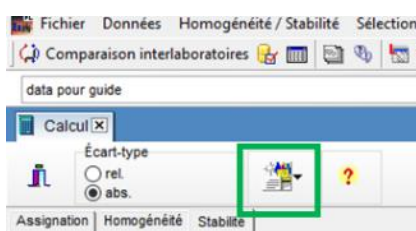
ISO 13528:2015	ISO 13528:2015 - Critère élargi	Protocole Harmonisé / ISO 13528:2015 - t test
OK	OK	Pas OK
Pas OK	Pas OK	Pas OK
Pas OK	Pas OK	Pas OK
OK	OK	OK

Si les échantillons sont stables, la conclusion est "**OK**" ; dans le cas contraire, "**pas OK**".

À noter que ces calculs sont strictement identiques entre la version 2015 et la version 2022 de la norme NF ISO 13528.

D.3 Rapport

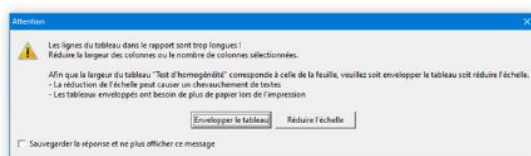
En cliquant sur le bouton  il est possible de visualiser, imprimer et sauvegarder les rapports :



Une fenêtre s'ouvre alors, et il est possible de voir:

- soit un tableau avec toutes les ESEA (combinaisons) ;
- soit le graphique et le texte correspondant à l'ESEA choisie.

Dans le cas où le tableau contient un nombre de colonnes ou un nombre de lignes trop élevé, l'écran suivant apparaît :



Si « envelopper le tableau » est sélectionné, le tableau est réparti sur plusieurs lignes :

DATE: 2015/03/02

Test de stabilité

Appréhension de la stabilité des données de stabilité

Paramètre	Mesure	Unité	Moyenne Test de l'hétérogénéité	Incertitude de la moyenne Test de l'hétérogénéité	Moyenne Test de stabilité	Incertitude de la moyenne Test de stabilité
ESEA2	10A71	µg/g	7,034	0,032	0,032	0,032
ESEA2	10A72	µg/g	40,815	0,283	35,000	0,283
ESEA2	10A73	µg/g	174,185	2,288	206,393	2,288
ESEA4	10A74	µg/g	313,316	3,986	291,983	3,986

Évaluation

Paramètre	Mesure	Unité	Protocole Harmonisé ISO 1528:2015 - Critère étroit	Protocole Harmonisé ISO 1528:2015 - Critère élargi
ESEA2	10A71	µg/g	OK	OK
ESEA2	10A72	µg/g	OK	OK
ESEA2	10A73	µg/g	OK	OK
ESEA4	10A74	µg/g	OK	OK

Si « réduire l'échelle » est sélectionné le tableau est dimensionné pour la page :

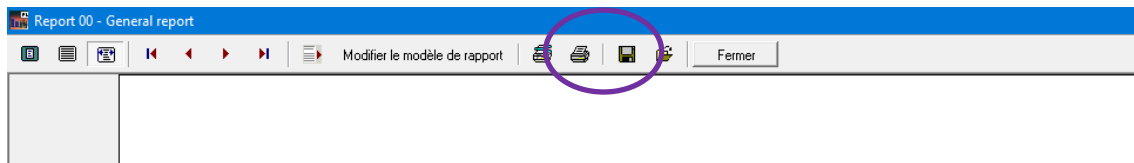
DATE: 2015/03/02

Test de stabilité

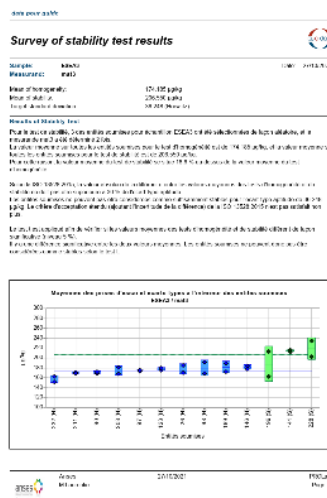
Appréhension de la stabilité des données de stabilité

Paramètre	Mesure	Unité	Moyenne Test de l'hétérogénéité	Incertitude de la moyenne Test de l'hétérogénéité	Moyenne Test de stabilité	Incertitude de la moyenne Test de stabilité	CT ISO 1528:2015 - Critère étroit	CT ISO 1528:2015 - Critère élargi	Protocole Harmonisé ISO 1528:2015 - Critère étroit	Protocole Harmonisé ISO 1528:2015 - Critère élargi
ESEA2	10A71	µg/g	7,034	0,032	0,032	0,032	OK	OK	OK	OK
ESEA2	10A72	µg/g	40,815	0,283	35,000	0,283	OK	OK	OK	OK
ESEA2	10A73	µg/g	174,185	2,288	206,393	2,288	OK	OK	OK	OK
ESEA4	10A74	µg/g	313,316	3,986	291,983	3,986	OK	OK	OK	OK

Une fois ce choix effectué, vous pouvez sauvegarder  en format pdf le tableau ou l'imprimer  :



Si vous sélectionnez la deuxième option « Le graphique et le texte correspondant pour une ESEA choisie », vous obtiendrez :



D.4 Exemple de jeu de données

Les données fournies sont les données qui ont permis de réaliser l'ensemble des calculs décrits :

date	code entité soumise	valeur mesurée 1	valeur mesurée 2	Unité	facteur Horwitz	Index de Production	Index de mesure VM 1	Index de mesure VM 2
29/11/2018	172	6.76	6.64	µg/kg			1	3
29/11/2018	221	7.04	6.61	µg/kg			2	5
29/11/2018	159	6.52	6.53	µg/kg			4	6
29/11/2018	172	36.4	37.3	µg/kg			1	3
29/11/2018	222	35.6	33	µg/kg			2	5
29/11/2018	159	35.7	35	µg/kg			4	6
22/11/2018	141	212.3	215.7	µg/kg			6	2
22/11/2018	196	161.9	212.7	µg/kg			5	3
22/11/2018	228	234.3	202.4	µg/kg			4	1
19/11/2018	57	292.7	268.8	µg/kg			1	3
19/11/2018	126	297.2	292.3	µg/kg			2	5
19/11/2018	190	346	290.9	µg/kg			4	6

Annexe E : Homogénéité et stabilité - Calculs réalisés avec QualiScore

E.1 Exemple de calculs d'homogénéité

Pour les essais qualitatifs utilisant le S-score, l'homogénéité des ESEA2, ESEA3 et ESEA4 peut être vérifiée statistiquement en utilisant la loi binomiale au risque de 5 %. QualiScore permet de calculer un score d'homogénéité (S_{HOM}). Les résultats obtenus sont présentés dans l'exemple ci-dessous avec une VA assignée égale à 75% qui correspond au pourcentage moyen de résultats positifs attendus sur chacune des 2 séries de 10 ESEA. Le score d'homogénéité (1,35) inférieur à 2 indique que les ESEA sont homogènes.

Code EILA : EILA Evaluation de l'homogénéité

Valeur assignée Type ESEA			75 % (P)									
			ESEA2	ESEA2	ESEA2	ESEA2	ESEA2	ESEA2	ESEA2	ESEA2	ESEA2	ESEA2
Codes laboratoires	S-score	Homogène	Tube 1	Tube 2	Tube 3	Tube 4	Tube 5	Tube 6	Tube 7	Tube 8	Tube 9	Tube 10
Prise d'essai 1	1,30	Oui	N	N	N	P	P	P	P	P	P	P
Prise d'essai 2	1,40	Oui	P	P	P	N	N	N	N	P	P	P



S_{HOM} 1,35
Vérification de l'homogénéité : Loi binomiale (n= 10, p= 75%) au risque 5 % [5 à 10] *
Conclusion : ESEA homogènes

NB: un score ≥ 2 indique au risque 5 % que le pourcentage de résultats corrects est significativement différent du pourcentage attendu de 75 %

E.2 Exemple de calculs de stabilité

Pour les essais qualitatifs utilisant le S-score, QualiScore permet de calculer un score de stabilité (S_{STAB}) sur les ESEA2, ESEA3 et ESEA4 selon la même procédure statistique utilisée pour vérifier l'homogénéité (loi binomiale, risque à 5 %).

Les résultats obtenus sont présentés dans l'exemple ci-dessous avec une VA assignée égale à 90% : le nombre de résultats corrects n'est pas acceptable (Prise d'essai 2) avec un S-score égal à 2,40. Un score de stabilité supérieur ou égal à 2 indique que les ESEA ne sont pas stables.

Code EILA : EILA Evaluation de la stabilité

Valeur assignée Type ESEA	90 % (P)											
	ESEA2	ESEA2	ESEA2	ESEA2	ESEA2	ESEA2	ESEA2	ESEA2	ESEA2	ESEA2		
Codes laboratoires	S-score	Stable	Tube 1	Tube 2	Tube 3	Tube 4	Tube 5	Tube 6	Tube 7	Tube 8	Tube 9	Tube 10
Prise d'essai 1	1,30	Oui	N	N	N	P	P	P	P	P	P	P
Prise d'essai 2	2,40	Non	P	P	P	N	N	N	N	P	P	P



S_{STAB} 2,35

Vérification de la stabilité : Loi binomiale (n= 10, p= 90%) au risque 5 % [7 à 10] *

Conclusion : ESEA NON stables

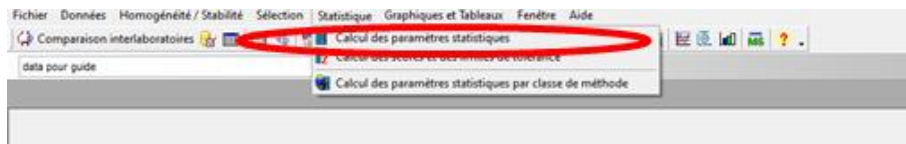
NB: un score ≥ 2 indique au risque 5 % que le pourcentage de résultats corrects est significativement différent du pourcentage attendu de 90 %

Annexe F : Exemple de calcul des scores z et z' avec le logiciel Prolab

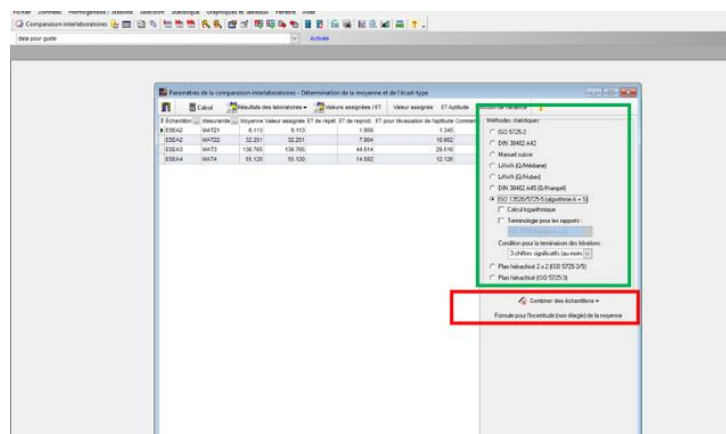
L'évaluation de la performance se fait en deux étapes : une première permettant de déterminer la valeur assignée et l'ET d'aptitude, la seconde permettant de calculer les scores.

F.1 Valeur assignée et ET aptitude

Dans le module "Statistique", sélectionner "Calcul des paramètres statistiques" :



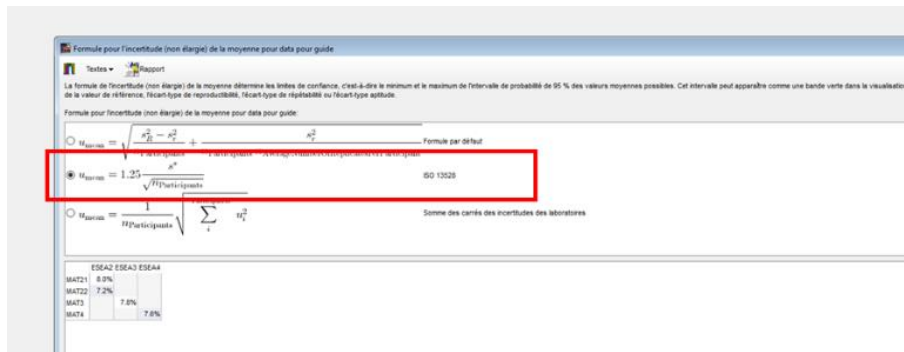
La fenêtre suivante apparaît :



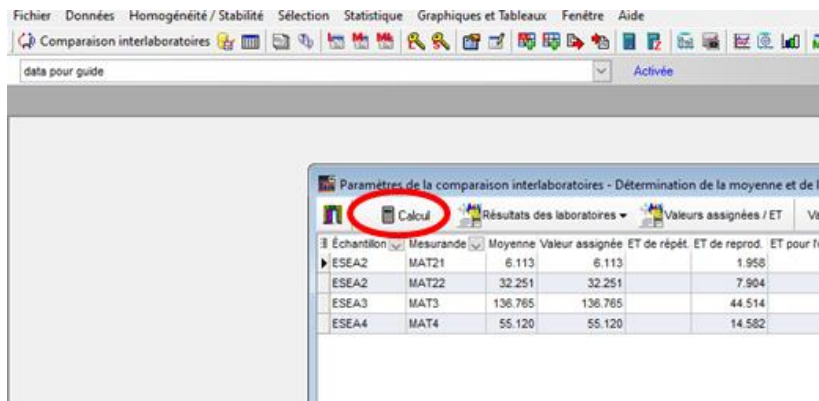
Sélectionner la méthode statistique parmi les propositions qui figurent à droite (encadré vert).

Pour sélectionner une approche consensuelle, cocher "ISO13528/5725-5 (Algorithme A+S)" avec "Terminologie pour les rapports" fixée à "ISO13528" et "Condition pour la terminaison des itérations" fixée à 4 chiffres significatifs (au moins 30 itérations).

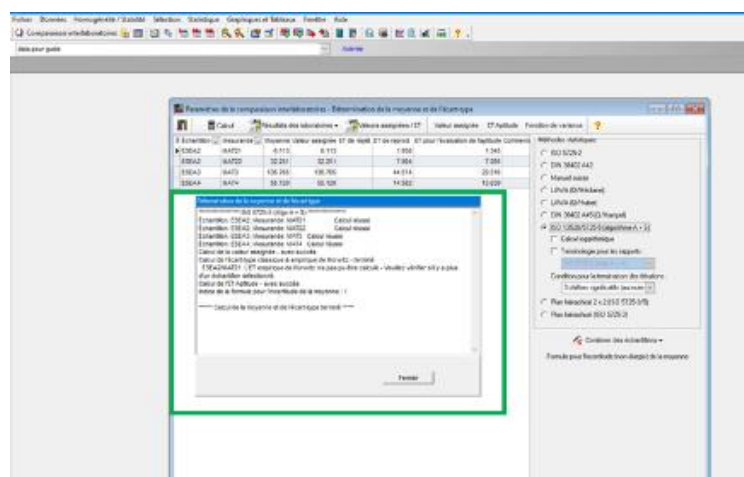
S'assurer également que le calcul de l'incertitude de la valeur assignée est correctement effectué. Pour cela, s'assurer dans la rubrique "**Formule pour l'incertitude (non élargie) de la moyenne**" (encadré rouge) que la formule n°2 est sélectionnée :




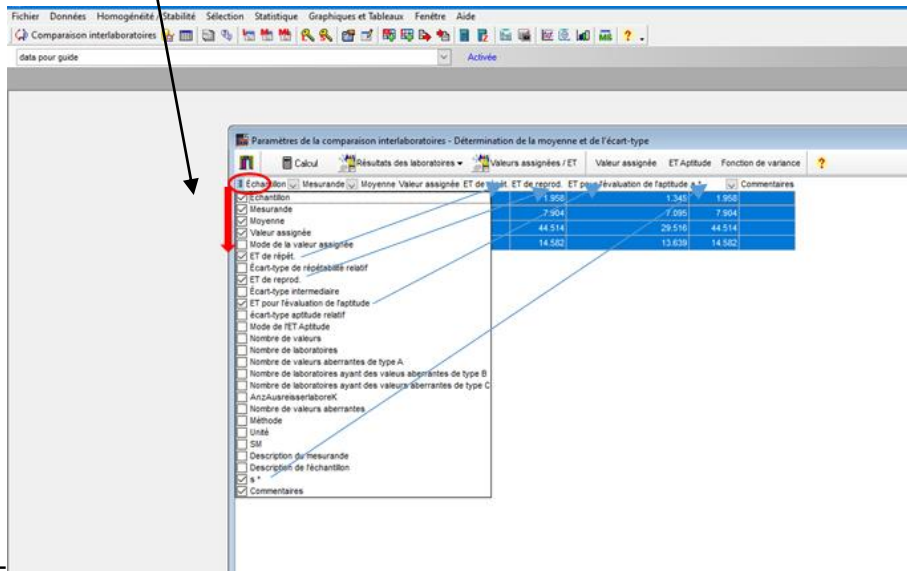
Cette vérification effectuée, effectuer les calculs en cliquant sur "**Calcul**".



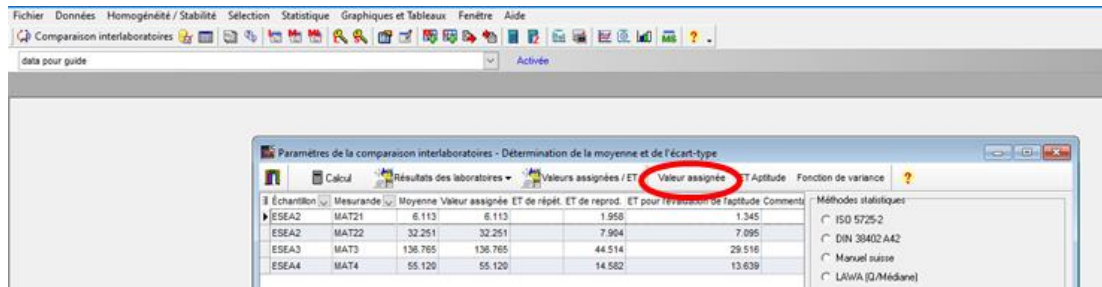
Une fois les calculs effectués, la valeur assignée, son incertitude et l'ET pour l'évaluation de l'aptitude s'affichent à l'écran. De plus, une fenêtre apparaît indiquant le bon déroulement des opérations. Fermer cette fenêtre pour passer à la suite des opérations (rectangle vert) :



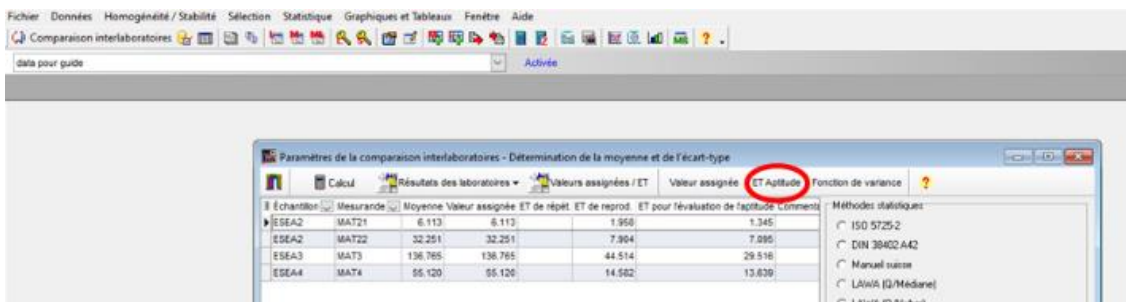
À noter que le bouton  situé en haut à gauche (ellipse rouge) ouvre un sous-menu qui permet de sélectionner les colonnes visibles à l'écran. Quel que soit la fenêtre ouverte, ce bouton permet de sélectionner les colonnes de données qui seront affichées à l'écran (voir ci-dessous) :



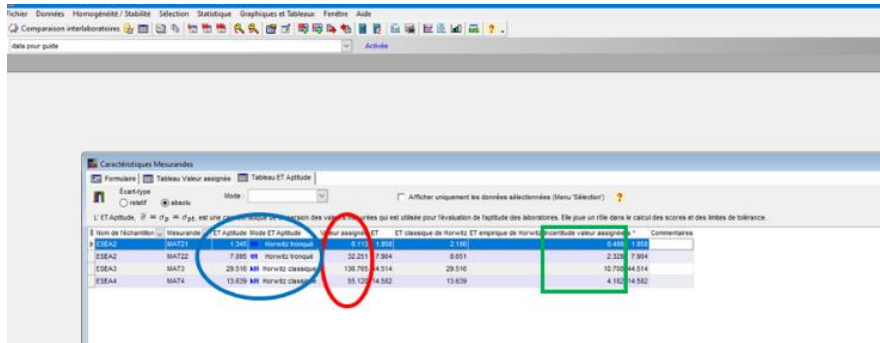
En cliquant sur "**Valeur assignée**" (ellipse rouge) :



ou sur "**Tableau ET Aptitude**" (ellipse rouge) :



Le même tableau fournit la valeur assignée (ellipse rouge), son incertitude (rectangle vert) et l'ET pour l'évaluation de l'aptitude (ellipse bleue) :

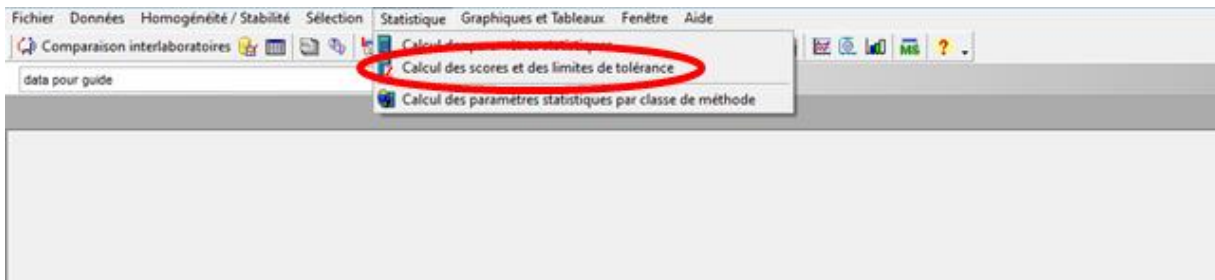


Écart-type	Mode	Ajustement		ET		Valeur assignée		Incertitude	
ESAE1	MAT1	1,326	Horvitz bonjour	32,201	7,854	8,051	2,328	7,954	1,326
ESAE2	MAT2	7,895	Horvitz bonjour	32,201	7,854	8,051	2,328	7,954	1,326
ESAE3	MAT3	29,510	Horvitz classique	136,785	4,514	29,510	10,700	44,514	29,510
ESAE4	MAT4	13,626	Horvitz classique	55,125	4,552	13,626	4,162	14,552	13,626

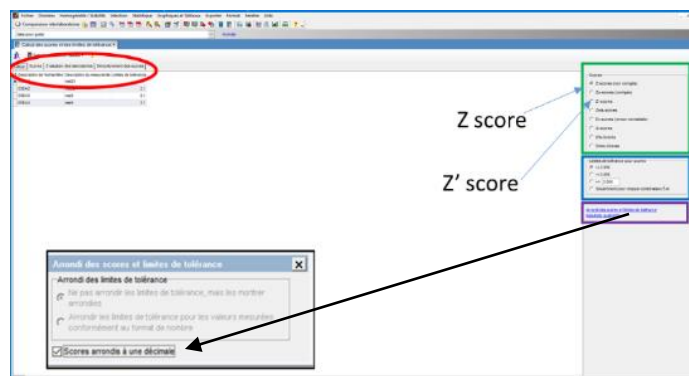
L'étape suivante consiste à calculer les scores.

F.2 Calcul des scores

Dans le module "Statistique", sélectionner le sous-module "Calcul des scores et des limites de tolérance" :



La fenêtre suivante apparaît :



La partie entourée en rouge est composée de 4 onglets :

- "Calcul" ;
- "Scores" ;
- "Évaluation des participants" ;
- "Dénombrement des scores".

Pour effectuer le calcul, définir au préalable le score qui va être utilisé. Pour cela, sélectionner à droite (rectangle vert) le score (z, z', etc.). En première intention utiliser le z-score.

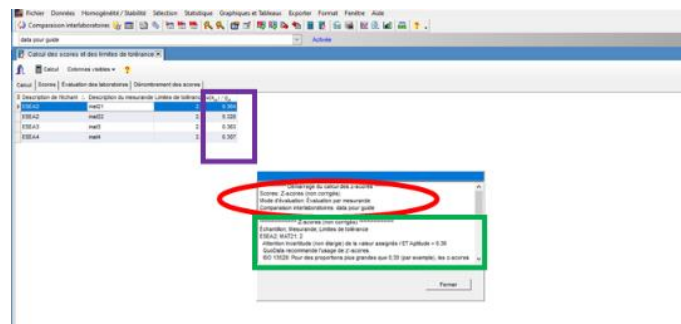
Vérifier ensuite les "**Limites de tolérance pour les scores**", situées à droite (cadre bleu). Vérifier également l'"**Accord des scores et limites de tolérance**" (cadre violet). Un sous-menu apparaît dans lequel vous pourrez définir soit les arrondis des limites de tolérance, soit l'arrondi des scores. Choisir "**Scores arrondis à une décimale**".

Lorsque ces points ont été vérifiés, effectuer les calculs en cliquant sur "**Calcul**".

Une fenêtre de calcul s'ouvre (ellipse rouge) pour signaler ou non la réussite du calcul.

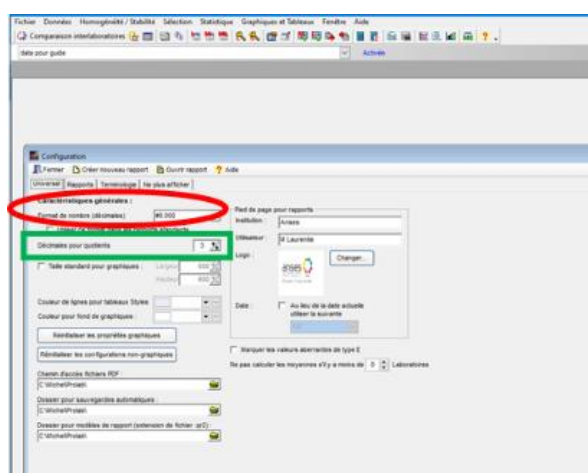
En particulier l'utilisation du score z n'est possible que si l'incertitude de la valeur assignée est inférieure à 0,3 écart-type d'aptitude ($u(x_{pt}) < 0.3\sigma_{pt}$) :

- Dans le cas où cette condition n'est pas satisfaite, un message d'alerte apparaît (rectangle vert) avec une recommandation d'utiliser le score z'.
- Dans le cas où la condition est vérifiée la réussite du calcul est indiquée.

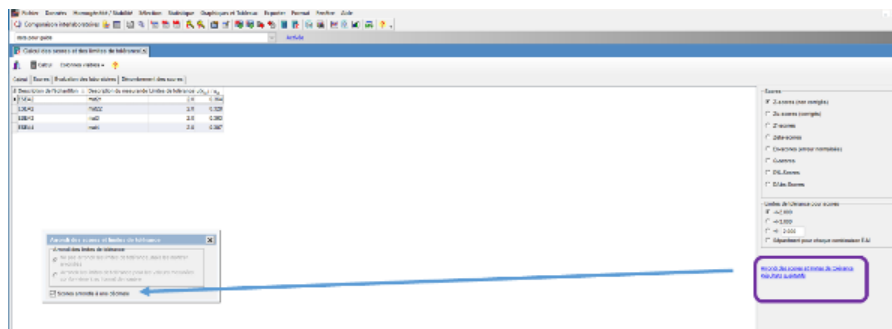


À noter que la valeur du rapport $u(x_{pt}) / \sigma_{pt}$ est indiquée (rectangle violet).

Vous pouvez définir le nombre de décimales des rapports via le module "**Fichier**", sous-module "**Configuration**".



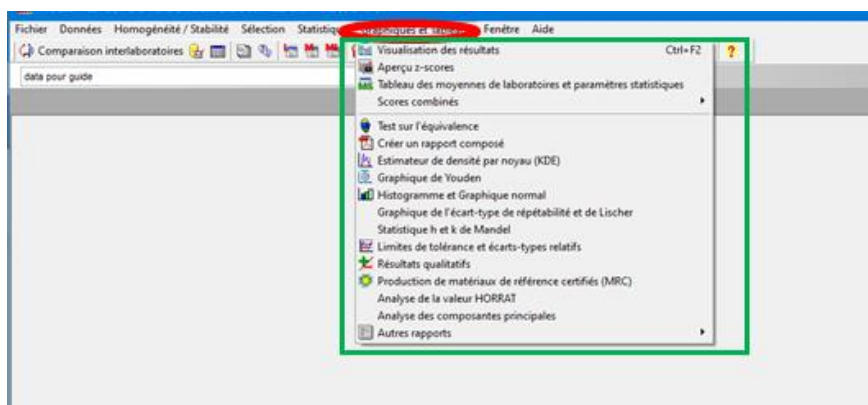
Le nombre de décimales des résultats est fixé par «**format de nombre (décimales)**» (ellipse rouge). Vous pouvez également définir le nombre de décimales pour les quotients via «**Décimales pour quotients**» (rectangle vert). À noter que pour les scores ou le rapport $u(x_{pt}) / \sigma_{pt}$, une décimale est suffisante. Cependant pour les scores, utiliser plutôt le sous-module "**Calcul des scores et des limites de tolérance**" pour définir les arrondis.



L'étape suivante est la visualisation des résultats et l'édition des tableaux et graphiques pour le rapport.

F.3 Visualisation des résultats

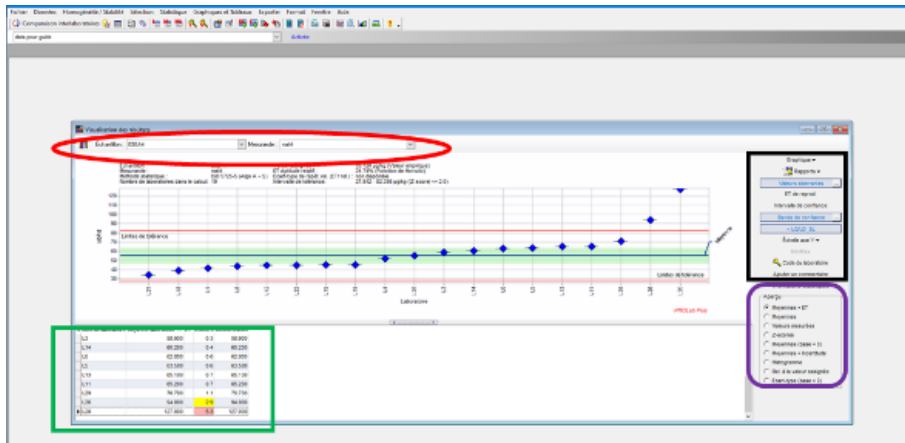
Pour visualiser les résultats, ouvrir le module "**Grahiques et Tableaux**" (ellipse rouge). Dans ce module, différentes options sont proposées (rectangle vert) :



Seules les plus utilisées sont présentées :

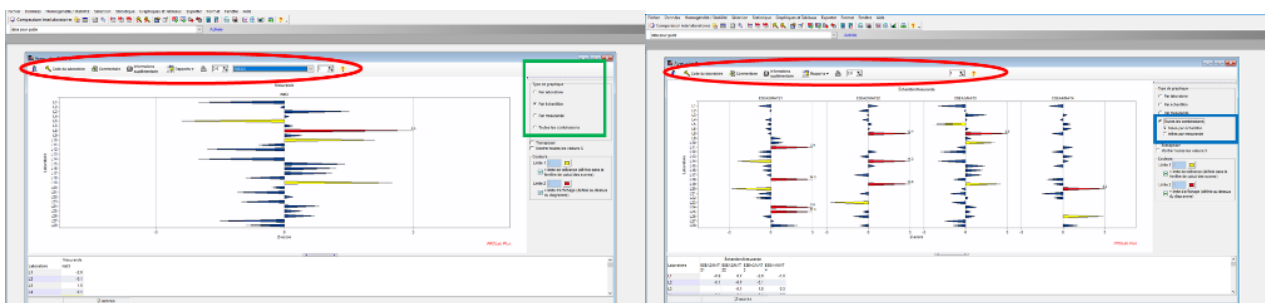
- "**Visualisation des résultats**" : cette option permet de visualiser les résultats sous forme d'histogrammes. Les résultats peuvent être des moyennes, des valeurs mesurées, des scores.... Effectuer le choix en bas à droite dans la fenêtre (rectangle violet). En haut à

gauche, des menus déroulants permettent de sélectionner l'ESEA et l'analyte à visualiser (ellipse rouge). En-dessous du graphique, la moyenne par participant, les valeurs individuelles et les scores sont indiqués (rectangle vert) :

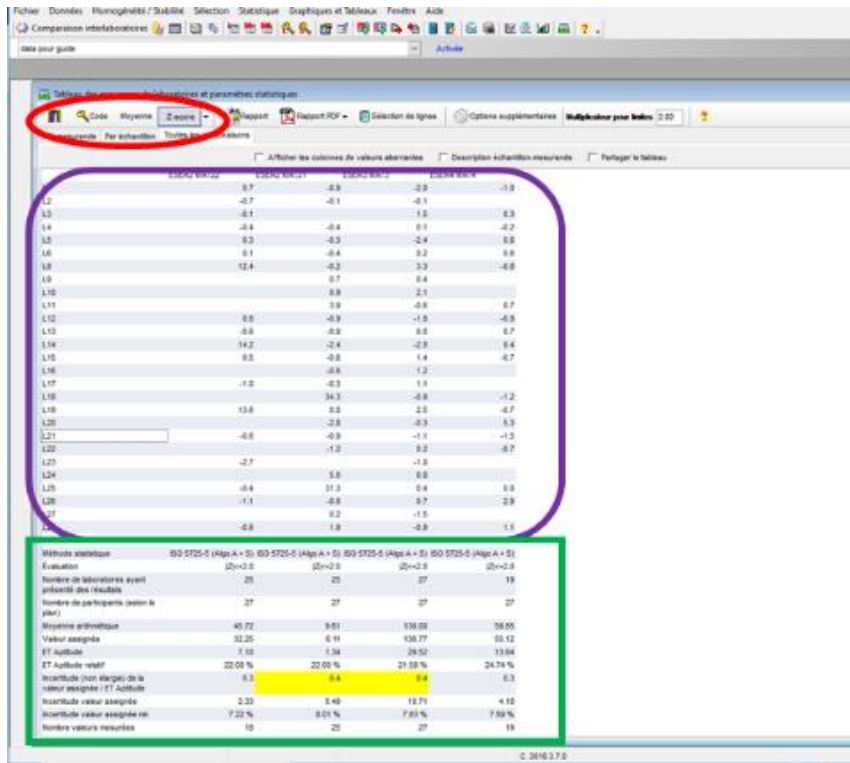


À droite, d'autres options sont également disponibles. Elles permettent de personnaliser le graphique (rectangle noir).

- **"Aperçu des z-scores"** : cette option permet de visualiser directement et uniquement les scores. Des options sont disponibles en haut du graphique pour personnaliser le graphique (ellipse rouge). La représentation peut se faire soit par participant, par ESEA, par analyte (mesurande) ou de façon globale (rectangle bleu) :



"Tableau des moyennes de participants et paramètres statistiques" : dans cette fenêtre sont fournis les résultats obtenus pour chaque participant sous forme de tableau (encadré rouge). Il est possible d'afficher à la fois les résultats obtenus et les z-scores correspondants en sélectionnant cet affichage au niveau de **"Code"**, **"Moyenne"**, **"z-score"** (rectangle violet). Dans cette fenêtre, est également fourni un tableau résumant l'ensemble des paramètres statistiques utilisés (rectangle vert). À noter que le rapport $u(x_{pt}) / \sigma_{pt}$ apparaît et est surligné en jaune quand il est ≥ 0.3 :



	ES02	ES03	ES04	ES05
L1	0,7	-0,5	-0,8	-1,5
L2	-0,7	-0,1	-0,1	
L3	-0,1		1,0	0,3
L4	-0,4	-0,4	0,1	-0,2
L5	0,3	-0,3	-0,4	0,0
L6	0,1	-0,4	0,2	0,0
L8	12,4	-0,2	3,3	-0,8
L9		0,7	0,4	
L10		0,9	2,1	0,7
L11		3,9	-0,4	0,7
L12	0,0	-0,9	-1,0	-0,5
L13	-0,0	-0,9	0,0	0,7
L14	14,2	-0,4	-0,9	0,4
L15	0,3	-0,0	1,4	-0,7
L16		-0,6	1,2	
L17	-1,0	-0,3	1,0	
L18		14,3	-0,8	-1,2
L19	13,0	0,0	2,0	-0,7
L20		-0,8	-0,3	0,3
L21	-0,0	-0,9	-1,1	-1,3
L22		-1,0	0,2	0,7
L23	-0,7		-1,0	
L24		5,0	0,0	
L25	-0,4	11,3	0,4	0,0
L26	-1,1	-0,0	0,7	2,9
L27	-0,0	1,0	-0,9	1,1

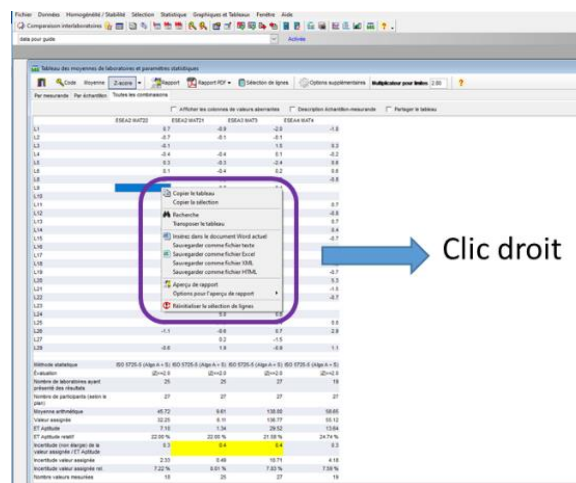
Méthode statistique	ISO 5725-5 (Alpa A = 5)	ISO 5725-5 (Alpa A = 5)	ISO 5725-5 (Alpa A = 5)	ISO 5725-5 (Alpa A = 5)
Évaluation	20<+2,0	20<+2,0	20<+2,0	20<+2,0
Nombre de laboratoires ayant participé aux résultats	25	25	27	19
Nombre de participants (selon le plan)	27	27	27	27
Moyenne arithmétique	45,72	0,01	130,00	30,00
Valeur assignée	32,25	0,19	126,77	33,12
ET Assignée	7,10	1,34	29,52	13,64
ET Assignée relatif	22,00 %	22,00 %	21,00 %	24,74 %
Incertitude (sur largeur de la valeur assignée) ET Assignée	0,3	0,4	0,4	0,3
Incertitude valeur assignée	2,33	0,40	10,71	4,10
Incertitude valeur assignée rel	7,22 %	0,01 %	7,85 %	7,50 %
Nombre valeurs mesurées	10	25	27	19


F.4 Édition de résultats

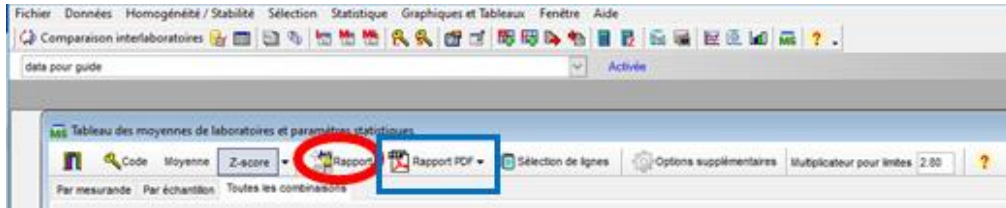
F.4.1 Aspects généraux

Plusieurs possibilités sont disponibles pour éditer les rapports.

Vous pouvez copier et coller vers Word tous les graphiques et tableaux, en effectuant un clic droit sur le tableau ou le graphique. Un sous-menu apparaît avec plusieurs choix (rectangle violet) :



Vous pouvez créer les rapports en cliquant sur le bouton "**Rapport**" (ellipse rouge) dans le bandeau des boutons. Ce bouton  est disponible dans la plupart des fenêtres et permet d'avoir un aperçu du rapport avant de le sauvegarder, il permet également de faire des modifications :



F.4.2 score z

Le rapport généré apparaît selon le modèle présenté avec le tableau des scores z et le tableau de synthèse :

data pour guide

Résumé de z-scores

Échantillon ESEA2

	ESEA2 MAT22	ESEA2 MAT21	ESEA2 MAT3	ESEA4 MAT4
U98	0,7	-0,9	-0,9	-1,0
L1	-0,7	-0,1	-0,1	-0,3
L3	-0,1	-0,4	1,5	0,3
L4	-0,4	-0,4	0,1	-0,3
L5	0,3	-0,3	2,4	0,6
L6	0,1	0,4	0,2	0,6
L8	12,4	-0,2	3,3	-0,8
L9	0,7	0,4		
L10	0,9	2,1		
L11	0,9	0,6		0,7
L12	0,6	-0,9	-1,8	-0,8
L13	-0,8	-0,9	0,0	0,7
L14	14,2	2,4	2,0	0,4
L15	0,5	-0,6	1,4	-0,7
L16	-0,8	-0,8	1,2	
L17	-1,0	-0,3	-1,1	-1,2
L18	13,6	34,3	2,5	0,7
L20	0,0	-2,8	-0,3	5,3
L21	-0,8	-0,9	-1,1	-1,5
L22	-1,2	-1,2	0,2	-0,7
L23	-2,7	5,0	-1,0	
L24	0,4	31,3	0,4	0,0
L25	-1,1	-0,6	0,7	3,9
L27		0,2	-1,5	

ANSES M Laurent 28/10/2021 Page 1

data pour guide

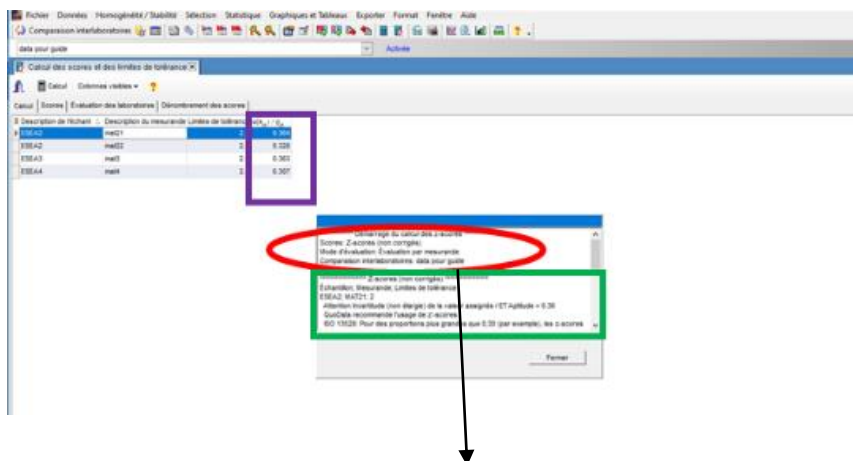
Échantillon ESEA2

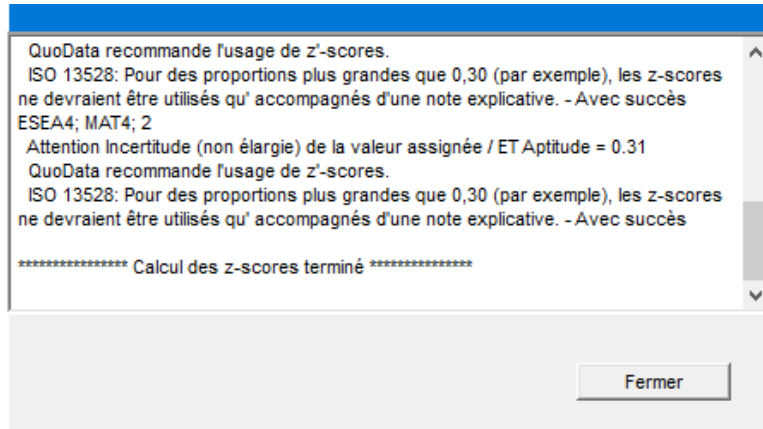
	ESEA2 MAT22	ESEA2 MAT21	ESEA2 MAT3	ESEA4 MAT4
L29	-0,8	1,3	-0,9	1,1
-	-	-	-	-
Méthode statistique	ISO 5725-5 (Algo A + 3)	ISO 5725-5 (Algo A + 3)	ISO 5725-5 (Algo A + 3)	ISO 5725-5 (Algo A + 3)
Évaluation	20+20	20+20	20+20	20+20
Nombre de laboratoires ayant participé aux Mesures	25	25	27	19
Nombre de participants (selon le plan)	27	27	27	27
Moyenne arithmétique	45,72	8,61	108,00	58,65
Valeur assignée	32,25	8,11	108,77	55,12
ET Assignée	7,10	1,34	28,92	13,66
ET Assignée relatif	22,00 %	27,00 %	31,58 %	24,61 %
Incertitude (non élargie) de la valeur assignée / ET Assignée	0,3	0,4	0,4	0,3
Incertitude valeur assignée	2,33	0,49	10,71	4,18
Incertitude valeur assignée rel.	7,22 %	8,01 %	7,83 %	7,58 %
Nombre valeurs mesurées	18	25	27	19

Vous pouvez créer directement les rapports en cliquant sur "Rapport PDF" dans le bandeau des boutons. Ce bouton (rectangle bleu) est disponible dans la plupart des fenêtres. L'activation du bouton génère un tableau ou un graphique en format PDF qui est sauvegardé.

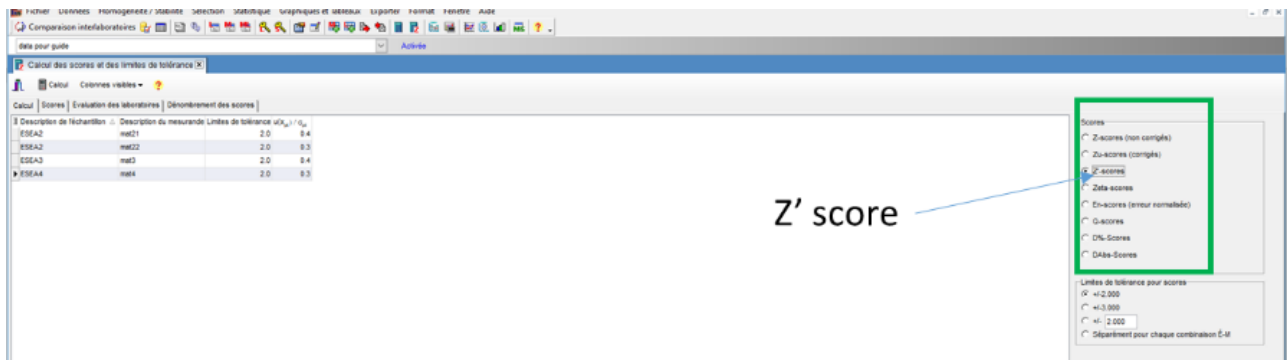
F.4.3 score z'


Dans le cas où l'incertitude de la valeur assignée ne vérifie pas la condition $u(x_{pt} < 0.3\sigma_{pt})$, un message d'alerte apparaît (rectangle vert) avec une recommandation d'utiliser le score z' :

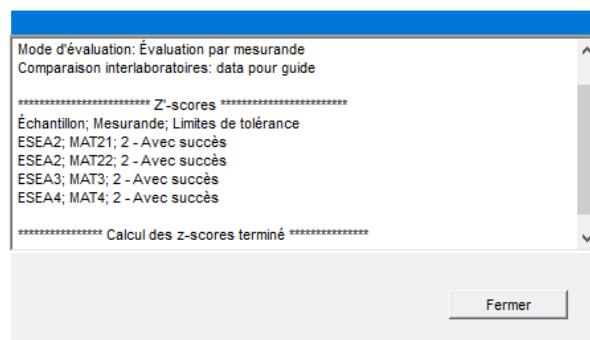




Relancer les calculs en sélectionnant le score z' (rectangle vert) :



Une fois sélectionné, les calculs sont relancés par le bouton  et le message suivant apparaît si les l'utilisation du score z' permet de vérifier la condition :



Dans cet exemple, l'utilisation du score z' permet de vérifier la condition.

Une fois les calculs du score z' effectués, la visualisation des graphiques, des tableaux et l'édition des rapports est la même que pour le score z (F.4.2).

F.4.4 Exemple de jeu de données

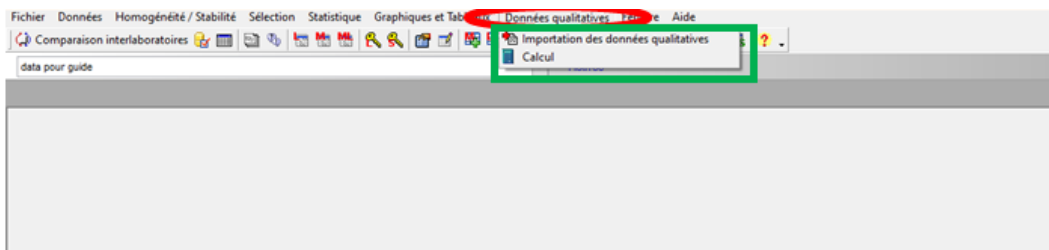
Les données fournies sont les données qui ont permis de réaliser l'ensemble des calculs décrits :

code lab	ESEA	Mesurande	Valeur mesurée 1	Unité	Code préparation échantillon	code lab	ESEA	Mesurande	Valeur mesurée 1	Unité	Code préparation échantillon
L1	ESEA2	MAT21	4.9	µg/kg	11	L1	ESEA2	MAT22	37.2	µg/kg	11
L2	ESEA2	MAT21	5.93	µg/kg	185	L2	ESEA2	MAT22	27.4	µg/kg	185
L4	ESEA2	MAT21	5.61	µg/kg	5	L3	ESEA2	MAT22	31.2	µg/kg	79
L5	ESEA2	MAT21	5.72	µg/kg	77	L4	ESEA2	MAT22	29.5	µg/kg	5
L7	ESEA2	MAT21	5.62	µg/kg	215	L5	ESEA2	MAT22	34.1	µg/kg	77
L8	ESEA2	MAT21	5.83	µg/kg	76	L7	ESEA2	MAT22	33	µg/kg	215
L9	ESEA2	MAT21	7.04	µg/kg	225	L8	ESEA2	MAT22	120	µg/kg	76
L10	ESEA2	MAT21	7.32	µg/kg	129	L12	ESEA2	MAT22	36.7	µg/kg	107
L11	ESEA2	MAT21	11.4	µg/kg	60	L13	ESEA2	MAT22	27.9	µg/kg	183
L12	ESEA2	MAT21	4.93	µg/kg	107	L14	ESEA2	MAT22	133	µg/kg	101
L13	ESEA2	MAT21	4.96	µg/kg	183	L15	ESEA2	MAT22	35.7	µg/kg	53
L14	ESEA2	MAT21	2.93	µg/kg	101	L17	ESEA2	MAT22	25.2	µg/kg	149
L15	ESEA2	MAT21	5.34	µg/kg	53	L19	ESEA2	MAT22	129	µg/kg	93
L16	ESEA2	MAT21	5.36	µg/kg	193	L21	ESEA2	MAT22	27.7	µg/kg	132
L17	ESEA2	MAT21	5.74	µg/kg	149	L23	ESEA2	MAT22	13.1	µg/kg	74
L18	ESEA2	MAT21	52.2	µg/kg	224	L24	ESEA2	MAT22	29.2	µg/kg	103
L19	ESEA2	MAT21	6.17	µg/kg	93	L26	ESEA2	MAT22	24.7	µg/kg	19
L20	ESEA2	MAT21	2.31	µg/kg	58	L27	ESEA2	MAT22	28.3	µg/kg	56
L21	ESEA2	MAT21	4.96	µg/kg	132						
L22	ESEA2	MAT21	4.5	µg/kg	41						
L24	ESEA2	MAT21	12.9	µg/kg	122						
L25	ESEA2	MAT21	48.2	µg/kg	103						
L25	ESEA2	MAT21	5.26	µg/kg	19						
L27	ESEA2	MAT21	6.44	µg/kg	27						
L29	ESEA2	MAT21	8.65	µg/kg	56						
code lab	ESEA	Mesurande	Valeur mesurée 1	Unité	Code préparation échantillon	code lab	ESEA	Mesurande	Valeur mesurée 1	Unité	Code préparation échantillon
L1	ESEA3	MAT3	77.9	µg/kg	177	L29	ESEA3	MAT3	109	µg/kg	6
L2	ESEA3	MAT3	134	µg/kg	157	L1	ESEA4	MAT4	41.7	µg/kg	106
L3	ESEA3	MAT3	180	µg/kg	160	L3	ESEA4	MAT4	58.9	µg/kg	118
L4	ESEA3	MAT3	140.1	µg/kg	229	L4	ESEA4	MAT4	52.3	µg/kg	138
L5	ESEA3	MAT3	65.3	µg/kg	45	L5	ESEA4	MAT4	63.5	µg/kg	34
L6	ESEA3	MAT3	141.8	µg/kg	28	L7	ESEA4	MAT4	62.8	µg/kg	105
L8	ESEA3	MAT3	233	µg/kg	86	L8	ESEA4	MAT4	43.9	µg/kg	208
L9	ESEA3	MAT3	150	µg/kg	36	L11	ESEA4	MAT4	65.2	µg/kg	212
L10	ESEA3	MAT3	198.2	µg/kg	167	L12	ESEA4	MAT4	44.6	µg/kg	233
L11	ESEA3	MAT3	120	µg/kg	35	L13	ESEA4	MAT4	65.1	µg/kg	172
L12	ESEA3	MAT3	82.5	µg/kg	10	L14	ESEA4	MAT4	60.2	µg/kg	69
L13	ESEA3	MAT3	136	µg/kg	164	L15	ESEA4	MAT4	45.4	µg/kg	65
L14	ESEA3	MAT3	78.7	µg/kg	113	L18	ESEA4	MAT4	39.1	µg/kg	163
L15	ESEA3	MAT3	179	µg/kg	156	L19	ESEA4	MAT4	45.4	µg/kg	109
L16	ESEA3	MAT3	171	µg/kg	15	L20	ESEA4	MAT4	127	µg/kg	180
L17	ESEA3	MAT3	170	µg/kg	8	L21	ESEA4	MAT4	34.3	µg/kg	20
L18	ESEA3	MAT3	109	µg/kg	91	L22	ESEA4	MAT4	45	µg/kg	48
L19	ESEA3	MAT3	212	µg/kg	145	L25	ESEA4	MAT4	55.2	µg/kg	213
L20	ESEA3	MAT3	126.6	µg/kg	54	L26	ESEA4	MAT4	94	µg/kg	50
L21	ESEA3	MAT3	103	µg/kg	174	L29	ESEA4	MAT4	70.7	µg/kg	192
L22	ESEA3	MAT3	143	µg/kg	85						
L23	ESEA3	MAT3	106	µg/kg	9						
L24	ESEA3	MAT3	160	µg/kg	139						
L25	ESEA3	MAT3	150	µg/kg	168						
L26	ESEA3	MAT3	156	µg/kg	166						
L27	ESEA3	MAT3	93.8	µg/kg	218						

Annexe G : Exemples de calcul du L-score (Prolab)

G.1 Aspects généraux

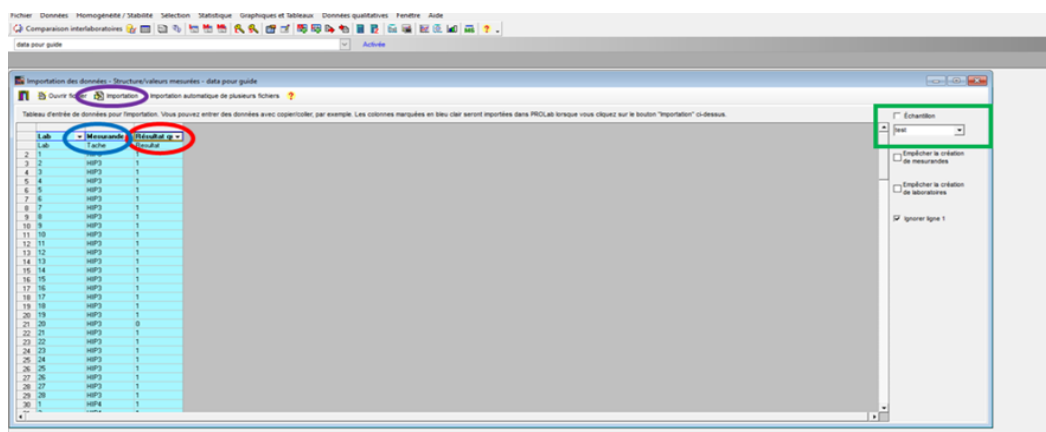
Pour calculer le L-score avec le logiciel Prolab, sélectionner le module **"Données qualitatives"** (ellipse rouge) qui permet d'accéder aux sous-modules **"Importation des données"** et **"Calcul"** (rectangle vert) :



Vous pouvez importer les données via un fichier Excel ou par un copier/coller à partir du document Excel **F/EA/FE/370**.

Il est important de définir les points suivants :

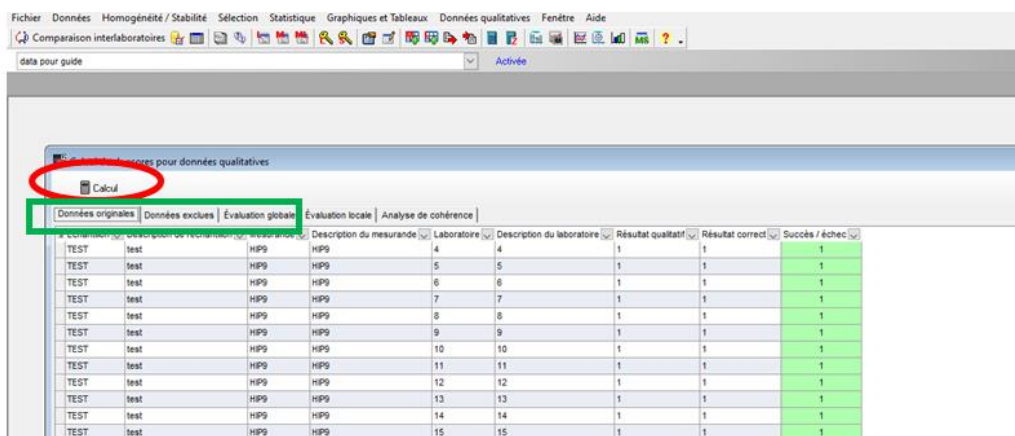
- Si un seul type d'échantillon (matrice) est utilisé dans l'EILA, par exemple du lait, sélectionner la case **"Echantillon"** et définir l'échantillon. Il n'est pas nécessaire d'avoir une colonne avec l'échantillon (rectangle vert).
- Définir la « tâche », c'est-à-dire l'analyse qui est réalisée dans l'EILA (par exemple, présence ou absence (ou détection/non détection) d'une molécule ou d'un microorganisme, identification d'une bactérie), dans la colonne **"Mesurande"** (ellipse bleue).
- La valeur mesurée, qui est binaire, est traduite par des "0" et des "1" : "0" correspond à un résultat incorrect par rapport à ce qui est attendu, "1" à un résultat correct attendu. De ce fait, sélectionner l'intitulé de la colonne dans le menu déroulant et le mettre dans **"Résultat qualitatif"** (ellipse rouge) :



Une fois ces points définis, déclencher l'importation par le bouton "**Importation**" (ellipse violette).

G.2 Calculs

L'étape suivante est la réalisation de calculs. L'activation du sous-module "**Calcul**" entraîne l'ouverture de la fenêtre suivante :

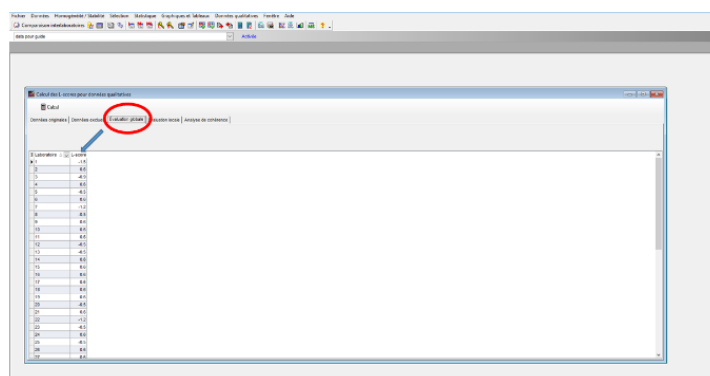


Parmi les cinq onglets qui apparaissent, 3 sont plus particulièrement intéressants (rectangle vert) :

- "**Données originales**" ;
- "**Évaluation globale**" ;
- "**Évaluation locale**".

L'onglet "**Données originales**" permet de visualiser les résultats importés. Une fois que l'importation a été vérifiée, lancer les calculs via le bouton "**Calcul**" (ellipse rouge).

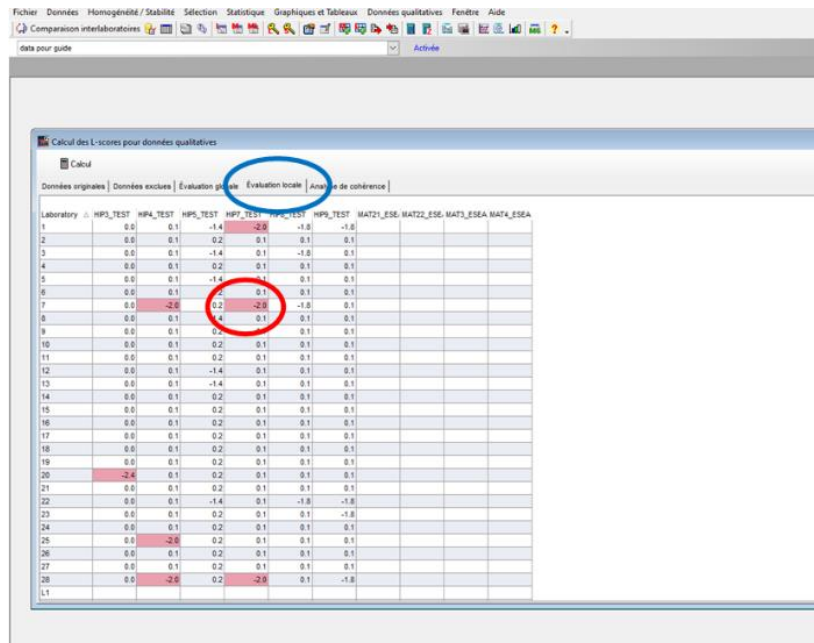
Une fois les calculs effectués, la fenêtre "**Évaluation globale**" permet d'afficher les L-scores globaux des participants. Ce score global prend en compte toutes les tâches. Pour pouvoir le calculer, il faut au moins 6 tâches et 10 participants :



Pour mémoire, interpréter les L-scores de la façon suivante :

- $L < -2$: la compétence du participant est inférieure à la compétence moyenne de l'ensemble des participants ;

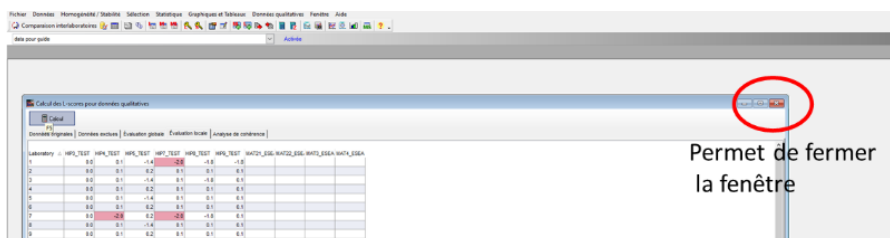
- $|L| \leq 2$: la compétence du participant est dans la compétence moyenne de l'ensemble des participants ;
- $L > 2$: la compétence du participant est supérieure à la compétence moyenne de l'ensemble des participants.
- Un score $L < -2$ sera surligné en rouge (ellipse rouge).
- Pour visualiser les L-scores par participant et par tâche, sélectionner la fenêtre "Évaluation locale" :



Laboratoire	HPI4_TEST	HPI4_TEST	HPI4_TEST	HPI4_TEST	HPI4_TEST	HPI4_TEST	MAT21_ESE	MAT22_ESE	MAT3_ESEA	MAT4_ESEA
1	0.0	0.1	-1.4	-2.0	-1.8	-1.8				
2	0.0	0.1	0.2	0.1	0.1	0.1				
3	0.0	0.1	-1.4	0.1	-1.8	0.1				
4	0.0	0.1	0.2	0.1	0.1	0.1				
5	0.0	0.1	-1.4	-1	0.1	0.1				
6	0.0	0.1	0.2	0.1	0.1	0.1				
7	0.0	-2.0	0.2	-2.0	-1.8	0.1				
8	0.0	0.1	0.2	0.1	0.1	0.1				
9	0.0	0.1	0.2	0.1	0.1	0.1				
10	0.0	0.1	0.2	0.1	0.1	0.1				
11	0.0	0.1	0.2	0.1	0.1	0.1				
12	0.0	0.1	-1.4	0.1	0.1	0.1				
13	0.0	0.1	-1.4	0.1	0.1	0.1				
14	0.0	0.1	0.2	0.1	0.1	0.1				
15	0.0	0.1	0.2	0.1	0.1	0.1				
16	0.0	0.1	0.2	0.1	0.1	0.1				
17	0.0	0.1	0.2	0.1	0.1	0.1				
18	0.0	0.1	0.2	0.1	0.1	0.1				
19	0.0	0.1	0.2	0.1	0.1	0.1				
20	-2.0	0.1	0.2	0.1	0.1	0.1				
21	0.0	0.1	0.2	0.1	0.1	0.1				
22	0.0	0.1	-1.4	0.1	-1.8	-1.8				
23	0.0	0.1	0.2	0.1	0.1	-1.8				
24	0.0	0.1	0.2	0.1	0.1	0.1				
25	0.0	-2.0	0.2	0.1	0.1	0.1				
26	0.0	0.1	0.2	0.1	0.1	0.1				
27	0.0	0.1	0.2	0.1	0.1	0.1				
28	0.0	-2.0	0.2	-2.0	0.1	-1.8				
L1										

L'interprétation des L-scores individuels est identique à l'interprétation des L-scores pour l'évaluation globale.

Precision : pour fermer les fenêtres dans cette partie du programme, activer la croix dans le coin droit en haut :



Permet de fermer la fenêtre

Pour visualiser/éditer un tableau, faire un clic droit avec la souris sur le tableau pour faire apparaître un sous-menu. Sélectionner  "Aperçu de rapport".

Des graphiques ne sont pas générés, cependant les scores peuvent être copiés ou exportés sous Word ou Excel, et être représentés sous format graphique.

Annexe H : Principe et calcul du S-score

H.1 Principe

Le S-score est construit en combinant les notes attribuées aux résultats corrects et incorrects. La partie décimale (DP) du score correspond au taux global de résultats incorrects, tandis que la partie entière (IP) indique la performance du laboratoire avec une interprétation similaire au score z utilisé dans les EILA quantitatifs, excepté pour la valeur IP=2 (satisfaisante pour le z-score). En notant g_1 le nombre de ESEA1, g_2 le nombre de ESEA2, g_3 le nombre de ESEA3 et g_4 le nombre de ESEA4 pour lesquels la valeur assignée (VA) est définie, et g leur somme :

$$g = g_1 + g_2 + g_3 + g_4$$

Pour effectuer le calcul du S-score, le résultat obtenu par le laboratoire i pour l'ESEA j est converti en une note X_{ij} où $1 \leq i \leq p$ avec p le nombre de laboratoires et $1 \leq j \leq g$ avec g le nombre d'ESEA

$X_{ij} = 1$ si le résultat est différent de la VA (résultat incorrect)

$X_{ij} = 0$ si le résultat correspond à la VA (résultat correct)

La performance des participants est évaluée à partir :

- des résultats individuels (corrects ou incorrects) pour les ESEA1, ESEA3, et ESEA4,
- d'un intervalle acceptable de résultats corrects pour le groupe ESEA2.

L'intervalle acceptable de résultats corrects pour les ESEA2 est déterminé a posteriori à l'aide de la loi binomiale (c'est-à-dire la densité de probabilité non cumulative) en fonction de la proportion observée :

$$\hat{\theta} = \frac{m}{M}$$

Avec :

- m le nombre d'occurrences de la modalité binaire majoritaire observée ($\geq 50\%$)
- M le nombre de laboratoires p multiplié par le nombre g_2 d'ESEA2 :

$$M = p \times g_2$$

Le nombre de résultats corrects n_0 est défini, pour le groupe ESEA2, par le nombre de notes égales à 0 :

$$n_0 = \sum_{j=1}^{g_2} \mathbb{1}_{\{X_{ij}=0\}}$$

où $\mathbb{1}$ représente la fonction indicatrice pour la condition $X_{ij}=0$.

Ainsi, un intervalle acceptable de résultats corrects $R = [Y_{min}, Y_{max}]$ est défini au risque $\alpha = 0,05$ et $\alpha = 0,01$ dans les tableaux J.1 et J.2, respectivement.

Les règles de décision qui déterminent Y_{min} et Y_{max} sont les suivantes :

$$Y_{min} = \min\{y < g_2 \in \mathbb{N} \text{ such as } P(Y = y) \geq \alpha\}$$

$$Y_{max} = \max\{y < g_2 \in \mathbb{N} \text{ such as } P(Y = y) \geq \alpha\}$$

Où Y est une variable aléatoire, $Y \sim \text{Bin}(g_2; \hat{\theta})$. Sa valeur est donnée dans les tableaux J.1 et J.2 pour chaque proportion observée $\hat{\theta}$, allant de 50 % à 100 % avec un pas de 0,1 %.

La détermination de la valeur assignée pour les ESEA4 est effectuée à partir d'une distribution bêta unilatérale qui permet de vérifier si le consensus des laboratoires participants (LC) est significativement supérieur à 50 % : voir Baudry 2024 et tableau J.3.

H.2 Calcul du S-score

Pour chaque laboratoire et chaque ESEA, deux paramètres, l'un appelé RR (Required Results) correspondant aux résultats exigés et l'autre AR (Average Results) correspondant à la moyenne des notes attribuées aux résultats corrects et incorrects, sont calculés.

Paramètre	ESEA	Formule
$RR1_i$	ESEA1	$RR1_i = \sum_{j=1}^{g_1} X_{ij}$
$RR2_i$	ESEA2	<p>Selon les équations Y_{min} et Y_{max} avec $\alpha = 0.05$,</p> <p>Si $Y_{min} \leq n_0 = \sum_{j=1}^{g_2} \mathbb{1}_{\{X_{ij}=0\}} \leq Y_{max} \rightarrow RR2_i = 0$</p> <p>Selon les équations Y_{min} et Y_{max} avec $\alpha = 0.01$,</p> <p>Sinon si $Y_{min} \leq n_0 = \sum_{j=1}^{g_2} \mathbb{1}_{\{X_{ij}=0\}} \leq Y_{max} \rightarrow RR2_i = -1$</p> <p>Sinon $RR2_i = -2$</p>
$AR1_i$	ESEA1, ESEA2, ESEA3, ESEA4	$AR1_i = \sum_{j=1}^g X_{ij} / g$
$AR2_i$	ESEA3, ESEA4	$AR2_i = \left[\sum_{j=1}^{g_3} X_{ij} + \sum_{j=1}^{g_4} X_{ij} \right] / (g_3 + g_4)$

Le calcul final du S-score d'un laboratoire dépend des valeurs des paramètres et des différentes conditions énumérées ci-dessous. Un panel doit inclure au moins une ESEA1 ou un groupe ESEA2.

Conditions	S-score
C1: $RR1_i = 0$ et $RR2_i = 0$ et $AR2_i = 0$	$1 + AR1_i$
C2: $RR1_i = 0$ et $RR2_i = -1$ et $AR2_i = 0$ ou C3: $RR1_i = 0$ et $RR2_i \neq -2$ et $AR2_i \neq 0$	$2 + AR1_i$
C4: $RR1_i \geq 1$ ou $RR2_i = -2$	$3 + AR1_i$

Les points suivants doivent être pris en compte pour calculer le S-score :

- a) ESEA2 non répliquées n'ayant pas la même modalité binaire majoritaire (par exemple, positive) : les ESEA2 ne sont pas prises en compte dans le calcul du S-score.
- b) Pas de VA majoritaire définie dans le groupe ESEA2 car le LC est égal à 50 % : une VA doit être définie arbitrairement (par exemple positive) pour attribuer une note à chaque résultat et calculer le score participant. Dans ce cas, les résultats incorrects ESEA2 ne sont pas comptabilisés dans la partie décimale du S-score.
- c) Condition C3 avec $AR1_i$ égal à 1 (100% de résultats incorrects) : la valeur du score de 3,0 est remplacée par 2,99. Ce cas peut se produire uniquement si, dans un panel sans ESEA1, un LC inférieur à 53,6% est observé dans un groupe de 6 ESEA2 (voir tableau J.2).
- d) Pas de VA définie pour une ESEA4 car consensus minimum acceptable n'est pas atteint : ESEA non prise en compte dans le calcul du S-score.

Annexe I : Exemples de calcul du S-score pour les EILA qualitatifs

I.1 Exemple n°1 : panel composé des 4 types d'ESEA

Le premier exemple est obtenu à partir de données simulées provenant d'un EILA concernant la capacité des laboratoires à détecter la présence ou l'absence de médicaments vétérinaires interdits dans les denrées alimentaires d'origine animale (par exemple, le lait, les œufs, la viande). La nature de l'analyte est anonymisée et les ESEA sont nommées de A à M. Les ESEA A et B sont supérieures à la limite de détection (LD) et 11 ESEA (C à H) sont proches de la LD.

Le tableau I.1 présente un exemple de panel comprenant tous les différents types d'ESEA.

Tableau I.1 : Exemple d'un panel composé de 13 ESEA contenant des analytes interdits en médecine vétérinaire. Il y a six ESEA2 répliquées (de C à H) avec un consensus de laboratoire (LC) de 67,8 % (+) ; la VA ESEA4 (L) n'est pas définie (ND) car le LC de 66,7 % (+) est inférieur à 80 % (voir tableau J.3). Les résultats incorrects, c'est-à-dire différent de la VA, sont affichés sur fond orange. Les S-scores insatisfaisants sont en gras.

ESEA VA Lab	ESEA1		ESEA2						ESEA3			ESEA4		S-score
	A (-)	B (+)	C	D	E (+)	F	G	H	I (-)	J (+)	K (+)	L ND	M (-)	
01	-	+	-	+	+	-	-	+	-	+	+	+	-	1,25
02	-	+	-	-	-	-	-	-	-	-	+	-	+	3,67
03	-	+	+	+	+	+	+	+	-	-	-	+	-	2,17
04	-	+	+	+	+	+	-	-	-	+	+	-	-	1,17
05	-	+	+	+	+	-	+	+	-	+	+	+	-	1,08
06	-	+	-	-	-	-	-	+	-	+	+	-	-	2,42
07	-	+	+	-	+	+	+	-	+	-	-	-	-	2,42
08	-	-	-	+	+	+	+	-	-	+	+	+	-	3,25
09	-	+	+	-	+	-	+	-	+	-	-	-	+	2,58
10	-	+	+	+	+	+	+	+	-	+	+	+	-	1,00
11	-	+	+	-	+	+	+	+	+	+	+	+	-	2,17
12	-	+	+	+	+	+	+	-	-	+	+	+	-	1,08
13	-	+	+	+	+	+	+	+	-	+	+	+	-	1,00
14	+	-	+	-	+	-	+	-	+	-	-	+	-	3,67
15	-	+	+	+	+	+	+	+	-	+	+	+	-	1,00

L'organisateur d'EILA utilise deux ESEA1 (A et B) pour lesquels des résultats corrects sont exigés. Deux laboratoires (08 et 14) avec un S-score supérieur ou égal à 3 (3,25 et 3,67) ont une performance insatisfaisante car ils ont obtenu au moins un résultat incorrect sur ces deux ESEA.

Un groupe de six ESEA2 répétés (de C à H) est ajouté dans le panel. Ce sont des échantillons proches de la LD qui représentent des tâches difficiles exigées. La VA est positive (+) et le LC observé est de 67,8 %. Le nombre de résultats positifs requis doit être compris entre 1 et 6 ; intervalle acceptable de résultats positifs indiqué dans le tableau J.2 ($\alpha = 0,01$). Le laboratoire 02 n'ayant obtenu aucun résultat positif a une performance insatisfaisante (S-score = 3,67). Le laboratoire 06, avec un seul résultat positif sur six, a une performance discutable (S-score = 2,42), car le nombre de résultats positifs doit être compris entre 2 et 6 selon le tableau J.1 ($\alpha = 0,05$).

L'organisateur complète son panel avec trois ESEA3 (I à K) et deux ESEA4 (L à M) pour lesquels des résultats corrects sont attendus mais non requis. Les ESEA (L et M) proches de la LD ont été déclarées en type 4 : la VA est définie par la modalité majoritaire observée si le LC est statistiquement significatif. Pour l'ESEA4 (L), le LC observé n'est que de 66,7 % (10/15) et inférieur au LC minimal (12/15) indiqué dans le tableau J.3. Par conséquent, la VA ne peut pas être définie et les résultats ne sont pas pris en compte dans le calcul du S-score. Quatre laboratoires (laboratoires 03, 07, 09 et 11) ayant obtenu au moins un résultat incorrect sur les trois ESEA3 et l'ESEA4 (M) ont une performance discutable avec une valeur de S-score comprise entre $2 \leq S\text{-score} < 3$.

Les résultats obtenus avec l'outil de calcul Excel QualiScore sont présentés ci-dessous. A noter que la codification binaire utilisée est la suivante : N pour (-) et P pour (+).

Code EILA : CodeEILA Evaluation de la performance EILA : S-scores participants, ESEA et EILA



Valeur assignée			N	P	P						N	P	P	NA	N
Type ESEA			ESEA1	ESEA1	ESEA2	ESEA2	ESEA2	ESEA2	ESEA2	ESEA2	ESEA3	ESEA3	ESEA3	ESEA4	ESEA4
Consensus participants														66,7%	86,7%
Consensus acceptable														80,0%	80,0%
Codes laboratoires	S-score	Conclusion	A	B	C	D	E	F	G	H	I	J	K	L	M
L01	1,25	Satisfaisant	N	P	N	P	P	N	N	P	N	P	P	P	N
L02	3,67	Insatisfaisant	N	P	N	N	N	N	N	N	N	N	P	N	P
L03	2,17	Discutable	N	P	P	P	P	P	P	P	N	N	N	P	N
L04	1,17	Satisfaisant	N	P	P	P	P	P	N	N	N	P	P	N	N
L05	1,08	Satisfaisant	N	P	P	P	P	N	P	P	N	P	P	P	N
L06	2,42	Discutable	N	P	N	N	N	N	N	P	N	P	P	N	N
L07	2,42	Discutable	N	P	P	N	P	P	P	N	P	N	N	N	N
L08	3,25	Insatisfaisant	N	N	N	P	P	P	P	N	N	P	P	P	N
L09	2,58	Discutable	N	P	P	N	P	N	P	N	P	N	N	N	P
L10	1,00	Satisfaisant	N	P	P	P	P	P	P	P	N	P	P	P	N
L11	2,17	Discutable	N	P	N	P	P	P	P	P	P	P	P	P	N
L12	1,08	Satisfaisant	N	P	P	P	P	P	P	N	N	P	P	P	N
L13	1,00	Satisfaisant	N	P	P	P	P	P	P	P	N	P	P	P	N
L14	3,67	Insatisfaisant	P	N	P	N	P	N	P	N	P	N	N	P	N
L15	1,00	Satisfaisant	N	P	P	P	P	P	P	P	N	P	P	P	N
S _{ESEA}			3,07	3,13	1,27	1,40	1,13	1,40	1,27	1,47	2,27	2,33	2,27		2,13
S _{EILA}			3,26												

Les scores S_{ESEA} et S_{EILA}, respectivement pour chaque ESEA et sur l'ensemble des données, sont calculés selon le même principe que le S-score participant et constituent des informations complémentaires.

I.2 Exemple n°2 : utilisation des ESEA2 non répliquées

Le S-score a été appliqué aux données fournies par Uhlig *et al.* Le panel, envoyé à 28 laboratoires, était constitué de 9 ESEA, chacune contaminée avec des microorganismes pathogènes appelés HIP (« Highly Infectious Pathogen ») 1 à 9.

Le résultat individuel "+" indique un résultat correct (c'est-à-dire détecté), tandis que "-" indique un résultat incorrect (non détecté).

Pour comparer le S-score avec les deux autres scores, (L-score et a-score), nous avons choisi de considérer les 9 microorganismes pathogènes (HIP1 à HIP9) comme un seul groupe ESEA2. Il s'agit d'échantillons différents donc non répliqués, Il a été vérifié qu'ils ont tous la même valeur assignée (positive), avec un consensus pour chaque ESEA2 supérieur à 50%. C'est la situation que l'on peut observer dans l'exemple du tableau I.2. Avec un LC de 90,1 % de résultats positifs (+), 6 résultats positifs sur 9 sont exigés au risque $\alpha=0,01$ (sinon performance insatisfaisante) selon le tableau J.2, et 7 résultats positifs sur 9 sont attendus (sinon performance discutable) au risque $\alpha=0,05$ selon le tableau J.1. Ainsi, le laboratoire 01, avec cinq résultats positifs, est le seul laboratoire dont la performance avec le S-score est insatisfaisante, tandis que les laboratoires 07, 22 et 28, avec seulement six résultats positifs, ont une performance discutable.

Tableau I.2 : Comparaison L-score [3], a-score [2] et S-score associés à un EILA constitué d'un groupe de neuf ESEA2 non répliquées. Les résultats incorrects (signe moins "-") sont affichés sur fond orange; le a-score est fourni sans tenir compte de HIP5 [2] ; les scores correspondant à des performances insatisfaisantes en caractères gras.

Lab	HIP1	HIP2	HIP3	HIP4	HIP5	HIP6	HIP7	HIP8	HIP9	L-score	a-score	S-score
01	+	+	+	+	-	+	-	-	-	-2,2	74,8	3,44
02	+	+	+	+	+	+	+	+	+	1,1	0	1,00
03	+	+	+	+	-	+	+	-	+	-1,3	23,6	1,22
04	+	+	+	+	+	+	+	+	+	1,1	0	1,00
05	+	+	+	+	-	+	+	+	+	-0,6	0	1,11
06	+	+	+	+	+	+	+	+	+	1,1	0	1,00
07	+	+	+	-	+	+	-	-	+	-1,8	78,9	2,33
08	+	+	+	+	-	+	+	+	+	-0,6	0	1,11
09	+	+	+	+	+	+	+	+	+	1,1	0	1,00
10	+	+	+	+	+	+	+	+	+	1,1	0	1,00
11	+	+	+	+	+	+	+	+	+	1,1	0	1,00
12	+	+	+	+	-	+	+	+	+	-0,6	0	1,11
13	+	+	+	+	-	+	+	+	+	-0,6	0	1,11
14	+	+	+	+	+	+	+	+	+	1,1	0	1,00
15	+	+	+	+	-	+	+	+	+	-0,6	0	1,11

Lab	HIP1	HIP2	HIP3	HIP4	HIP5	HIP6	HIP7	HIP8	HIP9	L-score	a-score	S-score
16	+	+	+	+	-	+	+	+	+	-0,6	0	1,11
17	+	+	+	+	+	+	+	+	+	1,1	0	1,00
18	+	+	+	+	+	+	+	+	+	1,1	0	1,00
19	+	+	+	+	+	+	+	+	+	1,1	0	1,00
20	+	+	-	+	+	+	+	+	+	-0,6	38,5	1,11
21	+	+	+	+	+	+	+	+	+	1,1	0	1,00
22	+	+	+	+	-	+	+	-	-	-1,8	47,2	2,33
23	+	+	+	+	+	+	+	+	-	-0,6	23,6	1,11
24	+	+	+	+	+	+	+	+	+	1,1	0	1,00
25	+	-	+	-	+	+	+	+	+	-1,3	66,2	1,22
26	+	+	+	+	+	+	+	+	+	1,1	0	1,00
27	+	+	+	+	+	+	+	+	+	1,1	0	1,00
28	+	+	+	-	+	+	-	+	-	-1,8	78,9	2,33

Tableau I.3 : Interprétations associées aux L-score [3], a-score [2] et S-score

Lab	L-score		a-score		S-score	
	Valeur	Interpretation	Valeur	Interpretation	Valeur	Interpretation
01	-2,2	Insatisfaisante	74,8	Insatisfaisante	3,44	Insatisfaisante
02	1,1	Satisfaisante	0	Satisfaisante	1,00	Satisfaisante
03	-1,3	Satisfaisante	23,6	Insatisfaisante	1,22	Satisfaisante
04	1,1	Satisfaisante	0	Satisfaisante	1,00	Satisfaisante
05	-0,6	Satisfaisante	0	Satisfaisante	1,11	Satisfaisante
06	1,1	Satisfaisante	0	Satisfaisante	1,00	Satisfaisante
07	-1,8	Satisfaisante	78,9	Insatisfaisante	2,33	Discutable
08	-0,6	Satisfaisante	0	Satisfaisante	1,11	Satisfaisante
09	1,1	Satisfaisante	0	Satisfaisante	1,00	Satisfaisante
10	1,1	Satisfaisante	0	Satisfaisante	1,00	Satisfaisante
11	1,1	Satisfaisante	0	Satisfaisante	1,00	Satisfaisante
12	-0,6	Satisfaisante	0	Satisfaisante	1,11	Satisfaisante
13	-0,6	Satisfaisante	0	Satisfaisante	1,11	Satisfaisante
14	1,1	Satisfaisante	0	Satisfaisante	1,00	Satisfaisante
15	-0,6	Satisfaisante	0	Satisfaisante	1,11	Satisfaisante

Lab	L-score		a-score		S-score	
	Valeur	Interpretation	Valeur	Interpretation	Valeur	Interpretation
16	-0,6	Satisfaisante	0	Satisfaisante	1,11	Satisfaisante
17	1,1	Satisfaisante	0	Satisfaisante	1,00	Satisfaisante
18	1,1	Satisfaisante	0	Satisfaisante	1,00	Satisfaisante
19	1,1	Satisfaisante	0	Satisfaisante	1,00	Satisfaisante
20	-0,6	Satisfaisante	38,5	Insatisfaisante	1,11	Satisfaisante
21	1,1	Satisfaisante	0	Satisfaisante	1,00	Satisfaisante
22	-1,8	Satisfaisante	47,2	Insatisfaisante	2,33	Discutable
23	-0,6	Satisfaisante	23,6	Insatisfaisante	1,11	Satisfaisante
24	1,1	Satisfaisante	0	Satisfaisante	1,00	Satisfaisante
25	-1,3	Satisfaisante	66,2	Insatisfaisante	1,22	Satisfaisante
26	1,1	Satisfaisante	0	Satisfaisante	1,00	Satisfaisante
27	1,1	Satisfaisante	0	Satisfaisante	1,00	Satisfaisante
28	-1,8	Satisfaisante	78,9	Insatisfaisante	2,33	Discutable

Annexe J : Tableaux pour le calcul du S-score

Les tableaux J.1 et J.2 permettent de vérifier si le nombre de résultats corrects observé pour un laboratoire est dans l'intervalle acceptable de résultats corrects défini au risque de 5% dans la table J.1 et au risque de 1% dans le tableau J.2).

Tableau J.1 : Intervalle acceptable de résultats corrects $R = [Y_{\min}, Y_{\max}]$ basée sur la densité de probabilité binomiale ($\geq 5\%$), et dépendant du nombre g_2 de ESEA2 et de la proportion observée de la modalité binaire majoritaire $\hat{\theta}$ avec la borne inférieure incluse et la borne supérieure exclue. Par exemple, si $\hat{\theta} (\%) = 57,1$; $R = [2 \text{ à } 5]$.

$g_2 = 6$		$g_2 = 7$		$g_2 = 8$		$g_2 = 9$		$g_2 = 10$	
$\hat{\theta} (\%)$	R	$\hat{\theta} (\%)$	R	$\hat{\theta} (\%)$	R	$\hat{\theta} (\%)$	R	$\hat{\theta} (\%)$	R
50,0-57,1	1 to 5	50,0-50,9	1 to 6	50,0-54,2	2 to 6	50,0-53,3	2 to 7	50,0-51,1	3 to 7
57,1-60,7	2 to 5	50,9-64,4	2 to 6	54,2-58,4	2 to 7	53,3-58,3	3 to 7	51,1-58,7	3 to 8
60,7-71,7	2 to 6	64,4-65,2	2 to 7	58,4-68,8	3 to 7	58,3-63,7	3 to 8	58,7-61,8	4 to 8
71,7-83,9	3 to 6	65,2-76,2	3 to 7	68,8-69,5	3 to 8	63,7-71,7	4 to 8	61,8-67,8	4 to 9
83,9-93,4	4 to 6	76,2-86,3	4 to 7	69,5-79,5	4 to 8	71,7-73,3	4 to 9	67,8-74,2	5 to 9
93,4-99,2	5 to 6	86,3-94,4	5 to 7	79,5-88,2	5 to 8	73,3-81,9	5 to 9	74,2-76,2	5 to 10
$\geq 99,2$	6	94,4-99,3	6 to 7	88,2-95,1	6 to 8	81,9-89,5	6 to 9	76,2-83,9	6 to 10
		$\geq 99,3$	7	95,1-99,4	7 to 8	89,5-95,7	7 to 9	83,9-90,6	7 to 10
				$\geq 99,4$	8	95,7-99,5	8 to 9	90,6-96,1	8 to 10
						$\geq 99,5$	9	96,1-99,5	9 to 10
								$\geq 99,5$	10

Tableau J.2 : Intervalle acceptable de résultats corrects $R = [Y_{\min}, Y_{\max}]$ basée sur la densité de probabilité binomiale ($\geq 1\%$), et dépendant du nombre g_2 de ESEA2 et de la proportion observée de la modalité binaire majoritaire $\hat{\theta}$ avec la borne inférieure incluse et la borne supérieure exclue. Par exemple, si $\hat{\theta} = 53,6$; $R = [1 \text{ à } 6]$.

$g_2 = 6$		$g_2 = 7$		$g_2 = 8$		$g_2 = 9$		$g_2 = 10$	
$\hat{\theta}$ (%)	R	$\hat{\theta}$ (%)	R	$\hat{\theta}$ (%)	R	$\hat{\theta}$ (%)	R	$\hat{\theta}$ (%)	R
50,0-53,6	0 to 6	50,0-51,8	1 to 6	50,0-56,3	1 to 7	50,0-53,9	1 to 8	50,0-50,2	2 to 8
53,6-70,2	1 to 6	51,8-63,9	1 to 7	56,3-58,5	1 to 8	53,9-60,0	2 to 8	50,2-60,4	2 to 9
70,2-82,3	2 to 6	63,9-75,9	2 to 7	58,5-70,1	2 to 8	60,0-64,9	2 to 9	60,4-63,1	3 to 9
82,3-91,4	3 to 6	75,9-85,4	3 to 7	70,1-79,6	3 to 8	64,9-74,3	3 to 9	63,1-69,5	3 to 10
91,4-97,3	4 to 6	85,4-92,8	4 to 7	79,6-87,6	4 to 8	74,3-82,4	4 to 9	69,5-77,5	4 to 10
97,3-99,9	5 to 6	92,8-97,7	5 to 7	87,6-93,8	5 to 8	82,4-89,1	5 to 9	77,5-84,4	5 to 10
$\geq 99,9$	6	97,7-99,9	6 to 7	93,8-98,0	6 to 8	89,1-94,5	6 to 9	84,4-90,4	6 to 10
		$\geq 99,9$	7	98,0-99,9	7 to 8	94,5-98,3	7 to 9	90,4-95,1	7 to 10
				$\geq 99,9$	8	98,3-99,9	8 to 9	95,1-98,5	8 to 10
						$\geq 99,9$	9	98,5-99,9	9 to 10
								$\geq 99,9$	10

Le tableau J.3 permet de déterminer si la valeur assignée correspondant à la modalité majoritaire est statistiquement significative ($> 50\%$). Elle a été construite à partir de la loi bêta (Baudry et al, 2024).

Tableau J.3. : Consensus minimal m_{min} pour différentes valeurs de M qui est le nombre de laboratoire pour les ESEA4.

M	m_{min}	(%)	M	m_{min}	(%)	M	m_{min}	(%)	M	m_{min}	(%)
1	-	-	26	18	69,2	51	32	62,8	76	46	60,5
2	-	-	27	19	70,4	52	33	63,5	77	47	61,1
3	-	-	28	19	67,9	53	33	62,3	78	47	60,3
4	-	-	29	20	69,0	54	34	63,0	79	48	60,8
5	5	100,0	30	20	66,7	55	35	63,6	80	48	60,0
6	6	100,0	31	21	67,7	56	35	62,5	81	49	60,5
7	7	100,0	32	22	68,8	57	36	63,2	82	49	59,8
8	7	87,5	33	22	66,7	58	36	62,1	83	50	60,2
9	8	88,9	34	23	67,7	59	37	62,7	84	51	60,7
10	9	90,0	35	23	65,7	60	37	61,7	85	51	60,0
11	9	81,8	36	24	66,7	61	38	62,3	86	52	60,5
12	10	83,3	37	24	64,9	62	38	61,3	87	52	59,8
13	10	76,9	38	25	65,8	63	39	61,9	88	53	60,2
14	11	78,6	39	26	66,7	64	40	62,5	89	53	59,6
15	12	80,0	40	26	65,0	65	40	61,5	90	54	60,0
16	12	75,0	41	27	65,9	66	41	62,1	91	54	59,3
17	13	76,5	42	27	64,3	67	41	61,2	92	55	59,8
18	13	72,2	43	28	65,1	68	42	61,8	93	55	59,1
19	14	73,7	44	28	63,6	69	42	60,9	94	56	59,6
20	15	75,0	45	29	64,4	70	43	61,4	95	57	60,0
21	15	71,4	46	30	65,2	71	43	60,6	96	57	59,4
22	16	72,7	47	30	63,8	72	44	61,1	97	58	59,8
23	16	69,6	48	31	64,6	73	45	61,6	98	58	59,2
24	17	70,8	49	31	63,3	74	45	60,8	99	59	59,6
25	18	72,0	50	32	64,0	75	46	61,3	100	59	59,0

Exemple : pour $M = 28$ laboratoires participants, au moins $m_{min} = 19$ (soit 67.9 %) résultats correspondant à la modalité binaire majoritaire sont nécessaires pour définir la valeur assignée pour chaque ESEA4 du panel.

Annexe K : Calcul du score z composé

Deux méthodes de calcul peuvent être utilisées pour le calcul d'un score z composé :

- **Somme des scores z pondérés (Sum of Weighted z-scores - SWz)**

$$SW_Z = \frac{\sum_{i=1}^n |z_i| \cdot \omega(z_i)}{n_z}$$

Avec :

- n_z : nombre de scores
- $\omega(z_i) = \begin{cases} 1 & \text{si } |z_i| \leq 2,0 \\ 3 & \text{si } 2,0 < |z_i| < 3,0 \\ 5 & \text{si } |z_i| \geq 3,0 \end{cases}$

Cette méthode consiste à faire l'addition de :

- la somme des scores z compris entre 0 et 2 avec une pondération de 1
- la somme des scores z supérieurs à 2 mais inférieurs ou égal à 3 avec une pondération de 3
- la somme des scores z supérieurs à 3 avec une pondération de 5.

Diviser la somme totale par le nombre de scores n_z .

- **Somme des scores z carrés (Sum of squared z-scores - Sz²)**

$$S_Z^2 = \frac{\sum_{i=1}^n |z_i|^2}{n_z}$$

avec n_z : nombre de scores

Dans cette méthode, chaque score z est pondéré par lui-même (équivalent à considérer $\omega(z_i)=|z_i|$ dans la formule précédente). La pondération progresse donc de façon linéaire en fonction de z, plutôt que par paliers arbitraires.

Dans les deux cas, classer les laboratoires selon l'interprétation habituelle des scores.

Bibliographie

NF ISO 13528 : octobre 2022 - Méthodes statistiques utilisées dans les essais d'aptitude par comparaisons inter-laboratoires.

NF EN ISO /CEI 17043 : avril 2010 - Évaluation de la conformité - Exigences générales concernant les essais d'aptitude

NF EN ISO /CEI 17025 : décembre 2017 – Exigences générales concernant la compétence des laboratoires d'étalonnage et d'essais.

Thompson M., Ellison S.L.R., Wood R., (2006). The international harmonized protocol for the proficiency testing of analytical chemistry laboratories (IUPAC technical report), *Pure. Appli. Chem.*, 78, 1, 145-196.

NF EN ISO 22117 : mars 2019 – Microbiologie de la chaîne alimentaire — Exigences spécifiques et recommandations relatives aux essais d'aptitude par comparaison inter-laboratoires

NF U 47 400 : décembre 2014 – Méthode d'analyse en santé animale - Guide pour l'organisation des essais inter-laboratoires d'aptitude en santé animale

Thompson M., Lowthian P.J., (1996). Statistical aspect of proficiency testing in analytical laboratories-Part 2 testing for sufficient homogeneity. *Analyst*, 121, 1593-1596.

ILAC Discussion Paper on Homogeneity and Stability Testing. Tholen D. (2006), ILAC Proficiency Testing Consultative Group, 2nd meeting.

FD ISO Guide 35 : mars 2018 – Matériaux de référence – Lignes directrices pour la caractérisation et l'évaluation de l'homogénéité et de la stabilité.

Fearn T., Thompson, M., (2001). A new test for "sufficient homogeneity" *Analyst*, 126, 1414-1417.

Heisterkamp S.H., Hoeskstra J.A., van-Strip-Lockfeer N.G.W.M., Havelaar A.H., Mooijman K.A. 1993 in Veld P.H. et al., *Statistical analysis of certification trials for microbiological reference material*. Luxembourg: Commission des Communautés Européenne, (Report EUR 15008 EN), 41p

NF ISO 5725-2 : août 2020 - Exactitude (justesse et fidélité) des résultats et méthodes de mesure - Partie 2 : méthode de base pour la détermination de la répétabilité et de la reproductibilité d'une méthode de mesure normalisée.

Ulhig S., Blaul C., Frost K., Sgorzaly S., Colson B., Simon K. (2015). Qualitative PT data analysis with easy-to-interpret scores. *Accreditation Quality Assurance*, 20, 5, 347-353

Beavis G., Wilson J., Sykes M. (2019). Quantitative scores for binary qualitative proficiency testing. *Accreditation Quality Assurance*, 24, 1, 263–269.

Baudry C., Jadé G., Rayneau P., Lombard B., Feinberg M., Laurentie M. (2024) S-score: a new score for binary qualitative proficiency testing schemes interpretable as the z-score. *Accreditation and Quality Assurance*, vol 29, issue 2, 103-113.doi.org/10.1007/s00769-023-01561-y

Senyuva H.Z., Gilbert J., (2006). Assessment of the performance of pesticide-testing laboratories world-wide through proficiency testing. *Trends in Analytical Chemistry*, 25, 6, 554-562.

Medina-Pastor P., Rodriguez-Torreblanca C., Andersson A., Fernandez-Alba A.R., (2010a). European Commission proficiency tests for pesticide residues in fruits and vegetables. *Trends in Analytical Chemistry*, 29, 1, 70-83

Medina-Pastor P., Mezcua M., Rodríguez-Torreblanca C., Fernández-Alba A.R., (2010b). Laboratory assessment by combined z score values in proficiency tests: experience gained through the European Union proficiency tests for pesticide residues in fruits and vegetables. *Analytical and Bioanalytical Chemistry*, 397, 3061–3070.

General protocol for EU Proficiency Tests for Pesticide Residues in Food and Feed. 4rd Edition, approved November 2013. EU Reference Laboratories for Pesticide Residues. <http://www.eurl-pesticides.eu>