

Analyses statistiques validation d'un questionnaire

Utilisation du logiciel SPSS

Dr. Charlotte Beudart, Coordinatrice Scientifique, CHU Liège, CNRF, c.beudart@chuliege.be en collaboration avec le Pr. Christophe Demoulin, Département des sciences de la motricité Kinésithérapie spécifique et réadaptation motrice, ULiège

Introduction

Guide d'aide à la réalisation des statistiques visant à vérifier les propriétés psychométriques d'un questionnaire/d'une échelle/d'un outil.

Les questionnaires de santé sont de plus en plus employés en recherche médicale et en pratique clinique. Il est indispensable d'adopter une méthodologie scientifique pour développer ce type de questionnaire ou pour traduire des questionnaires de santé préexistants dans la littérature scientifique. Mais le développement/traduction d'un questionnaire ne suffit pas, une étape recherche supplémentaire reste fondamentale une fois le questionnaire développé ou traduit : l'étape de validation du questionnaire de santé. Il est en effet nécessaire de s'assurer que notre questionnaire développé ou traduit possède un niveau de qualité suffisant : il doit être démontré, par une analyse qualitative et quantitative, que notre questionnaire mesure bien ce qu'il est censé mesurer, que les mesures établies par le questionnaire sont reproductibles et que ce questionnaire permet de mettre en évidence des changements de l'état de santé.

Une méthodologie rigoureuse, issue des principes de l'épidémiologie, est proposée depuis 2007 par Terwee et al.¹ pour garantir un processus de validation du questionnaire robuste et scientifique. Il s'agit de la méthode de référence, publiée dans le « Journal of Clinical Epidemiology », qui a été citée à plus de 5000 reprises.

Base de données et import dans SPSS

Une fois l'ensemble de vos données recueillies, la base de données doit préalablement être construite dans Excel. Chaque ligne représente un individu et chaque colonne représente une variable. **Il est indispensable de donner un identifiant à chaque participant !** Une seule base de données doit être développée et doit comprendre l'ensemble des données récoltées pour les analyses ; données socio-démographiques et cliniques des participants, encodage du questionnaire (chaque item séparément + total), encodage des autres questionnaires/outils utilisés lors de la validation de l'outil, encodage du questionnaire (chaque item séparément + total) lors du re-test, etc. Il est important de correctement nommer les variables pour éviter toute confusion. Parallèlement, il est recommandé de développer un codebook visant à définir exhaustivement chaque nom de variable. Il est également important de nommer les variables sans espaces (astuce : utilisez _ pour indiquer un espace. Exemple : « PASI_ITEM1 »). Une fois votre base de données encodée dans le logiciel Excel, enregistrez-la en format Excel 97-2003 (format reconnu par SPSS).

ID	Age	Sexe	BMI	Niv_education			SF36_MCS	SF36_PCS	PASI_retest			
				PASI_item1	PASI_item2	PASI_Total			Item1	item2	Total	
1	45	0	22,1	1	7	4	11	74,1	58,4	7	4	11
2	54	1	25,4	3	6	5	11	69,5	74,1	5	5	10
3	37	0	19,7	3	4	4	8	81,2	70,1	3	4	7
4	61	0	28,1	2	7	1	8	41,7	67,4	7	1	8
5	45	1	25,1	4	7	5	12	68,4	78,4	7	5	12

Exemple d'une BD construite dans Excel pour vérification de propriétés psychométriques d'un questionnaire (PASI). Dans cet exemple, le MCS et PCS du SF-36 a été utilisé pour mesurer la validité de construit. Les résultats de ce questionnaire sont donc également encodés dans la BD Excel.

L'ensemble des statistiques visant à vérifier les propriétés psychométriques d'un questionnaire/d'une échelle/d'un outil peuvent se faire sur n'importe quel logiciel statistique. Ce guide décrit les analyses statistiques à réaliser avec le logiciel **SPSS**. Ce logiciel est toutefois payant. Nous vous invitons à télécharger la version d'essai du logiciel pour la réalisation de vos analyses. Attention que cette version d'essai gratuite n'est valable que 30 jours. Veillez donc bien à avoir une base de données prête et un plan d'analyses statistiques prêt avant de télécharger votre version d'essai.

<https://www.ibm.com/account/reg/be-en/signup?formid=urx-19774>

Une fois le logiciel installé, pour importer votre BD Excel dans SPSS, il suffira de cliquer sur « Fichier », « importer des données », « Excel ».

Vérification normalité des variables

La toute première étape consiste à vérifier la normalité des variables continues de votre base de données.

4 techniques à appliquer

1. Test statistique Shapiro-Wilk
2. Vérification et comparaison des moyennes et des médianes
3. Vérification des histogrammes
4. Vérification des QQPlot

Sur SPSS :

Analyse, explorer, sélection des variables quantitatives continues, cocher « tracés de répartition gaussienne avec tests » + « histogrammes » → permet de vérifier 1, 3 et 4

Analyse, Analyse descriptive, Fréquence, sélectionner les variables quantitatives continues, cocher « moyenne et écart type » + « médiane et quartiles » → permet de vérifier 2

Statistiques descriptives population

Décrire la population, présentation d'une table (voir exemple ci-dessous) comprenant différentes données socio-démographiques et cliniques.

Variables continues suivant une distribution normale → mean ± SD

Variables continues ne suivant pas une distribution normale → median P25-P75

Variables qualitative/binaires : fréquences (n, %)

Sur SPSS :

Analyse, Analyse descriptive, Fréquence, sélectionner les variables quantitatives continues, cocher « moyenne et écart type » OU « médiane et quartiles » OU ne rien cocher pour les variables qualitatives/binaires.

	Total sample n=101
Sex	
Women	63 (62.4)
Age (years)	25.0 (21.5-41.0)
Professional status	
Active	45 (44.6)
Student	35 (34.7)
Incapacity	8 (7.9)
Retired	6 (5.9)
Unemployed	7 (6.9)
BMI (kg/m ²)	24.1 (20.9-26.4)
Site of pain	
Neck	15 (14.9)
Back	27 (26.7)
Shoulder	18 (17.8)
Knee	25 (24.8)
Multi-site	16 (15.8)
ÖMP5Q-short	42.1 ± 16.7
KSMT (0/12)	5.74 ± 2.56

Variables that displayed normal distribution (ÖMP5Q-short and KSMT) are reported as mean (standard deviation). Non-normal variables (age and BMI) are reported as median and interquartile range (P25–P75). Qualitative and binary variables are expressed as absolute and relative frequencies (n, (%)).

Exemple d'une table descriptive de la population reprenant certaines variables sociodémographiques, cliniques, ainsi que les résultats de 2 questionnaires administrés pour évaluer la validité de construit.

Analyses psychométriques

1/ Cohérence interne

Deux éléments à mesurer

1. Alpha de Cronbach

L'alpha de Cronbach est une mesure de l'homogénéité de l'échelle et des différents items/domaines de l'échelle entre eux. Idéalement, on calculera une valeur d'alpha de Cronbach pour l'échelle complète et une mesure de l'alpha de Cronbach en supprimant, individuellement, un à un les items. L'idée de cette deuxième analyse est d'identifier un item qui aurait éventuellement un impact important sur l'homogénéité de l'échelle.

L'alpha de Cronbach varie entre 0 et 1. Un Alpha de Cronbach se situant entre 0.7 et 0.95 indiquera une bonne cohérence interne de l'échelle². En effet, une valeur supérieure à 0.95 pourrait indiquer une homogénéité trop importante des items ce qui pourrait correspondre à une redondance des items entre eux.

2. Corrélation item vs score total (sans l'item)

On mesure aussi, comme reflet de la cohérence interne, la corrélation entre chaque item et le score total de l'échelle (sans cet item).

Une corrélation supérieure à 0.6 et significative est considérée comme une corrélation importante³.

Sur SPSS : Analyse, Echelle, fiabilité, sélection de l'ensemble des items (mais pas le score total !) + cocher « alpha de Cronbach ». Dans statistiques, cocher « éléments », « échelle », « échelle sans l'élément » et « corrélation ».

On obtient alors l'alpha de Cronbach « total » ainsi qu'une table comme celle-ci-dessous reprenant plusieurs éléments.

Statistiques de total des éléments

	Moyenne de l'échelle en cas de suppression d'un élément	Variance de l'échelle en cas de suppression d'un élément	Corrélation complète des éléments corrigés	Carré de la corrélation multiple	Alpha de Cronbach en cas de suppression de l'élément
Item1	4,6139	3,599	,577	,398	,555
Item2	9,5050	11,052	,358	,157	,532
Item3	9,5743	10,667	,454	,368	,514
Item4	9,9406	11,736	,099	,126	,569
Item5	9,5446	11,230	,274	,172	,543
Item6	9,5446	11,150	,302	,287	,539
Item7	10,0891	11,902	,089	,056	,570
Item8	9,9604	10,518	,518	,350	,504
Item9	9,9604	11,158	,295	,164	,540
Item10	9,5842	11,185	,275	,284	,542

Exemple de table de résultat obtenue sur SPSS pour la mesure de l'Alpha de Cronbach. Les troisièmes et cinquièmes colonnes sont particulièrement intéressantes pour l'évaluation de la cohérence interne de l'échelle.

La dernière colonne renseigne l'alpha de Cronbach en excluant individuellement chacun des items.

La troisième colonne renseigne la corrélation de chaque item avec le score total de l'ensemble des 9 autres items. Attention toutefois que cette méthode de corrélation de chaque item avec le score total de l'échelle ne fonctionne que dans le cas où le score global de votre questionnaire correspond à une addition des différents items de l'échelle. Dans le cadre où le score total de l'échelle nécessite une statistique plus complexe (pondération plus importante pour certains items, par exemple), il conviendra d'utiliser une autre méthode de calcul de corrélation.

Pour ce faire, il sera nécessaire de calculer x nouveaux scores totaux, auxquels on retirera à chaque fois un item séparément. On mesurera ensuite une corrélation entre l'item 1 et le (score total – item 1) ; entre l'item 2 et le (score total – item 2), etc.

Sur SPSS :

1. Transformer, Calculer la variable, score_moins_item1 = (calcul du score total selon la pondération utilisée en supprimant l'item 1)
2. Analyse, corrélation, bivariée, sélection de chaque item séparément et (score total de l'échelle – ces items), cocher « Spearman » (en cas de distribution non normale des variables -préalablement testée) ou « Pearson » (en cas de distribution normale des variables)

2/ Fiabilité test-retest

En cas de variable continue (exemple : score total d'un questionnaire sur une échelle quantitative continue) : mesure d'un Coefficient de Corrélation Intraclasse (ICC) et de son intervalle de confiance à 95% entre le test (T1) et le retest (T2) sur le score total de l'échelle mais aussi sur les items/domaines individuels. Un ICC supérieur à 0.7 est considéré comme acceptable¹.

Sur SPSS : Analyse, Echelle, Sélectionner total questionnaire T1 et total questionnaire T2 (ou chaque item du questionnaire T1 avec chaque item du questionnaire T2 séparément), Statistiques, ICC two-way mixed, absolute agreement + interprétation de l'ICC obtenu en mesures uniques.

Coefficient de corrélation intra-classe

	Corrélation intra-classe ^b	Intervalle de confiance à 95 %		Test F avec valeur réelle 0			
		Borne inférieure	Borne supérieure	Valeur	ddl1	ddl2	Sig
Mesures uniques	,973 ^a	,953	,984	72,499	50	50	,000
Mesures moyennes	,986 ^c	,976	,992	72,499	50	50	,000

Modèle à effets mixtes à deux facteurs lorsque les effets sur les personnes sont aléatoires et les effets des mesures sont fixes.

Exemple de table de résultat obtenue sur SPSS pour la mesure de l'ICC et de son intervalle de confiance à 95%.

En cas de variable binaire (exemple : score total de l'échelle donne une donnée binaire - outil de diagnostic par exemple) : mesure d'un Kappa de Cohen et de sa p-valeur. Un Kappa de Cohen supérieur à 0.41 est considéré comme acceptable⁴.

Sur SPSS : Analyse, Fréquence, Tableau croisé, Sélectionner total questionnaire T1 et total questionnaire T2 (ou chaque item du questionnaire T1 avec chaque item du questionnaire T2 séparément), Kappa Cohen.

Mesures symétriques

		Valeur	Erreur asymptotique standard ^a	T approximatif ^b	Signification approximative
Mesure d'accord	Kappa				
		,891	,075	6,363	,000
N d'observations valides		51			

a. L'hypothèse nulle n'étant pas considérée.

b. Utilisation de l'erreur asymptotique standard en envisageant l'hypothèse nulle.

Exemple de table de résultat obtenue sur SPSS pour la mesure d'un Kappa de Cohen et de sa p-valeur.

3/ Standard Error of Measurement (SEM)

Applicable uniquement pour une échelle avec un score total continu.

Le SEM permet en quelques sortes de calculer la quantité d'erreur de mesure présente dans un instrument et reflète donc la fiabilité de cet instrument. Le SEM peut être utilisé pour fournir une fourchette autour de la valeur observée dans laquelle se trouve la « vraie » valeur théorique. L'intervalle ± 1 SEM donne une probabilité à 68% de contenir la vraie valeur ; l'intervalle ± 2 SEM donne la probabilité à 95% de contenir la vraie valeur⁵.

Pour calculer un SEM, il y a théoriquement deux méthodes :

1. Pour la première méthode, il sera nécessaire de préalablement mesurer la différence entre le score de la première (T1) et de la seconde (T2) administration (score du test – le score du retest = *Différence*).

ID	PASI_item1	PASI_item2	PASI_Total	PASI_item1 _retest	PASI_item2 _retest	PASI_Total _retest	Difference
1	5	7	12	5	6	11	1
2	4	7	11	4	7	11	0
3	5	5	10	5	4	9	1
4	2	4	6	3	5	8	-2
5	2	7	9	3	7	10	-1
6	5	7	12	5	7	12	0
7	5	7	12	5	7	12	0
8	5	6	11	2	5	7	4
9	3	5	8	3	5	8	0
10	3	6	9	3	7	10	-1

Exemple d'une BD construite dans Excel permettant de calculer la *Difference* entre le test (PASI_Total) et le retest (PASI_Total_retest). Ce calcul peut être effectué directement dans SPSS par le calcul d'une nouvelle variable.

Ensuite, il sera nécessaire de calculer l'écart type de cette différence (*SDdifference*) et de diviser la valeur obtenue par la racine de 2.

$$SEM = \frac{SD\ difference}{\sqrt{2}}$$

2. Dans la seconde méthode on utilise l'ICC mesuré à l'étape précédente. Il faudra mesurer l'écart type des valeurs du test (*SDT1*) et des valeurs du retest (*SDT2*). Grâce à ces deux SDs, il sera possible de calculer un écart type poolé (*SD pooled*) via la formule qui suit : $SD_{pooled} = \sqrt{\frac{SD1^2 + SD2^2}{2}}$
 $SEM = SD_{pooled} \times \sqrt{(1 - ICC)}$

4/ Minimal Detectable Change (MDC)

Avec le SEM, on calcule généralement également le MDC.

Le MDC est défini comme le changement du score de l'instrument au-delà de l'erreur de mesure. Cela signifie que le MDC fournit une valeur pour le changement minimum qui doit être observé afin d'être sûr que le changement observé est réel et non, potentiellement, un produit de l'erreur de mesure de l'instrument. Le MDC peut être calculé pour des sujets individuels (MDCind) ainsi que pour l'ensemble de l'échantillon (MDCgroup). Les deux sont utiles : Le MDCind peut être utilisé dans la pratique clinique ou pour indiquer si, individuellement il y a une modification chez chaque sujet. Le MDCgroup fournit une aide à l'interprétation des scores moyens de l'échantillon d'étude².

$$MDC_{ind} = 1.96 * SEM * \sqrt{2}$$

$$MDC_{group} = \frac{MDC_{ind}}{\sqrt{n}} \text{ ou } n = \text{le nombre de sujets inclus dans l'analyse}$$

5/ Limit of Agreement (LOA)


Les Limit of Agreement (LoA) (également connues sous le nom de méthode de *Bland et Altman*)^{6,7} sont un indicateur de l'erreur de mesure et peuvent être utiles lorsque vous souhaitez vérifier la présence d'une erreur systématique entre deux observateurs, deux instruments ou dans une situation de test-retest. Les LoA représentent les limites dans lesquelles se trouveraient 95% des valeurs de la distribution des erreurs de mesure. La méthode de *Bland et Altman* est basée sur un graphique tel que celui présenté ci-dessous (Plus d'infos : <https://www.medcalc.org/manual/blandaltman.php> ; <https://www.ibm.com/support/pages/bland-altman-plot>). Ce graphique est construit avec les variables suivantes :

- Créer une nouvelle variable avec la formule $\frac{(test(T1) + retest(T2))}{2}$ qui représente le score moyen du questionnaire de chaque participant. Dans cet exemple, nous appellerons cette variable "*T1T2_mean*". Sur le *Bland et Altman plot*, cette variable sera représentée sur l'axe des X.
- Créer une nouvelle variable avec la formule $(T1 - T2)$ représentant la différence entre le test et le retest pour chaque participant et qui sera représentée sur l'axe Y du *Bland et Altman plot*. Dans cet exemple nous appellerons cette variable "*T1T2_diff*".

Sur SPSS : Créer un Scatter plot avec ces deux variables en cliquant sur Graphiques, Boîte de dialogue, Dispersion/points – Dispersion simple. Mettre la variable *T1T2_mean* sur l'axe des X et la variable *T1T2_Diff* sur l'axe des Y.

Vous avez maintenant un Scatter Plot mais les LoA ne sont pas encore visibles. Pour les obtenir, vous devez d'abord calculer la différence moyenne entre les variables T1 et T2. Dans SPSS, cliquer sur Analyse – Comparer les moyennes – Test t pour échantillons appariés. Sélectionner les variables T1 et T2. Vous obtiendrez ainsi une table appelée « Paired samples Test ». Dans cette table la « Mean » représente la différence moyenne standardisée entre les deux scores, aussi appelée \bar{d} . Vous trouverez le “Std Deviation” qui vous sera utile pour calculer les LoA et que nous appellerons $SD_{\text{différence}}$.

		Paired Samples Test							
		Paired Differences			95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Overall_SF1_recode_test with3 - Overall_SF1_recode_RT	-3.34880	10.86966	.79487	-4.91692	-1.78069	-4.213	186	.000

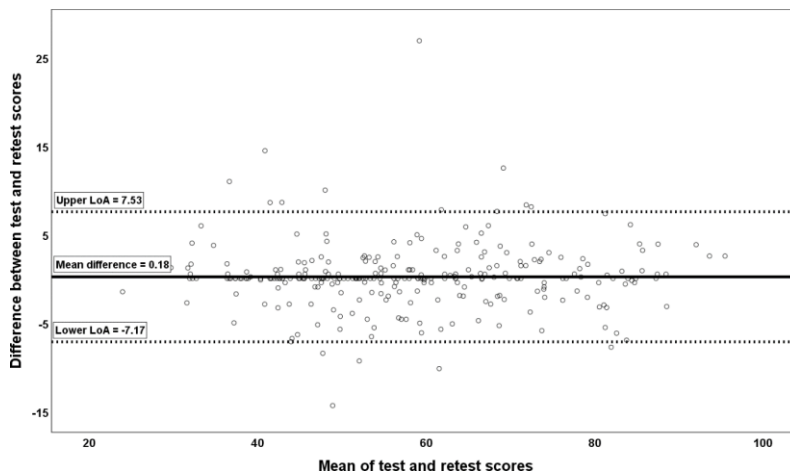
Sur SPSS, double cliquer sur le Scatter Plot pour ouvrir l'Editeur de Graphiques. Vous pouvez désormais insérer une ligne \bar{d} en cliquant sur ce symbole , ajouter une ligne de référence à l'axe Y. Il sera possible d'ajouter la valeur \bar{d} que vous avez préalablement calculé.

Vous pourrez ajouter les LoA de la même manière en utilisant les formules suivantes

- LoA supérieur = $\bar{d} + 1.96 \times SD_{\text{différence}}$
- LoA inférieur = $\bar{d} - 1.96 \times SD_{\text{différence}}$

Pour des raisons d'interprétation, une quatrième et dernière ligne de référence est souvent ajoutée à $y=0$.

Attention que le *Bland et Altman* nécessite que les variables T1 et T2 aient une distribution normale.



Exemple de graphique de Bland et Altman surlignant la Mean Difference et les LoA

6/ Validité de Construit / Contre un critère

Un questionnaire est dit « valide » si celui-ci mesure bien ce qu'il est supposé mesurer. Pour ce faire, il conviendra de vérifier que le score du questionnaire est bien corrélé à une mesure plus objective, dite Gold Standard (exemple= corrélation entre le score d'un questionnaire d'activité physique et une mesure plus objective d'activité physique par podomètre par exemple). Cette validité est **appelée validité contre critère**. Dans le cadre où un Gold standard n'existe pas (exemple : questionnaire de qualité de vie), il conviendra de mesurer la **validité de construit**. La validité de construit sera mesurée pour comparer l'échelle à une autre échelle ayant un construit similaire.

La validité se mesure à partir d'hypothèses qui sont développées en amont selon l'avis d'experts, l'expérience clinique ou des données obtenues par la littérature. Les statistiques consistent à mesurer la corrélation entre

les deux échelles (ou l'échelle et le Gold standard) pour vérifier les hypothèses développées. La validité de construit/reliée à un critère est confirmée si 75% des hypothèses sont vérifiées.

Une corrélation supérieure à 0.6 et significative est considérée comme une corrélation importante³.

Sur SPSS : Analyse, corrélation, sélection du score total du questionnaire A et soit le Gold standard ou le questionnaire B (ou de certains items/domaines du questionnaire, selon les hypothèses définies dans le protocole), Corrélation entre les deux mesures (Pearson ou Spearman selon la distribution des variables).

7/ Effets plancher/Plafond

Si >15% des répondants ont le score min / max au questionnaire.

8/ Responsiveness/Sensitivity to change

Ils existent plusieurs méthodes pour évaluer la sensibilité au changement d'un outil⁸. Deux approches sont présentées ci-dessous :

- Vérification d'hypothèses sur la corrélation entre les changements mesurés par l'outil/le questionnaire étudié (A) et les changements mesurés par un ou plusieurs autres outils de référence (B). La démarche est identique à celle pour mesurer la validité de construit, mais la mesure s'effectue sur les changements entre deux temps (T1 & T2).

Sur SPSS :

- 1) Créer une variable qui est la différence entre la valeur à T2 et T1 pour le questionnaire A: Transformer, calculer la variable, variable cible= $Diff_A$, expression numérique= $A_T2 - A_T1$. Identique pour outil B.
 - 2) Analyse, corrélation, sélection $Diff_A$ et $Diff_B$ (ou autres paires, selon les hypothèses définies dans le protocole), Corrélation entre les variables (Pearson ou Spearman selon la distribution des variables).
- Evaluation d'une taille d'effet (« effect size ») qui permet de savoir si les participants ont eu un réel changement sur la variable d'intérêt. L'idée sera de diviser l'échantillon de patients en trois catégories : 1) ceux qui ont amélioré leur résultat, 2) ceux qui ne l'ont pas changé/qui sont restés stables, et 3) ceux qui ont détérioré leur résultat. Dans cette méthode, l'idéal serait de demander aux participants s'ils ressentent une amélioration importante vs statut quo, vs détérioration de la variable d'intérêt (exemple : de leur qualité de vie « générale »). Il est donc nécessaire d'ajouter une question « d'ancrage » en plus du questionnaire.

Pour les trois catégories, on va ensuite calculer des « standardized response mean » (SRM) en utilisant la formule suivante : $\frac{\text{Moyenne}(Diff_A)}{\text{écart-type}(Diff_A)}$. Pour permettre d'utiliser les cut-offs de Cohen et al, on

transforme le SRM obtenu précédemment en utilisant la formule $\frac{SRM/\sqrt{2}}{\sqrt{(1-r)}}$ dans laquelle « r » est la corrélation entre les scores à T1 et à T2. Vous interprétez les SRM transformé comme suit : $SRM < 0.20$ est un effet négligeable, $0.20 \leq SRM < 0.49$ est un petit effet, $0.50 \leq SRM < 0.79$ est un effet modéré et $SRM \geq 0.80$ est un effet important⁹. Idéalement, on trouvera un effet important pour le groupe de personnes dont les résultats se sont soit améliorés ou détériorés et un effet négligeable dans le groupe de personne dont les résultats sont restés stables.

Sur SPSS :

- 1) Analyse, Analyse descriptive, Explorer, insérer le variable '*Diff_A*' dans la liste des variables dépendantes, et la variable catégorisant l'échantillon en trois groupes selon amélioration/stable/détérioration dans la liste des facteurs. Onglet « Options », cocher d'exclure seulement les composants non-valides. Vous trouverez ainsi la moyenne et l'écart-type de la variable *Diff_A* à utiliser dans les formules.
- 2) Analyse, corrélation, bivariée, sélection des variables *A_T1* et *A_T2*, cocher « Spearman » (distribution non normale) ou « Pearson » (distribution normale). Vous utilisez cette valeur pour transformer les SRM.

9/ Minimal Clinical Important Difference (MCID)

Le MCID est le changement minimal du score que le patient perçoit comme important, contrairement au SDC qui est le plus petit changement statistiquement significatif. L'un est une mesure statistique (puis-je détecter un changement ?), l'autre une mesure clinique (ce changement est-il important ?)¹⁰. Le MCID permet de vérifier si une petite différence après une intervention ou entre deux groupes est réellement cliniquement significative.

Une des techniques pour mesurer le MCID repose sur la méthode de l'ancrage, « Anchor-based approach »¹¹. Dans cette méthode on demandera aux participants s'ils ressentent une amélioration importante vs pas importante de la variable d'intérêt (exemple : de leur qualité de vie « générale »). Il est donc nécessaire d'ajouter une question « d'ancrage » en plus du questionnaire. On calculera ensuite un score moyen des sous-catégories de patients qui ont décrit une amélioration ou pas de la variable d'intérêt. On utilisera ensuite une ROC pour pouvoir distinguer amélioration importante vs pas importante.

Il existe d'autres techniques pour mesurer le MCID telles que la « Distribution based approach » ou encore, la combinaison des deux méthodes « Visual anchor-based MIC distribution ». Aucun consensus n'existe sur la méthode la plus appropriée à utiliser.

Références

1. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60(1):34-42. doi:10.1016/j.jclinepi.2006.03.012
2. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine.* Cambridge: Cambridge University Press; 2011. doi:10.1017/CBO9780511996214
3. Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures. Statistics and strategies for evaluation. *Control Clin Trials.* 1991;12(4 Suppl):142S-158S. doi:10.1016/S0197-2456(05)80019-4
4. McHugh ML. Interrater reliability: The kappa statistic. *Biochem Medica.* 2012;22(3):276-282. doi:10.11613/bm.2012.031
5. Standard Error of Measurement - Harvill - 1991 - Educational Measurement: Issues and Practice - Wiley Online Library. <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1745-3992.1991.tb00195.x>. Accessed March 24, 2021.
6. Martin Bland J, Altman DG. STATISTICAL METHODS FOR ASSESSING AGREEMENT BETWEEN TWO METHODS OF CLINICAL MEASUREMENT. *Lancet.* 1986;327(8476):307-310. doi:10.1016/S0140-6736(86)90837-8
7. Giavarina D. Understanding Bland Altman analysis. *Biochem Medica.* 2015;25(2):141-151. doi:10.11613/BM.2015.015
8. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: A critical review and recommendations. *J Clin Epidemiol.* 2000;53(5):459-468. doi:10.1016/S0895-4356(99)00206-1
9. Cohen J. *QUANTITATIVE METHODS IN PSYCHOLOGY A Power Primer.*
10. Middel B, Van Sonderen E. Statistical significant change versus relevant or important change in (quasi) experimental design: some conceptual and methodological problems in estimating magnitude of intervention-related change in health services research. *Int J Integr Care.* 2002;2(4). doi:10.5334/ijic.65
11. De Vet HCW, Ostelo RWJG, Terwee CB, et al. Minimally important change determined by a visual method integrating an anchor-based and a distribution-based approach. *Qual Life Res.* 2007;16(1):131-142. doi:10.1007/s11136-006-9109-9