

Position du problème.

What is the matter?

- Monsieur Dupont est âgé de 55 ans. Il est incapable de travailler depuis 1997, pour troubles dépressifs. Le médecin-conseil établit un rapport le 12/02/2002, concluant à une dépression sévère, avec algies psychosomatiques.
- Mr Smith is 55 years old. He has been unable to work since 1997, following a nervous breakdown. The medical advisor writes a report concluding: serious nervous breakdown with psychosomatic pain.

Position du problème.

What is the matter?

- Ce rapport est lu par 4 autres médecins -conseils qui doivent remettre une décision.
 - Le premier conclut que l'incapacité durera probablement encore un an et demande la transmission d'un rapport complémentaire.
- 4 other medical advisors read that report and have to take a decision.
 - The first one concludes that the patient will be unable to work for one year and asks for a new and more complete report .

Position du problème.

What is the matter?

- Le deuxième médecin estime que l'incapacité durera probablement encore 6 mois et demande un nouveau rapport.
- Le troisième médecin estime que l'incapacité sera définitive.
- The second one concludes that the patient will be unable to work probably for 6 months and asks for a new and more complete report.
- The third one thinks that the patient is permanently unable to work .

-
-
-

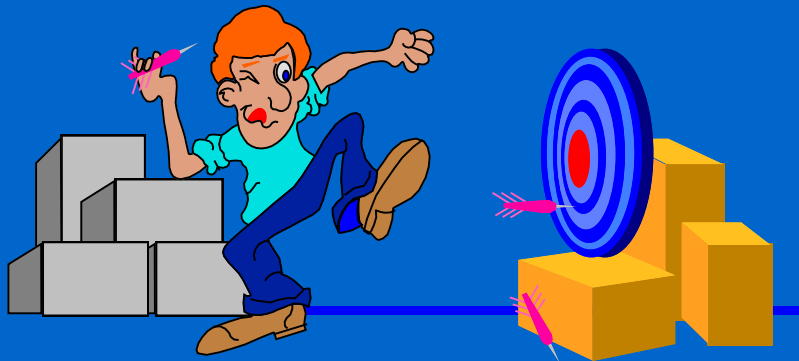
Position du problème.

What is the matter?

– Le quatrième enfin, conclut qu'il n'y a plus motif à incapacité de travail et décide de mettre fin aux indemnités.

– At last, the fourth one concludes that the patient is no longer unable to work and decides to end the payment of benefit allowance.

Position du problème. What is the matter?



- Un patient, 4 médecins, 4 approches différentes d'un même problème. Est-ce grave, docteur?
- One patient, 4 doctors and 4 different assessments. Is it serious, doctor?

Comment mesurer le problème?

How serious is the problem?

- La variabilité des pratiques médicales est un problème bien connu. Les médecins-conseils y échapperaient-ils dans leurs évaluations? Probablement non. Comment évaluer cette variabilité?
- Medical practise variability is a well known problem. Do medical advisors avoid it when assessing patients? Probably, they do not. How can we measure that variability?

Comment mesurer le problème? How serious is the problem?

- Différentes méthodes statistiques permettent d'estimer la concordance des décisions prises par différents évaluateurs.
- Different statistical methods allow us to estimate the decision concordance between different assessors.



•
•
•

Méthodes statistiques.

Statistical methods.

↳ Variables nominales et ordinales: coefficient κ de Cohen.

↳ Variables ordinales: γ de Goodman et Kruskal.

↳ Variables quantitatives: t test ou Wilcoxon.

↳ Nominal and ordinal variables: Cohen 's κ coefficient.

↳ Ordinal variables: Goodman, and Kruskal 's γ .

↳ ratio variables: t test or Wilcoxon.

Coefficient κ de Cohen

Cohen 's κ coefficient

- Permet d'évaluer la concordance de classement en variables nominales ou ordinales, par deux juges différents, ou, par plusieurs juges.
- It allows us to estimate the concordance between 2 or more assessors, rating the same population, by nominal or ordinal variables.

Coefficient κ de Cohen

Cohen 's κ coefficient

Ex1 / Ex2	Var 1	Var 2	Var 3	Total
Var 1	$l_i c_i$	n	n	l_i
Var 2	n	$l_i c_i$	n	l_i
Var 3	n	n	$l_i c_i$	l_i
Total	c_i	c_i	c_i	N

Coefficient κ de Cohen

Cohen 's κ coefficient

- Première évaluation de la concordance: le pourcentage d 'accord (population observée).
- First estimation of the concordance: the agreement percentage (observed population).

$$p_o = \frac{\sum \mathbf{1}_i \mathbf{C}_i}{\mathbf{N}}$$

•
•
•

Coefficient κ de Cohen

Cohen 's κ coefficient

↳ Le coefficient κ standardise la mesure en éliminant le hasard.

↳ Cohen 's κ gives a standardised measure by correcting chance agreement.

$$p_e = \frac{\sum \mathbf{1}_i \times \mathbf{c}_i}{\mathbf{N}^2}$$

⋮

Coefficient κ de Cohen

Cohen 's κ coefficient

$$\kappa = \frac{p_o - p_e}{1 - p_e}$$

⋮

Coefficient κ de Cohen

Cohen 's κ coefficient

Valeur de κ	Interprétation
0-0.19	Faible Low
0.20-0.39	Modéré Moderate
0.4-0.59	Bon Good
0.6-0.79	Très bon Very good
= 0.8	Excellent Excellent

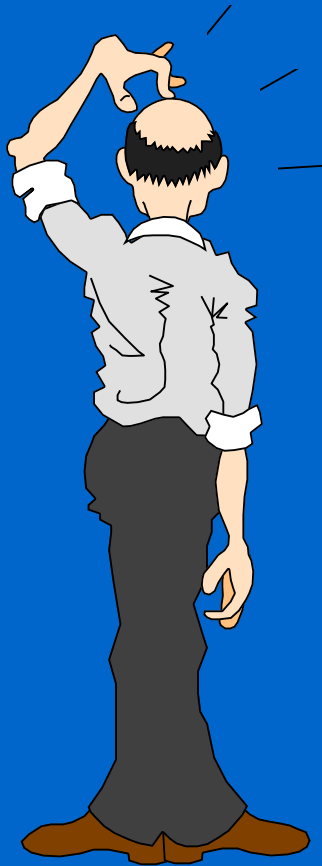
Goodman and Kruskal 's γ

Obs 1	Obs 2	
A	A	Concordant (P)
B	C	
B	B	Discordant (Q)
A	C	
B	B	Liées Tied.
B	A	

Goodman and Kruskal 's γ

$$\gamma = \frac{P - Q}{P + Q}$$

Goodman and Kruskal's γ



- Interprétation:
 - γ s'interprète comme un coefficient de corrélation, compris entre 0 et 1.
- Interpretation:
 - γ is a correlation coefficient with a value between 0 and 1.

-
-
-

Accord interjuges Interrater 's agreement.

- Premier exemple: évaluation de l 'invalidité par quatre médecins-conseils possédant une formation spécifique en médecine d 'assurance.
- First example: invalidity assessment by four medical advisors with specific training in insurance medicine.

Accord interjuges

Interrater 's agreement.

- Population:
 - échantillon aléatoire de 50 rapports d 'invalidité.
- Décisions:
 - probablement apte=1
 - renseignement complémentaire nécessaire =2
 - invalidité=3
- Population
 - random sample of 50 invalidity reports
- Decisions:
 - probably able=1
 - more information needed=2
 - invalidity=3

•
•
•

Accord interjuges Interrater 's agreement.

κ	Doc 2	Doc 3	Doc 4
Doc1	0.14	0.16	0.08
Doc2		0.34	0.4
Doc 3			0.34

-
-
-

Accord interjuges Interrater 's agreement.

- 2ème décision: pronostic.
 - Réinsertion immédiate = 1
 - réinsertion lointaine = 2
 - rééducation nécessaire = 3
 - inapte définitif = 4
 - ne se prononce pas = 5
- 2nd decision:
 - immediate reinsertion = 1
 - later reinsertion = 2
 - retraining needed = 3
 - definitively unable = 4
 - does not decide = 5

•
•
•

Accord interjuges Interrater 's agreement.

κ	Doc 2	Doc 3	Doc 4
Doc 1	0.37	0.44	0.31
Doc 2		0.4	0.24
Doc 3			0.3

•
•
•

Accord interjuges Interrater 's agreement.

Doc	Durée moyenne attendue Expected duration mean
1	8.3
2	6.1
3	9
4	5.3

•
•
•

Accord interjuges Interrater 's agreement.

p	Doc 2	Doc 3	Doc4
Doc 1	0.1 (NS)	0.57 (NS)	0.02(S)
Doc 2		0.67(NS)	0.24 (NS)
Doc 3			0.01(S)

•
•
•

Accord interjuges: test-retest.

Interrater 's agreement: test-retest.

- Le retest permet de vérifier la fidélité de l'évaluateur par rapport à lui-même.
- The retest allows us to estimate the assessor's reliability with himself.

Retest

Doctor	Décision Decision κ	Pronostic Prognostic κ	Moyenne1 Mean 1	Moyenne2 Mean 2	P
1	0.49	0.44	8.3	7.5	0.92(NS)
2	0.41	0.48	6.1	6.9	0.45 (NS)
3	0.26	0.45	9	8.9	0.77(NS)
4	0.65	0.55	5.3	5.5	0.78 (NS)

•
•
•

Accord interjuges: usages d 'échelles Interrater 's agreement: using scales

- Des échelles d 'évaluation peuvent, si elles sont valides, améliorer la concordance de la décision.
- Assessment scales may improve the concordance, if they are valid.

Accord interjuges: usages d 'échelles

Interrater 's agreement: using scales

Item AGGIR	κ	γ
Coherence	0.74	0.98
Orientation	0.75	0.98
Toilette(washing)	0.86	0.99
Habillage (dressing)	0.79	0.99
Manger (eating)	0.72	0.98
Eliminer (toileting)	0.81	0.99
Transfert (transfer)	0.75	0.98
Déplacement (moving)	0.69	1.00
GIR	0.76	0.99

Accord interjuges: usages d 'échelles

Interrater 's agreement: using scales

Items Katz	κ	γ
Laver (washing)	0.64	0.98
Habiller (dress ing)	0.54	0.96
Transferts déplacements Transfer moving	0.56	0.94
Toilettes toileting	0.58	0.96
Incontinence	0.59	0.97
Manger eating	0.44	0.91
Désorientation espace (space)	0.51	0.96
Désorientation temps (time)	0.51	0.97
Catégorie	0.78	0.96

Accord interjuges: usages d 'échelles

Interrater 's agreement: using scales

Item APA	κ	γ
Déplacements <i>moving</i>	0.43	0.90
Manger <i>eating</i>	0.31	0.68
Hygiène	0.58	0.96
Maison <i>household</i>	0.20	0.78
supervision	0.46	0.96
Contacts sociaux <i>Social contacts</i>	0.31	0.86
catégorie	0.47	0.91

Conclusions

- Pour l'évaluation de l'invalidité en Belgique, l'élaboration de lignes directrices et l'utilisation de la CIF-2 devraient être testées, en vue d'améliorer la fidélité des décisions.
- Des échelles valides peuvent être une aide efficace à la décision si elles sont fidèles.
- In Belgium, we have to test the use of guidelines in assessment and the use of ICF-2, in order to improve interrater reliability when assessing invalidity.
- Valid scales can be an efficient help to take decisions, if they are reliable.