



Tout ce que vous avez voulu
savoir sur la docimologie des
KFP,...en osant le demander

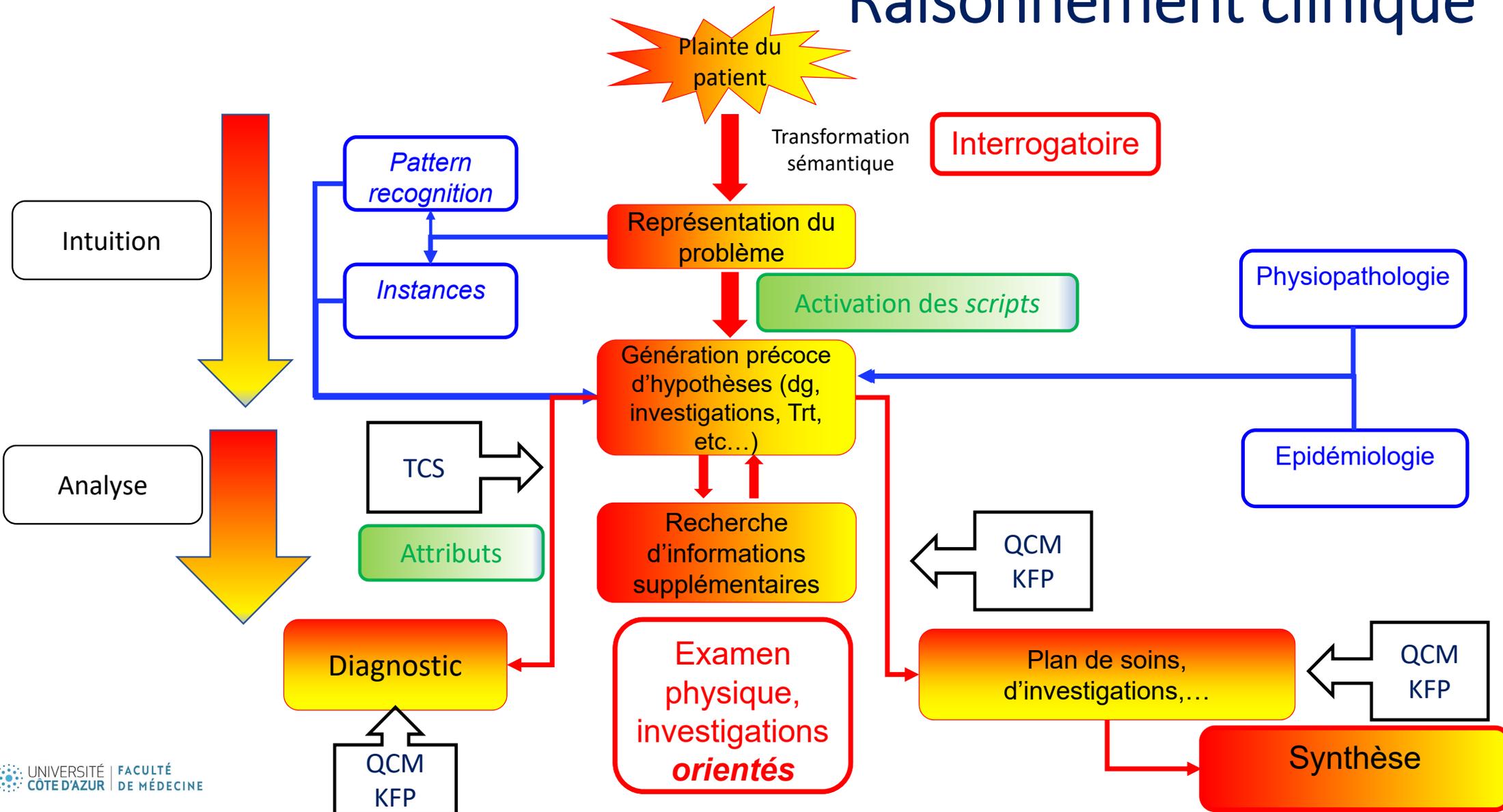
L. Sibert (Rouen), JP. Fournier (Nice)

Pourquoi introduire les KFP aux EDN ?

Assessment method: Definition	Clinical reasoning component						
	IG	HG	PR	DD	LD	DJ	MT
Non-workplace-based assessments							
Clinical or comprehensive integrative puzzles: An extended matching crossword puzzle designed to assess a learner's ability to relate clinical vignettes to specific diagnoses and diagnostic or therapeutic interventions.	0.4	0.3	0.6	1.1	1.9	0.4	1.3
Concept maps: A schematic method for learners to organize and represent their knowledge and knowledge structures through a graphical illustration of the complex processes and relationships between concepts within a subject domain.	0.4	0.4	1.2	1.0	0.4	0.8	0.9
Extended matching questions: A written exam format consisting of a lead-in question (clinical vignette) followed by multiple answer options in a list where more answer options are given than in multiple-choice questions (i.e., > 5).	0.2	0.3	0.2	0.8	1.7	0.3	1.3
Key feature examinations: Problems typically consist of a clinical vignette followed by 2–3 questions that assess the critical elements ("key features") or challenging decisions that clinicians must make.	0.9	0.5	0.4	1.5	1.4	0.6	1.4
Multiple-choice questions: A clinical vignette is followed by up to 5 alternatives. Questions may take the following formats: single best alternative, matching, true or false, and combinations of alternatives.	0.9	0.3	0.0	0.6	1.9	0.0	1.8
Modified essay questions: A method wherein serial information about a clinical case is presented chronologically. After each item, the learner must document a decision. The student cannot preview subsequent items until a decision is made.	1.3	1.2	1.0	1.6	1.7	1.3	1.7
Oral examinations: A verbal examination conducted by one or more faculty members through unscripted or semiscripted questions that assess clinical reasoning and decision-making abilities, as well as professional values.	1.3	1.3	1.1	1.8	1.8	1.9	1.9
Patient management problems: A clinical scenario is presented in real-life settings with specific resources available for diagnosis or management. The learner chooses among multiple alternatives. The results of actions (e.g., labs, images) are provided.	1.6	1.0	0.3	1.4	1.9	0.6	1.7
Script concordance tests: Clinical scenarios with uncertainty are followed by a series of questions (e.g., if you are thinking X and you find Y, the answer becomes more likely, less likely, or no change). Responses are compared with those of experts.	0.4	0.8	0.6	0.8	1.3	0.9	1.1
Short- or long-answer (essay) questions: A clinical vignette is followed by one or more questions. Learners provide free-text responses that range in length from a few words to several sentences.	0.8	1.2	1.2	1.8	1.7	1.8	1.7

IG : Information gathering
 HG : Hypothesis generation
 PR : Problem representation
 DD : Differential diagnosis
 LD : Leading diagnosis
 DJ : Diagnosis justifications
 MT : Management and treatment

Raisonnement clinique



Les KFP c'est quoi ?

- ❖ Etapes(s) essentielle(s) et difficile(s) de la résolution d'un problème
- ❖ Etapes difficiles(s) dans l'identification et la résolution d'un problème en pratique
- ❖ Etape(s) à l'origine d'erreurs habituelles de la part des étudiants évalués
 - ↪ Mesure *discriminante* de la compétence
 - ↪ On s'intéresse plus à l'objet de la mesure qu'à la mesure elle-même : situation-clé et étapes-clé
 - ↪ On s'intéresse à quelques étapes-clé de la résolution du problème
 - ↪ Les formats de questions (QRM vs QROC) sont choisis pour obtenir la meilleure discrimination

Les KFP c'est quoi ?

At 20:00h, the nurse in the emergency department asks you to see a 24-year-old woman, who has 2 children and is now 31 weeks pregnant, reporting of bright red bleeding coming from her vagina that began about two hours ago. The three sanitary napkins that she used were completely soaked. Her pregnancy has been normal, as were the two previous ones. She has not had any pain in her belly and she says that the fetus is moving as usual.

Her BP is 110/70 mm Hg and her pulse is 92/min. Your examination of the abdomen reveals a uterine height of 31 cm with a soft, non-tender uterus. The fetus is in a breech position and has a heart rate of 150/min. No bleeding has occurred since she arrived 25 minutes ago.

Question 1. What is your leading diagnosis at this time? List only one or write "normal" if you judge this situation to be within normal limits.

1. _____

Question 2. What steps will you take in your immediate assessment and management of this patient? Select as many as are appropriate.

1. Artificial rupture of membranes
2. Cervical swab for chlamydia
3. Complete blood count
4. Computed tomographic scan of abdomen & pelvis
5. Cross and match for 2 units of blood
6. Discharge home with instructions to come back if bleeding starts again
7. Fetal scalp monitor to assess status of fetus
8. Human immunodeficiency virus serology
9. International normalized ratio (INR)
10. Large bore intravenous line
11. Partial thromboplastin time (PTT)
12. Pelvic ultrasound
13. Pelvic examination
14. Syphilis serology
15. Ultrasound using vaginal probe
16. Vaginal swab for Group B streptococcus

Vignette clinique courte
(*Situation Clinique de
Départ*)

2 à 3 questions (= *Key-features*) :
❖ Questions à réponses ouvertes (QROC) ;
❖ "*Short menu*" : 15 à 25 options de réponses ;
❖ "*Long menu*" : jusqu'à 500 options de réponses.

Key-Feature
Problem

D'après

Page G. *Acad Med* 1995

Farmer E. *Med Educ* 2006

Medical council of Canada Guidelines for the development of key-feature problems and test cases, 2012

Ces objectifs sont-ils atteints ?

Quelles validités ?

❖ Validité apparente :

- ✓ Mesure de compétences
- ✓ Niveau de difficulté acceptable ?
 - Taux de réussite : 0,3 à 0,8
- ✓ Indice de discrimination correct ?
 - $\geq 0,25$ pour 12 KFP sur 15

❖ Validité de contenu :

- ✓ Tags et table de spécification
- ✓ Situations Cliniques de Départ : 356 !

❖ Validité externe :

- ✓ Corrélation faiblement significative aux autres formats de questions : 0,43 (ECOS) à 0,55 (QCM)

D'après

Fischer MR. *Med Teach* 2005
Bronander KA. *Med Teach* 2015
Huwendiek S. *Med Teach* 2017

Ces objectifs sont-ils atteints ?

Quelles validités ?

- ❖ Validité de construit :
 - ✓ Meilleure performance en utilisant un langage courant qu' avec un langage médical
 - ✓ Langue maternelle sans influence
 - ✓ Meilleure performance d'experts vs novices
 - ✓ Corrélation significative avec la réflexion à voix haute
 - ✓ Faible niveau de corrélation avec les formats explorant des connaissances pures
 - ✓ Faible corrélation avec les évaluations par patients standardisés

- ❖ Validité prédictive :
 - ✓ Corrélation avec la "vraie vie"

D'après

Hrynchak P. *Med Educ* 2014
Bordage G. *Acad Med* 1995
Eva K. *Med Educ* 2010
Farmer EA. *Aust Fam Physician* 2005
Rademakers J. *Med Teach* 2009
Schuwirth LWT. *Med Educ* 2001
Hatala R. *Med Educ* 2002
Fisher MR. *Med Teach* 2005
Tamblyn R. *Arch Intern Med* 2010
Tamblyn R. *JAMA* 2007
Huwendiek S. *Med Teach* 2017

Ces objectifs sont-ils atteints ?

Quelle stabilité des scores ?

- ✓ Cronbach alpha : 0,70 à 0,95 pour 25 à 40 cas (3 à 4 heures d'épreuve)
- ✓ Coefficient G : 0,45 à 0,83 ;
- ✓ Score non influencé par le format des questions (texte libre vs short ou long-menu).

D'après

Hrynchak P. *Med Educ* 2014
Fisher MR. *Med Teach* 2005
Huwendiek S. *Med Teach* 2017

Ces objectifs sont-ils atteints ?

Qu'en pensent les étudiants ?

- ✓ Format apprécié
- ✓ Impact didactique plus important que les QCM
- ✓ Format "short menu" plus difficile que le "long menu"
- ✓ Caractère séquentiel peu apprécié

Données psychométriques observées

- ❖ Examen Cumulatif Transversal juin 2021
- ❖ DFASM 1 et DFASM 2
- ❖ Pour chaque année : 4 DCP multi-disciplinaires et 8 KFP **communs**
- ❖ 3 heures d'épreuve (2 heures pour les DCP et 1 heure pour les 8 KFP)
- ❖ Administration *via* SIDES
- ❖ Etudiants prévenus, disposant d'un tutoriel succinct
- ❖ 183 étudiants en DFASM 1 et 186 en DFASM 2

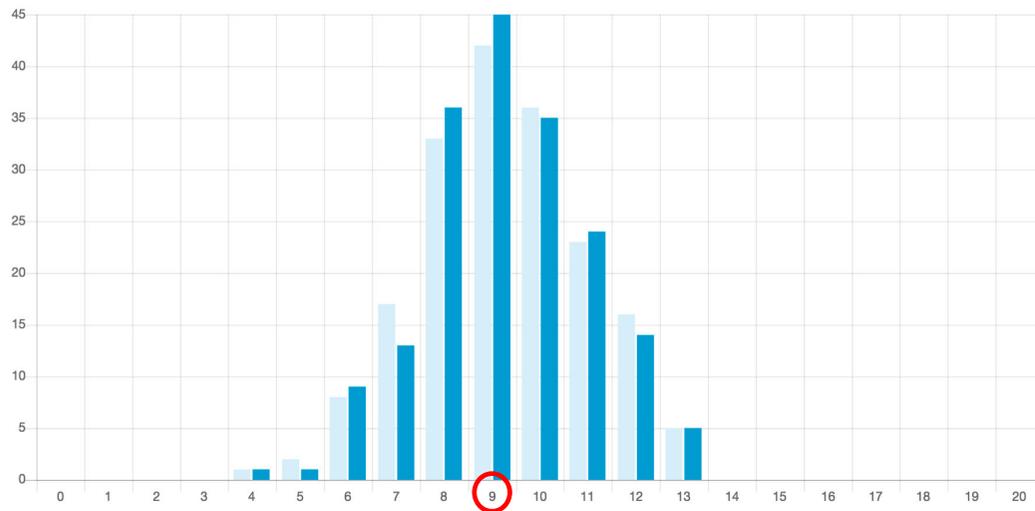
Données psychométriques observées

- ❖ Troubles de conscience chez l'éthylique (Gayet Wernické) :
SDD : 49, 119, 178, Questions : 76, 343, 351 (rangs A et B)
- ❖ « Insuffisance cardio-respiratoire » (insuffisance cardiaque congestive) :
SDD : 162, 178, 204, Questions : 203, 208, 234, 359 (rangs A et B)
- ❖ Hyponatrémie (potomanie) :
SDD : 120, 178, 202, 253, Questions : 342, 346 (rangs A et B)
- ❖ Hémorragie digestive haute (hypertension portale sur cirrhose) :
SDD : 14, 60, 178, 217 et 223, Questions : 213, 279, 280, 355 (rangs A et B)
- ❖ Diarrhée aiguë (colite à *C. difficile*)
SDD : 178, 249, Questions : 176, 286 (rangs A et B)
- ❖ Dyspnée aiguë (acido-cétose diabétique) :
SDD : 162, 178, 192, 197, Questions : 247, 267 (rangs A et B)
- ❖ Crise convulsive généralisée (sevrage éthylique) :
SDD : 120, 128, 178, Questions : 76, 342, 346 (rangs A et B)
- ❖ Grosse jambe douloureuse (thrombophlébite) :
SDD : 71,87, 178, 248, Questions : 226, 257, 350 (rangs A et B)

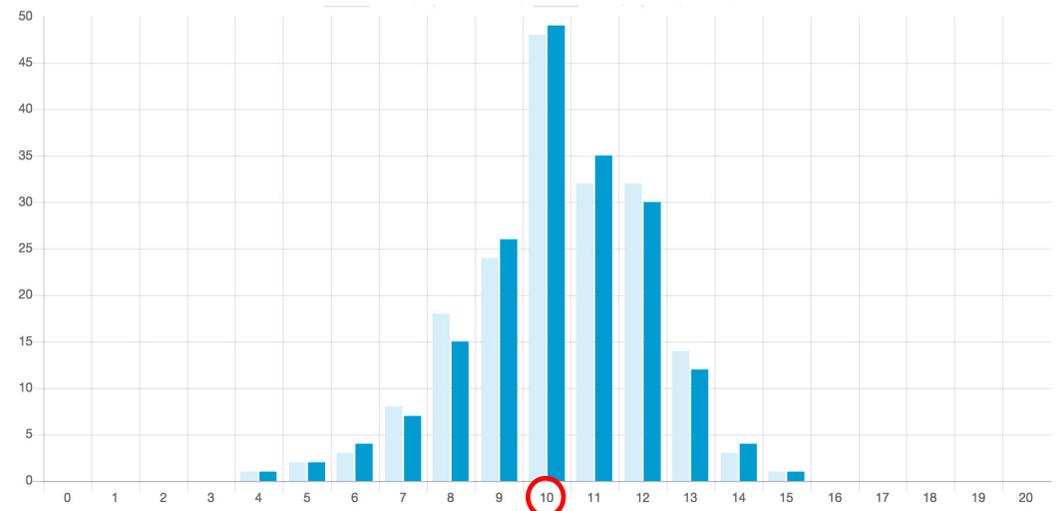
26 questions au total :

- ❖ 20 QRM (13 ± 2 options de réponses)
- ❖ 2 QRU (10 options de réponses)
- ❖ 4 QROC

Données psychométriques observées



DFASM 1



DFASM 2

Données psychométriques observées

	DFASM 1	DFASM 2	p
Temps de rédaction (min)	58,05 ± 2,69	58,85 ± 4,23	< 0,001
Taux de réussite	0,15 ± 0,18	0,18 ± 0,20	NS
Index de discrimination	0,29 ± 0,34	0,30 ± 0,32	NS
Score relatif	0,49 ± 0,09	0,53 ± 0,09	< 0,001
Cronbach alpha	0,35	0,39	

Données psychométriques observées

		DFASM 1	DFASM 2
QRM	Taux de réussite	0,11 ± 0,15	0,09 ± 0,11
	Index de discrimination	0,15 ± 0,18	0,14 ± 0,12
QRU	Taux de réussite	0,47 ± 0,01	0,49 ± 0,03
	Index de discrimination	0,89 ± 0,13	0,85 ± 0,04
QROC	Taux de réussite	0,23 ± 0,27	0,41 ± 0,20
	Index de discrimination	0,65 ± 0,45	0,81 ± 0,20

Données psychométriques observées

Q2- Parmi les examens suivants, lesquels seront les plus contributifs à votre prise en charge de ce patient ? (**QRM – 5 réponses attendues**)

A. clichés d'abdomen sans préparation

B. coloscopie

C. coproculture

D. C Reactive Protein

E. ECG

F. examen parasitologique des selles

G. gazométrie artérielle

H. hémogramme

I. formule sanguine

J. hémocultures

K. ionogramme sanguin

L. lactates veineux

M. recherche de toxines de *C. difficile* dans les selles

N. TDM abdominale injectée

O. TP-TCA-fibrinogène

DFASM 1 :

Taux de réussite : 0

Index de discrimination : 0

DFASM 2 :

Taux de réussite : 0,02

Index de discrimination : 0,04

Données psychométriques observées

QRM et KFP :

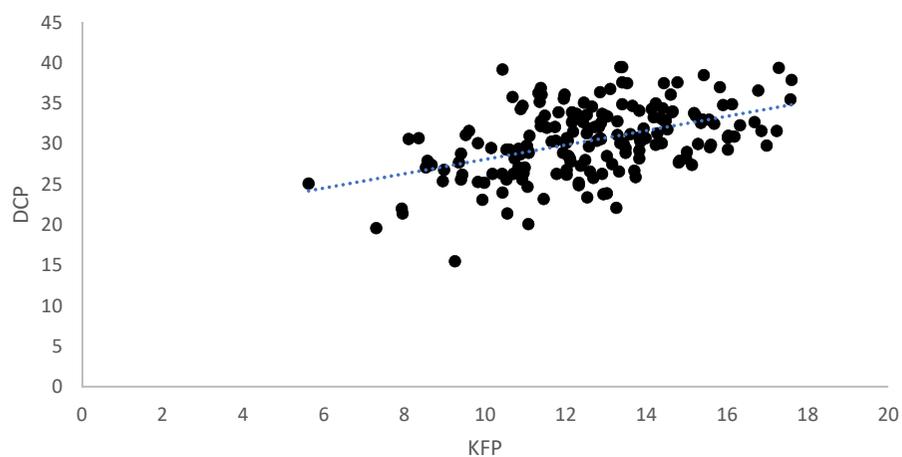
- ❖ Défaut d'entraînement des rédacteurs ?
- ❖ Défaut d'entraînement des étudiants ?
- ❖ Distracteurs ?
- ❖ Plus d'un objectif ?
- ❖ A réserver :
 - ✓ Antécédents
 - ✓ Histoire clinique
 - ✓ Examen clinique
 - ✓ Prescription d'examens complémentaires
 - ✓ Analyse de résultats d'examens
- ❖ QROC :
 - ✓ Diagnostic
 - ✓ Prise en charge
 - ✓ Prescriptions
- ❖ QRM moins discriminants que les QROC

Données psychométriques observées

	DFASM 1	DFASM 2	p
Effectif	183	186	
Ex aequos DCP	119 (0,65)	91 (0,49)	0,002
Ex aequos KFP	54 (0,29)	70 (0,37)	NS
Ex aequos DCP + KFP	28 (0,15)	27 (0,14)	NS

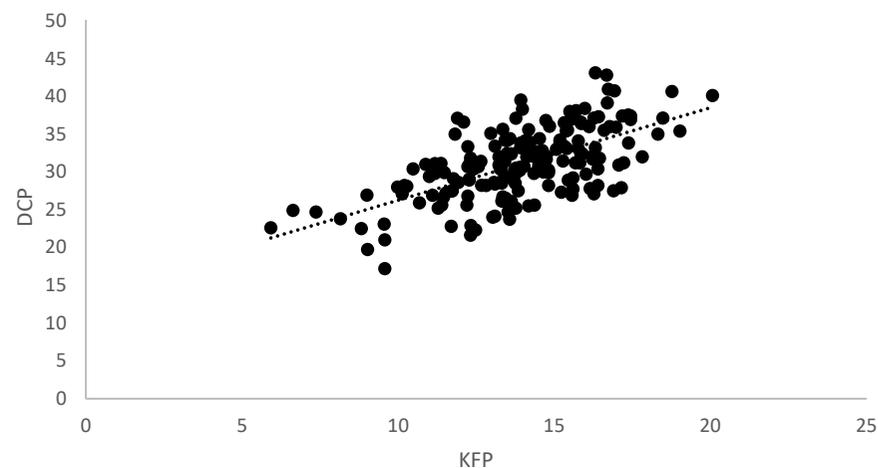
Données psychométriques observées

Corrélation DCP/KFP en DFASM1



Rho = 0,4662 [0,3444 – 0,5725]
p < 0,001

Corrélation DCP/KFP en DFASM2



Rho = 0,6266 [0,5306 – 0,7068]
p < 0,001

DCP ou KFP déguisés ?

En résumé

- ❖ Format séduisant
- ❖ Largement documenté
- ❖ Nécessité d'un entraînement pour étudiants et enseignants
- ❖ Apport raisonnable pour l'évaluation des compétences
- ❖ Limitation des ex aequos
- ❖ Nécessité de QROC
- ❖ Rédaction des QRM

En conclusion

- ❖ **Aucun format ne peut « tout faire » ;**
- ❖ 3 formats docimologiquement corrects ;
- ❖ 3 formats **complémentaires** :
 - ✓ Même approche des étudiants les plus performants que les cliniciens entraînés aux QCM à contexte riche ;
 - ✓ Mêmes erreurs des étudiants les moins performants sur des QCM à contexte riche que dans la « vraie vie » ;
 - ✓ Les QCM représentent l'aboutissement du raisonnement (prise de décision) ;
 - ✓ Les options des QCM, des KFP et des TCS correspondent à la phase d'intuition (génération des hypothèses pertinentes) ;
 - ✓ Les TCS mesurent le cheminement du raisonnement (phase analytique) ;
 - ✓ Les formes les plus évoluées de QCM à contexte riche permettent de représenter le problème (Bloom 5) et de mesurer le cheminement clinique (identification d'informations pertinentes – Bloom 4) ;
 - ✓ Les KFP permettent une approche « en profondeur » avec un impact didactique ;
 - ✓ KFP et QCM à contexte riche sont prédicteurs de la performance des futurs internes.

D'après

Van der Vleuten CPM. *BMJ* 2000
Evans JSBT *Pers Psychol Sci* 2013
Rangel RH. *Med Teach* 2017
Heist BJ. *J Grad Med Educ* 2014
Surry LT. *Med Educ* 2017
Lubarsky S. *Med Teach* 2013
Krathwol DR. *Theory into Pract* 2002
Daniel M. *Acad Med* 2019
Hrynychak P. *Med Educ* 2014
Huwendiek S. *Med Teach* 201
Lee M. *Med Teach* 2018
Sharma S. *J Grad Med Educ* 2019