



**HAL**  
open science

## Le Test de Concordance de Script comme outil d'évaluation en Néonatalogie

Amira Ben Hamida

► **To cite this version:**

Amira Ben Hamida. Le Test de Concordance de Script comme outil d'évaluation en Néonatalogie. Médecine humaine et pathologie. 2015. dumas-01301670

**HAL Id: dumas-01301670**

**<https://dumas.ccsd.cnrs.fr/dumas-01301670>**

Submitted on 12 Apr 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**UNIVERSITE VICTOR SEGALEN BORDEAUX 2**

ANNÉE UNIVERSITAIRE 2014/2015

---

# **MEMOIRE**

Pour l'obtention du

## **DIPLOME UNIVERSITAIRE DE PEDAGOGIE DES SCIENCES DE LA SANTE**

Présenté le 12/06/2015

par

Dr Amira Ben Hamida Nouaili

<b>TITRE</b>	<b>Le Test de Concordance de Script comme outil d'évaluation en Néonatalogie</b>
Mots-clés	Test de concordance de script, Evaluation, Compétence, Raisonnement, Néonatalogie.

**Jury :**

Pr Emmanuel Cuny

Dr Julien Naud

Pr André Quinton

**Directeur de Mémoire :** Pr André Quinton

## **Remerciements**

Je remercie le Pr André Quinton pour toute l'aide précieuse qu'il m'a apportée pour l'élaboration de ce travail.

Je remercie les membres du jury d'avoir accepté de juger notre travail.

A toutes les personnes qui ont contribué de près ou de loin à l'élaboration de ce travail.

## Plan

<b>I-</b>	Introduction.....	5
<b>II-</b>	Méthodes.....	7
	<b>II-1</b> Population d'étude.....	7
	<b>II-2</b> Rédaction du TCS.....	7
	<b>II-3</b> Validation des items et choix des experts.....	9
	<b>II-4</b> Elaboration des scores : méthode des scores combinés.....	9
	<b>II-5</b> Réalisation du test.....	10
	<b>II-6</b> Etude statistique.....	10
<b>III-</b>	Résultats.....	12
	<b>III-1</b> Résultats globaux .....	12
	<b>III-2</b> Résultats après élimination des items et ancrages erronés ou discutables.....	14
<b>IV-</b>	Discussion.....	18
	<b>IV-1</b> Evaluation des scores moyens.....	18
	<b>IV-2</b> Critères psychométriques du TCS.....	19
<b>V-</b>	Perspectives.....	21
<b>VI-</b>	Conclusion.....	24
	Références.....	26
	Annexes.....	29
	Résumé.....	44
	Abstract.....	45

# INTRODUCTION

**L'objectif de ce travail** était de construire un test de concordance de script (TCS) et de démontrer sa pertinence dans l'évaluation des compétences en Néonatalogie des étudiants du deuxième cycle et des internes de la faculté de médecine de Tunis.

Le TCS est un instrument d'évaluation du raisonnement en contexte d'incertitude [1-3]. Dans leur pratique de tous les jours, les professionnels de la santé rencontrent des problèmes variés, plus au moins complexes. Pour les résoudre, ils font appel à leurs connaissances, à leur expérience et à leur raisonnement clinique. Il peut s'agir de données bien définies, complètes et précises ainsi le but à atteindre est clair et une solution est disponible. Mais il peut s'agir aussi de données incomplètes, ambiguës ou imprécises où le but à atteindre n'est pas toujours clair. Les données peuvent faire l'objet de différentes interprétations et plusieurs solutions peuvent être proposées. Ce type de problème exige une capacité à raisonner en contexte d'incertitude [2].

Le TCS complète les méthodes d'évaluation traditionnelles déjà existantes comme les questions à choix multiples (QCM) ou les questions à réponses ouvertes et courtes (QROC), qui explorent bien les connaissances liées aux problèmes cliniques bien définis. Il permet ainsi d'enrichir de façon complémentaire l'évaluation des apprentissages, en tant que moyen d'approcher le raisonnement face à des problèmes mal définis, c'est-à-dire lorsque toutes les données nécessaires à la résolution du problème ne sont pas présentes ou lorsqu'il existe des données ambiguës, ce qui est la situation quotidienne en pratique clinique [2-4].

Le TCS permet d'approcher non plus les seules connaissances techniques mais le raisonnement clinique lui-même. La théorie des scripts, postule qu'en situation clinique les médecins mobilisent des réseaux de connaissance préétablis, des scripts, qui sont utilisés pour comprendre cette situation et agir en fonction de buts diagnostiques, d'investigation ou de traitement. Les scripts diagnostiques contiennent les associations que le clinicien a établies entre l'entité pathologique et ses différents attributs cliniques (les signes et symptômes notamment). Ce sont ces associations, ces liens, qui dans le contexte clinique permettent de prendre les décisions concernant la force ou la faiblesse d'une hypothèse ou de décider si un signe ou un symptôme n'est jamais associé à une maladie et donc que cette hypothèse doit être éliminée [5,6].

La validité du TCS est aujourd'hui bien établie [7]. Une progression des scores selon qu'il est administré respectivement à des étudiants en médecine, à des internes et à des praticiens expérimentés, les modalités techniques de construction et d'administration d'un test sont également bien connues [8-11].

Il n'est en revanche encore pas utilisé dans l'évaluation des étudiants en médecine en Tunisie.

## **METHODES**

## II- Méthodes :

### II-1. Population étudiée

Il s'agissait d'étudiants de quatrième année médecine (externes, DCEM2) et des internes de médecine générale de la faculté de médecine de Tunis. La durée de stage en néonatalogie pour les externes était de deux à trois semaines en fonction des groupes et de leur répartition et celle des internes de deux mois. Chaque groupe d'externes était composé de dix à 11 étudiants. Les groupes d'internes étaient composés de cinq à six étudiants.

Les stages se sont déroulés au service de néonatalogie de l'hôpital Charles Nicolle de Tunis, un service hospitalo-universitaire, de niveau III de soins néonataux offrant à côté des soins de réanimation néonatale des soins de pédiatrie de maternité.

Les différentes périodes de stage incluses, des étudiants de DCEM2, étaient :

- Du 29-09-2014 au 11-10-2014 : 10 étudiants.
- Du 13-10-2014 au 01-11-2014 : 11 étudiants.
- Du 03-11-2014 au 16-11-2014 : 10 étudiants.
- Du 17-11-2014 au 07-12-2014 : 11 étudiants.
- Du 08-12-2014 au 19-12-2014 : 10 étudiants.

Durant la période d'étude deux groupes d'internes ont été inclus, respectivement de six et de cinq membres.

### II-2. Rédaction du TCS

Les questions ont été élaborées en prenant en considération la liste des objectifs de stage de quatrième année Médecine (**Annexe 1 : Carnet de stage DCEM2 p:53-54**).

Les objectifs de stages des internes ne comprennent pas une liste spécifique d'objectifs pour la néonatalogie. Ces derniers sont confondus avec ceux de la pédiatrie. Ils rejoignent ceux des externes.

Les objectifs de stage des externes concernent fondamentalement le recueil de l'information clinique par l'interrogatoire et l'examen physique, le choix et l'interprétation des examens complémentaires, l'initiation au diagnostic et au traitement des cas simples de pathologie courante et de situations d'urgence ainsi que les attitudes d'autoformation, de relation avec le patient et avec l'équipe sanitaire.



L'élaboration des TCS a été effectuée selon les recommandations de l'université de Montréal, disponibles sur le site <http://www.cpass.umontreal.ca/index.html>.

Une liste de sujets a été définie, avec pour chacun le ou les domaines de connaissances à évaluer (diagnostique, d'investigation, thérapeutique et pronostique), conformément aux objectifs de stage.

### Tableau 1

Les vignettes composées chacune d'un seul item ont été rédigées par une seule personne (sénior en néonatalogie, maître de conférence agrégé) et ont été soumises à une relecture (sénior en néonatalogie, assistant hospitalo-universitaire).

Chaque vignette comprenait de façon explicite l'échelle de Likert, avec un score (ou ancrage) de -2, -1, 0, 1, et 2. Si l'information donnée rendait l'hypothèse, l'examen initialement prévu, la prescription ou le geste envisagé(e) plus probable ou plus justifié(e), selon le domaine de connaissance à évaluer, le score était de 1 ou 2 ; si elle était sans effet, il était de 0 ; si elle le (la) rendait moins probable ou moins justifiée, le score était de -1 ou -2.

**Tableau 1: Caractéristiques des 20 items du TCS**

Vignette	sujet	Domaine de connaissance à évaluer
1	Ictère hémolytique immunologique	Diagnostique
2	Ictère par hémolyse constitutionnelle	Diagnostique
3	Convulsions néonatales	Investigation
4	Cholestase néonatale	Investigation
5	Cardiopathies congénitales	Diagnostique
6	Réanimation en salle de naissance	Thérapeutique
7	Hyperplasie congénitale des surrénales	Diagnostique
8	Pédiatrie en maternité (Retard d'émission du méconium)	Diagnostique
9	Pédiatrie de maternité (retard à l'émission des urines)	Diagnostique
10	Allaitement maternel (Difficultés-conseils)	Thérapeutique
11	Syndrome hémorragique	Investigation
12	Infection materno-fœtale	Diagnostique
13	Prématurité	Diagnostique
14	Retard de croissance intra-utérin	Diagnostique
15	Détresse respiratoire du nouveau-né	Thérapeutique
16	Urgences chirurgicales néonatales	Diagnostique
17	Détresse respiratoire du nouveau-né	Diagnostique
18	Détresse respiratoire du nouveau-né	Diagnostique
19	Embryofetopathies infectieuses	Diagnostique
20	Infection materno-fœtale/méningite	Diagnostique

### II-3. Validation des items et choix des experts

Une grille de contrôle de la qualité *à priori* a été consultée pour la validation du TCS. Cette grille est disponible sur le site <http://www.cpass.umontreal.ca/index.html>. **Annexe3**

Le test a été soumis à 15 médecins séniors en néonatalogie. Le test était anonyme et ils ont été informés de son but et de son cadre. Dix seulement ont répondu. Le groupe d'experts était ainsi composé de dix séniors exerçant dans quatre services hospitalo-universitaires différents : deux professeurs, deux maitres de conférence agrégés, quatre assistants hospitalo-universitaires et deux médecins spécialistes hospitalo-sanitaires.

Au cours de l'analyse complémentaire, Les ancrages situés à distance de la réponse modale ont été considérés comme erronés. Si l'erreur était liée à un seul expert, sa réponse n'a pas été prise en compte et si elle était faite par plusieurs experts, l'ensemble de l'item n'a pas été pris en compte dans la notation. Les items où la distribution des réponses n'était pas centrée sur une partie de l'échelle de Likert, n'ont pas été également pris en compte.

### II-4. Établissement des scores : méthode des scores combinés

L'établissement des scores s'est basé sur la méthode des scores combinés qui utilise les réponses des experts recrutés pour construire un barème de scores, où chaque réponse des experts possède une valeur intrinsèque et reçoit une pondération en fonction du nombre d'experts qui l'ont choisie. Plus précisément, la réponse à un item la plus souvent choisie par les experts (la réponse modale) vaut un point ; les autres réponses des experts reflètent la variabilité dans l'interprétation des données et donnent droit à une fraction d'un point, en fonction du nombre d'experts qui les ont choisies. Par exemple à l'un des items, 7 experts répondent « -1 », 2 répondent « -2 », 1 répond « 0 » et aucune ne répond « 1 » ou « -2 ». Pour obtenir les scores, il s'agit de repérer la réponse qui a été la plus choisie par les experts (dans ce cas-ci, la réponse « -1 » a été choisie par 7 experts). Cette réponse donnera maintenant un point à toutes les personnes évaluées qui la choisiront. Ensuite, il faut diviser toutes les réponses par le nombre d'experts le plus élevé à avoir choisi la réponse. Dans cet exemple, il faut diviser par 7. Les réponses non choisies par les experts donnent 0 point et les autres réponses donnent un crédit en fonction du nombre d'experts qui les ont choisies, soit  $2/7 = 0,28$  point pour la réponse « -2 » et  $1/7 = 0,14$  pour la réponse « 0 ». Tableau 3

Le score global a été obtenu en ajoutant les scores de chaque question, puis les transformant en un score combiné sur 100.

**Tableau 2 : Méthode des scores combinés. Exemple de réponse à un item par dix experts.**

	-2	-1	0	1	2	Instructions
<b>Nombre de réponses parmi les experts</b>	2	7	1	0	0	Repérer la réponse modale la plus choisie)
<b>Création de scores</b>	2/7	7/7	1/7	0/7	0/7	Diviser par le nombre d'experts ayant donné la réponse la plus choisie
<b>Crédit pour l'item</b>	0.28	1	0.14	0	0	Donner les points aux étudiants examinés pour cet item

## II-5. La réalisation du test

Les différents groupes d'externes et d'internes présents au cours de la période d'étude ont été évalués par le TCS au dernier jour de leur stage. Le déroulement du test a eu lieu dans des conditions identiques, en salle de staff pendant 20 à 30 minutes. L'évaluation était anonyme et les étudiants étaient informés de son but. Un paragraphe expliquant les principes des TCS était placé en préambule.

## II-6. Analyse statistique

Les résultats ont été exprimés en moyennes±écart-type. Les comparaisons de plusieurs moyennes sur séries indépendantes ont été effectuées au moyen du test F de Snedecor d'analyse de la variance paramétrique (ANOVA à un facteur). En cas de différence significative, les comparaisons 2 à 2 ont été faites par la méthode de Bonferroni. La normalité des distributions a été étudiée par le test de Kolmogorov et Smirnovet graphiquement. L'égalité des variances par le test de Levene.

La cohérence interne (fidélité) du TCS a été évaluée par le coefficient  $\alpha$  de Cronbach à l'aide d'une feuille de calcul Excel téléchargeable sur le site de l'université de Montréal <http://www.cpass.umontreal.ca/>.

# RESULTATS

### III. Résultats

#### III-1 Résultats globaux des experts, externes et internes

Les réponses du groupe d'experts sont détaillées sur le **tableau3**. Les différents scores attribués sont illustrés sur l'**Annexe 4**.

Les **experts** ont eu un score moyen, rapporté sur 100, de **82.52 ± 7.35 IC<sub>95%</sub> [77.26-87.78]**, les **externes de 58.52 ± 9.72 IC<sub>95%</sub> [55.82-61.23]**, (**p<0.0001**) ; et les **internes de 63.17±11.36 IC<sub>95%</sub> [55.53-70.81]** (**p<0.0001**). Les coefficients  $\alpha$  de Cronbach pour les réponses des externes et des internes respectifs étaient de 0.68 et de 0.61. La différence des scores entre étudiants et internes n'était pas statistiquement significative,  $p=0.46$ . **Tableau 4 et Figure 1**

**Tableau 3: Réponses des experts au test de concordance de script**

Echelle Likert	Vignette 1	Vignette 2	Vignette* 3	Vignette* 4	Vignette 5	Vignette 6	Vignette 7	Vignette 8	Vignette 9	Vignette* 10
-2	0	0	0	1	2	9	0	8	6	0
-1	0	0	1	6	7	1	0	2	2	1
0	4	0	1	2	1	0	0	0	2	6
1	4	5	3	1	0	0	5	0	0	0
2	2	5	5	0	0	0	5	0	0	3
Echelle Likert	Vignette* 11	Vignette 12	Vignette 13	Vignette 14	Vignette 15	Vignette 16	Vignette 17	Vignette 18	Vignette 19	Vignette 20
-2	1	0	4	7	0	0	0	3	1	1**
-1	4	3	6	3	7	0	1	2	5	0
0	3	7	0	0	2	10	8	5	4	1
1	0	0	0	0	0	0	0	0	0	2
2	2	0	0	0	1**	0	1**	0	0	6

\*Items pouvant être considérées comme erronées ou discutables  
\*\* ancrages pouvant être considérés comme erronés ou discutables

**Tableau 4 : Répartition des scores Experts/Etudiants/Internes.**

Scores	Experts N=10	Externes N=52	Internes N= 11
Moyenne	82.52	58.52	63.17
IC <sub>95%</sub>	[77.26-87.78]	[55.82-61.23]	[55.53-70.81]
Médiane	80.67	59.72	64.15
Variance	54.11	94.55	129.25
Ecart type	7.35	9.72	11.36
Minimum	72.75	35.75	45.65
Maximum	95.05	77.8	83.15

<b>Coefficient <math>\alpha</math> de Cronbach</b>	0.68	0.61
--	------	------

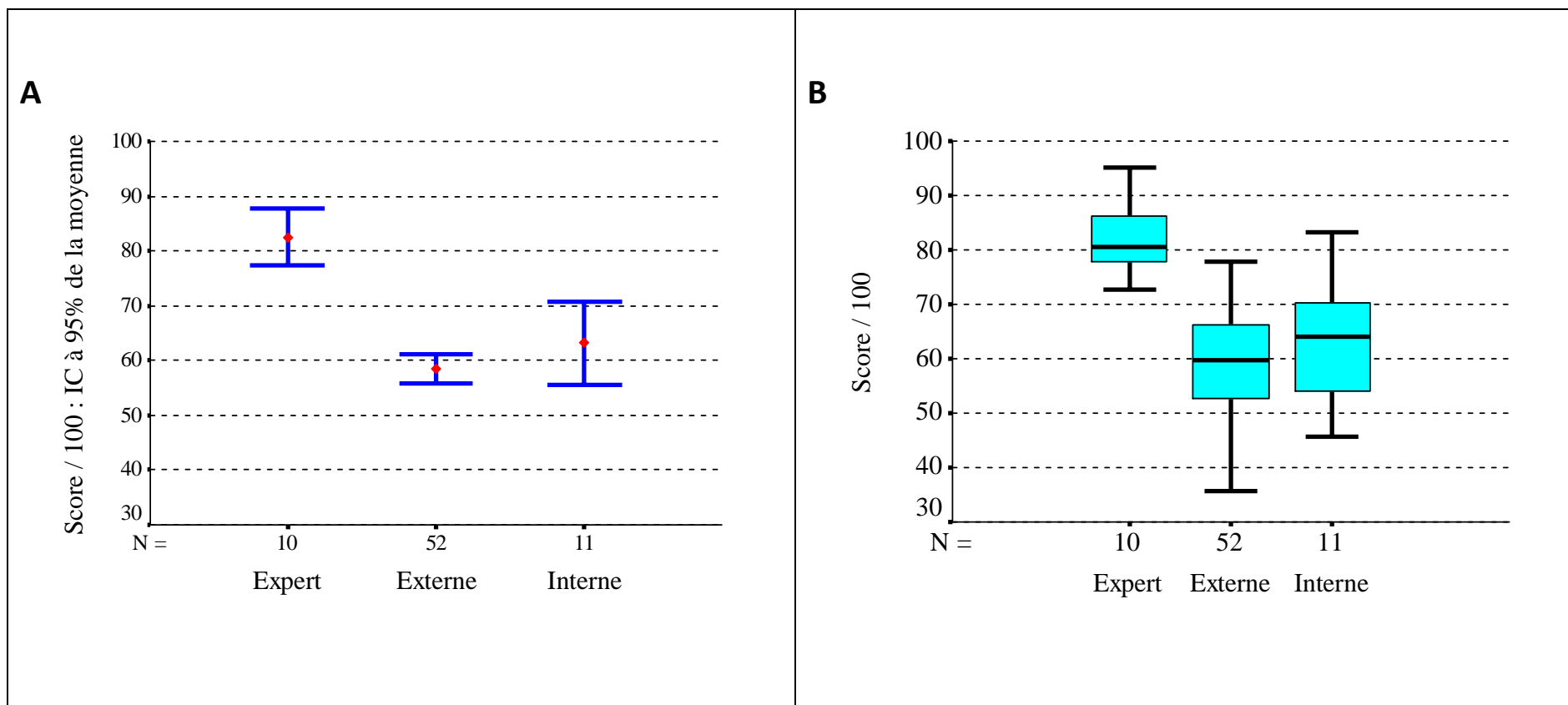


Figure 1: Résultats globaux du test de concordance de script exprimés en scores moyens avec barres d'erreurs (A) et en médianes et quartiles (B) pour les trois groupes : Expert, Externe et interne.

### III-2 Résultats des experts, externes et internes après élimination des items et ancrages erronés ou discutables.

Trois ancrages et quatre items ont été considérés comme erronés ou discutables et ont été retirés pour une analyse complémentaire. **Tableau 3**

La suppression des ancrages et des items erronés ou discutables a conduit aux résultats suivants : **Experts 84,49± 7,88 IC<sub>95%</sub> [78.91-90.19], externes 57,46±10,95 IC<sub>95%</sub> [54.41-61,51](p<0.0001), et internes 64,95±11,59 ](p<0.0001).** Les coefficients  $\alpha$  de Cronbach pour les réponses des externes et des internes respectifs étaient de 0.96 et de 0.99. La différence entre étudiants et internes n'était pas statistiquement significative, p=0.11. **Tableau 5 et Figure 2**

**Tableau 5 : Répartition des scores Experts/Étudiants/Internes après suppression des réponses et des items erronés ou discutables**

Scores	Experts N=10	Externes N=52	Internes N= 11
Moyenne	84,49	57,46	64,95
IC <sub>95%</sub>	[78,91 -90,19]	[54,41-61,51]	[57,16-72,74]
Médiane	84,49	59,09	61,75
Variance	62,18	119,99	134,48
Ecart type	7,88	10,95	11,59
Minimum	73,43	25,94	46,13
Maximum	97,87	73,81	85,50
Coefficient $\alpha$ de Cronbach		0.96	0.99

La dispersion des scores individuels dans chacun des groupes avant et après exclusion des items erronés ou discutables est représentée sur **la figure 3**.



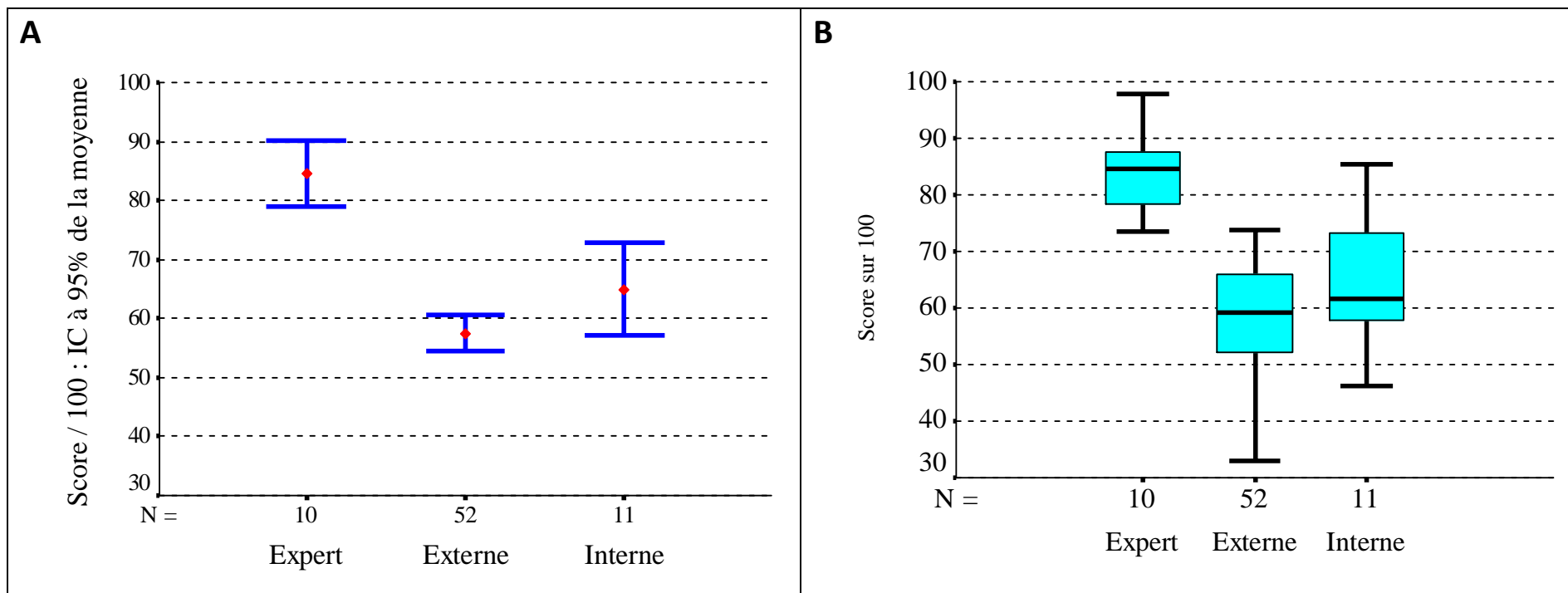


Figure 2 : Résultats du test de concordance de script après exclusion des items et ancrages erronés ou discutables, exprimés en scores moyens avec barres d'erreurs (A) et en médianes et quartiles (B) pour les trois groupes : Expert, Externe et interne.

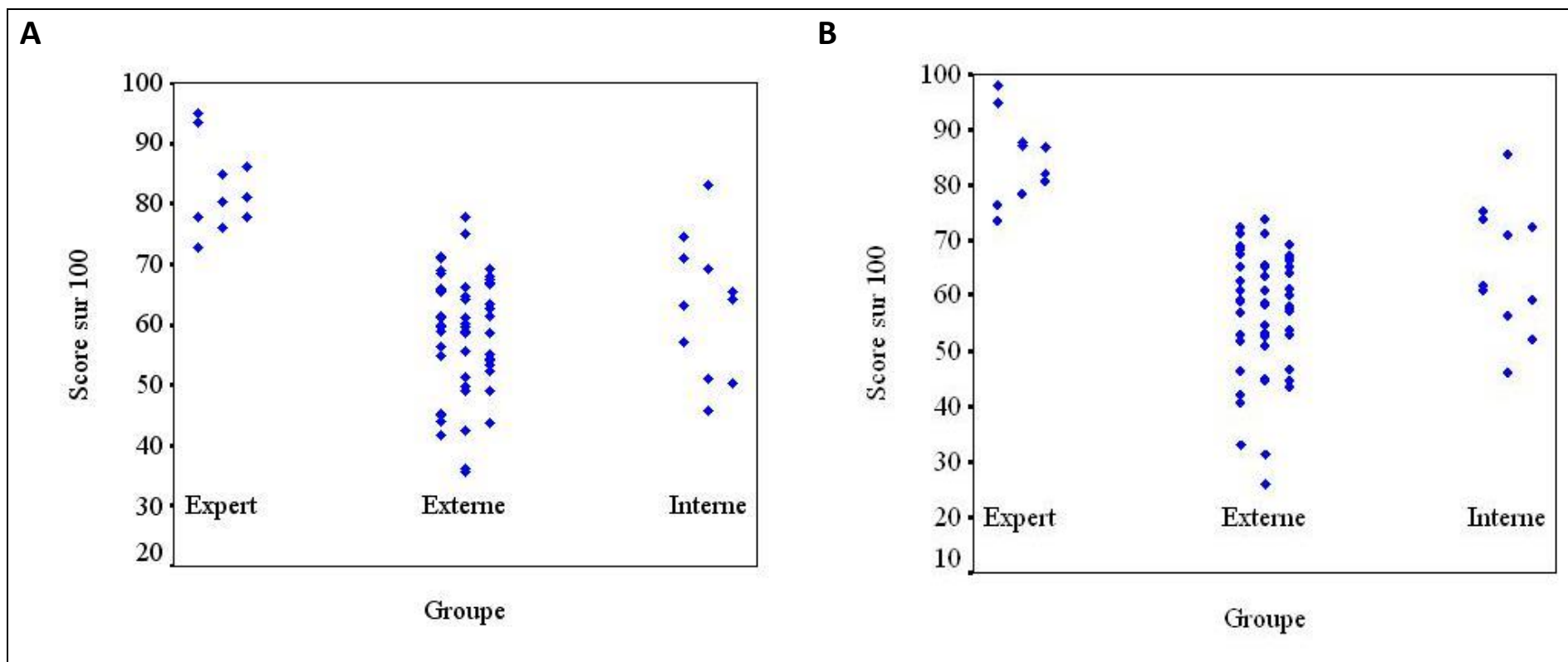


Figure3 : Dispersion des scores individuels dans chacun des groupes avant (A) et après exclusion (B) des items erronés ou discutables

# DISCUSSION

## **IV. Discussion**

### **IV-1 Evaluation des scores moyens**

Le TCS permet de différencier des groupes d'étudiants de niveaux de compétence différents. Il évalue non seulement la maîtrise des seules connaissances déclaratives mais également le processus de leur mobilisation dans le cadre du raisonnement clinique. Ceci a été démontré dans différentes spécialités [12-15].

Dans notre étude, nous avons retrouvé des différences statistiquement significatives entre experts et externes et entre experts et internes, ce qui était prévisible. Les scores moyens étaient plus élevés chez les internes que chez les externes mais cette différence n'était pas statistiquement significative. En effet, dans notre terrain de stage, les internes, ne bénéficiant que de deux mois de stage, ne diffèrent pas beaucoup des externes. Seule une année d'étude les sépare. Au niveau de l'internat va s'ajouter le sens de la responsabilité professionnelle impliquée par la prise en charge des patients. L'enseignement et l'encadrement prodigués en terrain de stage sont plus importants et plus réguliers pour les externes. Ceci pourrait expliquer l'absence de différence statistiquement significative au niveau de leurs scores moyens. Les résidents de spécialité n'ont pas été testés dans ce travail.

L'organisation des connaissances adaptées aux tâches cliniques se développe progressivement avec l'expérience acquise au cours de la formation et donc très tôt au cours de leurs stages hospitaliers. Squalli Houssaini et al. avaient évalué la pertinence d'un TCS dans l'évaluation des compétences en Néphrologie des étudiants du deuxième cycle de la faculté de Médecine de Fès. Ils ont comparé les résultats au TCS des étudiants ayant effectué un stage de néphrologie (groupe 1) à ceux n'ayant pas effectué un tel stage (groupe 2). Ces résultats ont aussi été comparés à ceux d'un examen de connaissances de néphrologie préalable aux stages. Les notes à l'examen de connaissances en néphrologie étaient comparables. Après le stage de néphrologie, les résultats du groupe 1 étaient significativement meilleurs. Cette différence était retrouvée aussi bien pour les items évaluant le diagnostic que pour les items évaluant la thérapeutique. Il n'y avait pas de corrélation entre les notes de l'examen de connaissances de 4<sup>ème</sup> année et celles du TCS, confirmant de façon objective le rôle bénéfique du stage hospitalier de néphrologie dans l'acquisition des compétences cliniques dans cette spécialité [16].

Des données comparables ont déjà été validées dans d'autres spécialités à différents niveaux de formation dont de DCEM2 [4,17,18].

La répétition du TCS par un même étudiant en début, en cours et en fin de stage permettrait d'évaluer sa progression ou de dépister des lacunes dans sa formation. Lorsque qu'il est administré en début de stage, le TCS permettrait de distinguer les étudiants les moins à l'aise en terme de

raisonnement clinique et appliqué en fin de stage, il pourrait constituer un moyen d'évaluer la progression [4].

Le TCS peut également constituer un moyen d'évaluation des méthodes d'apprentissage. Compère a étudié l'impact d'un enseignement par "groupe de pairs", tutoré par un senior et prodigué de façon hebdomadaire à un groupe d'internes en anesthésie dès le début de leur stage, par rapport à un groupe témoin bénéficiant d'un encadrement classique (périodes de stages distinctes), sur la performance à raisonner en situations d'incertitude. Les outils de mesure comprenaient en premier lieu un TCS. Les scores en fin de stage étaient significativement meilleurs dans le groupe de l'enseignement par "groupe de pairs" [19].

Des scores significativement plus bas chez les externes pourraient faire discuter la pertinence même de l'utilisation du TCS chez des étudiants de deuxième cycle, dans la mesure où qu'à leur niveau, il faut commencer par leur transmettre des certitudes, leur permettant d'acquérir des représentations mentales correspondant à la fois à des réalités et à des repères. Certes, dans la « vraie vie », au cours de leurs stages hospitaliers, ils vont être confrontés inéluctablement à des situations d'incertitudes, mais encore faut-il qu'ils apprennent à raisonner devant des situations typiques imposant un diagnostic de certitude.

Rédiger un TCS pour explorer le raisonnement clinique devant une situation typique revient à rédiger des QCM d'évaluation qualitative dont la notation dépend d'une norme et non de la variabilité d'un groupe d'experts. En effet, les QCM explorent bien le domaine du savoir théorique à ses différents niveaux. Dans ce domaine cognitif, la taxonomie de BLOOM hiérarchise six niveaux d'objectifs pédagogiques allant du simple au complexe : connaissance, compréhension, application, analyse, synthèse et évaluation. Les QCM évaluatrices de données permettraient d'évaluer le raisonnement lui-même [20]. L'originalité des TCS tient à ce que les réponses des étudiants sont comparées aux différentes modalités de réponses d'experts et non à une norme. Ces deux méthodes d'évaluation ont en outre un intérêt essentiel qui est l'automatisation de la correction, aux questionnements ouverts, écrits ou oraux avec des réponses spontanées des étudiants.

#### **IV-2. Qualités psychométriques du TCS**

Un test est dit fiable s'il donne avec constance un même résultat et s'il mesure de façon précise avec le moins d'erreur d'estimation possible. Le coefficient  $\alpha$  de Cronbach, un des indices de fidélité, mesure la cohérence interne du test, c'est-à-dire jusqu'à quel point les items contribuent à la mesure d'une même dimension dans le test. Il est communément admis qu'un test est fiable lorsque ce coefficient a une valeur supérieure à 0,80. Essentiellement, la fidélité d'un test dépend du nombre

d'items contenus dans le test. Pour optimiser au mieux le test, il faut augmenter le nombre de questions. Les tests comprenant 20 vignettes et 60 items, passés en une heure de temps, atteignent un coefficient de Cronbach  $>0.75$  [3, 8, 10,21].

Notre travail, malgré l'absence d'expérience antérieure, est à notre connaissance le premier test élaboré en néonatalogie et à l'échelle de notre faculté. Le caractère nouveau du test, aussi bien pour les étudiants que pour les enseignants était la principale raison ayant conduit à un nombre limité de vignettes avec des unités d'évaluations simples, comprenant un seul item pour chaque vignette.

Malgré ce nombre limité d'items, le coefficient de Cronbach était déjà satisfaisant à l'analyse globale, 0.68 pour les réponses d'externes et 0.61 pour celles des internes. Ces coefficients ont augmenté après exclusion des items jugés erronés ou discutables à variabilité très élevée entre experts. L'exclusion des items à haute variabilité enlèverait du test une part importante d'incertitude. Une variabilité élevée témoigne de la difficulté liée à l'incertitude, et la variabilité du test augmente la discrimination, effet recherché dans un TCS, mais quand la variabilité devient extrême ceci pourrait témoigner plutôt d'erreur [21]. Cette erreur peut émaner du panel d'experts ou alors être au niveau de l'élaboration même de la vignette clinique.

Dans un TCS destiné à des internes exerçant dans un service d'accueil des urgences pédiatriques, on avait fait passer un test comprenant 60 questions réparties en 38 vignettes mais les auteurs n'avaient retenu que 50 questions dans 30 vignettes, optimisant ainsi le score alpha de Cronbach à 0,77 [22].

Si on crée un seul item par vignette l'effort de réflexion est considérable pour les étudiants qui doivent se plonger dans une grande série de cas clinique [8]. La durée de notre test (30 mn) était suffisante pour répondre à 20 questions dans 20 vignettes cliniques. Les différents coefficients de Cronbach calculés étaient satisfaisants même après élimination de quatre vignettes jugées erronées, suggérant que c'est la qualité de la vignette elle-même et la variabilité des réponses qu'elle génère qui interviennent plutôt que le nombre seul.

L'élaboration des vignettes dans un contexte d'incertitude est une tâche difficile surtout quand on n'est pas formé et habitué à cet outil. Dans notre test, les vignettes étaient rédigées par un seul expert et validé par un autre, tous les deux sans expérience antérieure.

Il est préférable que trois experts participent à l'élaboration d'un TCS afin de multiplier les expériences et les jugements pour obtenir des vignettes et des questions optimales pour ce test, sans ancrage anormal et avec un certain degré de dispersion mais limité autour de la réponse modale, donc à variabilité correcte[10].

Les tests psychométriques montrent des résultats meilleurs à l'échelle individuelle pour les items à variabilité modérée, tandis que les items à grande variabilité discriminent mieux l'expérience clinique à l'échelle collective [21]. Ces constatations expliquent en partie l'absence d'amélioration des scores de nos externes et internes après élimination de certains items considérés erronés ou discutables, mais d'un autre côté l'amélioration des coefficients  $\alpha$  de Cronbach. En effet, les items à faible variabilité explorent les connaissances cliniques formelles plutôt que le raisonnement face à des données cliniques complexes et ambiguës. Le panel d'expert peut posséder plus de facilité à raisonner dans un contexte d'incertitude que les juniors, mais pas forcément devant des connaissances formelles [21].

Charlin, en étudiant l'effet de la variabilité au sein des réponses du panel de référence, a démontré que la variabilité des réponses était un élément clé du pouvoir discriminant du TCS. Il recommandait ainsi de construire délibérément des items pouvant induire une variabilité au sein du panel d'experts. Cependant, l'existence d'items à faible variabilité est nécessaire. Ces items évaluent les connaissances de problèmes bien définis et bien établis. Or ce type d'items s'apparente à une QCM à contexte riche sauf que le format et la tâche requise aux étudiants diffèrent [21].

Un nombre suffisant d'experts est nécessaire pour optimiser au mieux le test. Un nombre de 15 experts est considéré suffisant pour un examen à forts enjeux et un nombre de dix pour un examen formatif [9]. Dans notre travail, certains enseignants experts n'étaient pas réceptifs et n'ont pas répondu au test, d'autres l'ont fait mais en disant qu'il était « compliqué » ou même « non pédagogique ». En contre partie, même si nous ne l'avons pas évalué par écrit, les étudiants ont adhéré à ce test. Le caractère anonyme et non sanctionnable les a motivés. Ils étaient réceptifs et ils l'avaient jugé proche de la réalité professionnelle.

#### **V- Perspectives :**

- Veiller à améliorer la validité du test par une meilleure rédaction, une meilleure validation des items et par un choix optimal du panel de référence.
- La poursuite des recherches sur le TCS pour les résidents de spécialité mais aussi les autres professionnels de santé particulièrement les infirmiers spécialisés en soins néonataux et pédiatriques.
- Promotion du TCS auprès des responsables de formation d'autres spécialités médicales.
- Création de banques d'items par disciplines médicales.
- Développer des activités d'apprentissage (ateliers et travaux de groupes) au travers desquels l'étudiant pourra se confronter à ses pairs ou à des experts professionnels dans la résolution de problèmes cliniques et prendre conscience de ses propres processus de raisonnement clinique.

Le recours au TCS à côté des autres méthodes d'évaluation telles les QROC ou les QCM à contexte riche dans l'évaluation de ces méthodes d'apprentissage.

- La répétition du TCS par un même étudiant en début, en cours et en fin de stage afin d'évaluer sa progression ou de dépister des lacunes dans sa formation.
- Mise en place d'une évaluation en ligne par TCS pour les étudiants et les professionnels.



## **CONCLUSION**

## **VI- Conclusion :**

Notre travail a démontré la faisabilité du TCS en stage de néonatalogie. Cependant il nous a permis de reconnaître les difficultés de son élaboration, au niveau de la rédaction des vignettes et des items et de leur validation et au niveau du choix des experts et des conséquences qui en découlent. Ce test a permis de différencier les experts des internes et des étudiants. Les internes avaient des scores supérieurs aux externes sans pour autant que la différence ne soit statistiquement significative. Ce résultat attendu est en accord avec les différentes études réalisées dans diverses autres spécialités médicales ou chirurgicales.

L'élaboration du TCS impose de constituer un groupe d'enseignants formés et habitués progressivement à ce test pour respecter les recommandations actuelles. Il faut probablement deux à trois enseignants pour la rédaction des vignettes et items et un groupe d'experts de 10 à 15 pour constituer le panel de référence.

Le test de concordance de script ne remplace pas les autres méthodes d'évaluation telles les QROC ou les QCM à contexte riche, il complète ces stratégies d'évaluation par l'exploration du raisonnement en contexte d'incertitude. Sa présentation permet d'explorer une facette du raisonnement habituellement exclue des évaluations traditionnelles mais fréquemment rencontrées en pratique : raisonner devant un problème mal défini, des données incertaines avec plusieurs hypothèses ou solutions possibles. La diffusion de son utilisation pourrait contribuer à optimiser les activités d'apprentissage et inciter les étudiants au raisonnement au lit du malade.

La répétition du TCS par un même étudiant en début et en cours de stage permettra d'évaluer sa progression ou de dépister des lacunes dans sa formation. Et enfin le recours au TCS dans l'évaluation des méthodes d'apprentissage augmenterait davantage sa pertinence.

Le recours au TCS serait plus pertinent pour l'évaluation des internes et des résidents de spécialité confrontés à des situations parfois complexes et incertaines. Il l'est moins à un niveau d'études où l'étudiant doit encore acquérir des repères avant de manier l'incertitude.

## REFERENCES

## Références

1. Charlin B, Brailovsk y, CA, Brazeau-Lamontagne, L Samson, L Leduc, C. Script questionnaires: Their use for Assessment of Diagnostic Knowledge in radiology. *Medical Teacher* 1998 : 20 : 567-571.
2. Charlin B, Roy L, Brailovsky C, et al. The Script Concordance test: a tool to assess the reflective clinician. *Teach Learn Med* 2000; 12:189–95.
3. Charlin B, VleutenC Van Der. Standardized assessment of reasoning in context of uncertainty. The script concordance test approach. *Evaluation and the health professions* 2004, 27:304-3019.
4. Gibot S, Boulaert PE. Le test de concordance de script outil d'évaluation formative en reanimation médicale. *Pédagogie médicale* 2008 ; 9 :7-18.
5. Schmidt H.G., Norman G.R, Boshuizen, HP A . A Cognitive perspective on medical expertise: theory and Implications. *Acad Med* 1990; 65: 611-621.
6. Charlin B, Tardif J, Boshuize HPA. Scripts and Medical Diagnostic Knowledge: theory and applications for clinical reasoning instructions and resarch. *Acad Med* 2000; 75 : 182-90.
7. Lubarsky S, Charlin B, Cook DA, Chalk C, van der Vleuten CPM. Script concordance method: a review of published validity evidence. *Med Educ* 2011;45:329-38
8. Fournier J-P, Demeester A, Charlin B. Script Concordance Tests: Guidelines for Construction. *BMC Med inform and Decis Mak* 2008; 8:18.
9. Gagnon R, Charlin B, Coletti M, Sauvé E, Van der Vleuten C. Assessment in context of uncertainty: How many members are needed on the panel of reference of a script concordance test? *Med Educ* 2005;39:204-91.
10. Gagnon R, Charlin B, Lambert C, Carrière B, der Vleuten V. Script Concordance Testing: More Cases or More Questions? *Adv Health Sci Educ* 2009; 14:1382-96.
11. Charlin B, Gagnon R, Lubarsky S, Lambert C, Meterissian S, Chalk C, Goudreaut J, van der Vleuten C. Assessment in the context of uncertainty using the script concordance test: More meaning for scores. *Teach Learn Med* 2010; 22:180-6.
12. Marie I, Sibert L, Roussel F, Hellot MF, Lechevallier J, Weber J. Le test de concordance de script : un nouvel outil d'évaluation du raisonnement et de la compétence clinique en médecine interne. *Rev Med Int* 2005; 26:501-7. 9.
13. Brazeau-Lamontagne L, Charlin B, Gagnon R, Samson L, van der Vleuten C. Measurement of perception and interpretation skills along radiology training: utility of the script concordance approach. *Med Teach* 2004; 26:326-32. 10.

14. Caire F, Sol JC, Charlin B, Isidori P, Moreau JJ. Le test de concordance de script (TCS) comme outil d'évaluation formative des internes en neurochirurgie: implantation du test sur Internet à l'échelle nationale. *Pédagogie Médicale* 2004;5:87-94.
15. Talvarda M., Olivesa J.-P., Masa E. Evaluation des étudiants en médecine lors de leur stage en gastro-entérologie pédiatrique par un test de concordance de script. *Archives de Pédiatrie* 2014;21:372-376.
16. Sqalli Houssaini T, Bono W, Tachfouti N, Maillard D. pertinence d'un test de concordance de script dans l'évaluation des compétences en néphrologie des étudiants du deuxième cycle de la faculté de Médecine de Fès. *Les annales de médecine et de thérapeutique AMETHER.* 2009,1 :4-10.
17. Sibert L, Charlin B, Gagnon R, Corcos J, Khalaf A, Grise P. Evaluation du raisonnement clinique en urologie : l'apport du Test de Concordance de Script. *Progrès en Urologie* 2001 ;11 :1213-1219.
18. Caire F, Marin B, Cuny E. Utilisation du test de concordance de script au cours du 2<sup>ème</sup> cycle des études médicales. Expérience dans l'enseignement de la neuro-chirurgie. *Pédagogie médicale.* 2011 ;12(1) :29-35.
19. Compère V, Moriceau J, Gouin A et al. Residents in tutored practice exchange groups have better medical reasoning as measured by the script concordance test: A pilot study. *Anaesth Crit Care Pain Med* 2015, *sous presse.*
20. Bloom BS. Taxonomie des objectifs pédagogiques. Tome 1 domaine cognitif. Québec. Les presses de l'université du Québec, 1975.
21. Charlin B, Gagnon R, Pelletier J and al. Assessment of clinical reasoning in the context of uncertainty: the effect of variability within the reference panel. *Medical Education* 2006; 40: 848–854.
22. Carriere B, Gagnon R, Charlin B, et al. Assessing clinical reasoning in pediatric emergency medicine: validity evidence for a Script Concordance Test. *Ann Emerg Med* 2009;53:647–52.

# ANNEXES

## Annexe 1 : liste des objectifs de stage en Néonatalogie

### OBJECTIFS ET EVALUATION DU STAGE DE NEONATOLOGIE

OBJECTIFS <sup>1</sup>	Enseigné <sup>2</sup>	Évalué <sup>3</sup>	
		Note	Signature
<b>Objectifs relatifs à l'habileté technique :</b>			
1.Procéder à une anamnèse objective, complète et précise en se référant au formulaire ou au protocole d'interrogatoire en usage dans le service afin d'évaluer le risque périnatal du nouveau-né, par l'interrogatoire de la mère et la consultation du dossier obstétrical.			
2.Pratiquer les techniques d'examen physique, identifier les caractéristiques sémiologiques normales et pathologiques :			
2.1.Evaluer le score d'Apgar à la naissance ;			
2.2.Evaluer le degré de maturation du nouveau-né et de sa croissance in utéro d'après les données chronologiques, cliniques et les courbes de référence de croissance intra-utérine ;			
2.3.Apprécier par l'inspection l'état clinique du nouveau-né et pratiquer un examen somatique détaillé ;			
2.4.Inspector les organes génitaux externes et assigner un sexe au nouveau-né ;			
2.5.Dépister par les manœuvres appropriées :			
-une atrésie de l'œsophage par l'épreuve de la seringue ;			
-une perméabilité des choanes ;			
-une imperforation anale ;			
-une luxation congénitale de la hanche ;			
-une diminution voire une abolition des pouls fémoraux.			
3.Consigner les données cliniques sur le carnet de santé et sur le cahier d'observation.			
4. Interpréter les examens paracliniques suivants :			
- Radiographie thoraco-abdominale ;			
- Gaz du sang ;			
- Bilirubinémie ;			
- Numération formule sanguine ;			
- Marqueurs de l'inflammation (C-Reactive-Proteine) ;			

<sup>1</sup> A réaliser en situation réelle ou simulée (R ou S).

<sup>2</sup> C'est-à-dire, a fait l'objet d'une démonstration pratique (colonne à remplir par l'étudiant). Inscrire R ou S (ou 0 si pas de démonstration).

<sup>3</sup> A fait l'objet d'un ou plusieurs contrôle(s) dont seul le dernier a donné lieu à la note définitive inscrite et attestée par une signature. Inscrire NE si non évalué.

OBJECTIFS <sup>1</sup>	Enseigné <sup>2</sup>	Évalué <sup>3</sup>	
		Note	Signature
- Glycémie ;			
- Calcémie.			
<b>5. Pratiquer les soins suivants :</b>			
-Soins de l'ombilic ;			
-Vaccination au BCG ;			
-Injection de vitamine K ;			
-Glycémie au talon (GAT) ;			
-Préparer le matériel nécessaire pour mener correctement la réanimation d'un nouveau-né à la naissance en tenant compte de son âge gestationnel et de son poids.			
-Participer à la réanimation du nouveau-né en salle de naissance, selon le cercle Evaluation / Décision / Application.			
<b>Objectifs relatifs à la solution des problèmes de santé<sup>4</sup> :</b>			
<b>6.Suspecter ou reconnaître :</b>			<b>Epreuve spécifique. Cf. page 8 paragraphe 5.2.2.</b>
- les cas de pathologie courante ;			
- les urgences les plus fréquentes.			
<b>7.Proposer la prise en charge d'un nouveau-né dans les situations normales ou pathologiques courantes :</b>			
- traitement curatif ;			
- traitement préventif ;			
- mesures éducatives : .Prodiguer à une mère venant d'accoucher les conseils nécessaires à la réussite de l'allaitement maternel. .Prodiguer à une mère les conseils utiles à la prise en charge de son enfant à domicile. .Expliquer à une mère une ordonnance de sortie d'un nouveau-né normal.			
<b>Objectifs relatifs aux attitudes :</b>			
<b>8.Manifester les attitudes décrites dans le chapitre 2.2. de la page 5.</b>			

**PATHOLOGIES COURANTES / URGENTES :**

- Prématurité.
- Infection néonatale.
- Ictère du Nouveau-né.
- Détresse respiratoire.
- Convulsions néonatales.
- Asphyxie périnatale.
- Troubles métaboliques (hypoglycémie, hypocalcémie).
- Retard de croissance intra-utérin.
- Anomalies de la différenciation sexuelle.

<sup>4</sup> Voir liste ci-après présentée à titre indicatif sous réserve de modification par le service où se déroule le stage.  
DCEM2 (54)



## Annexe 2 : TCS

Les questionnaires auxquels vous allez répondre sont de 3 types : - questionnaire diagnostique - questionnaire d'investigation - questionnaire thérapeutique Le format des questions et des grilles de réponses vous sont présentés ci-dessous :

### Questionnaire diagnostique

Si vous pensez à une :	Et que l'on vous fournit :	L 'hypothèse diagnostique initiale est :
Hypothèse diagnostique	Une nouvelle information clinique	-2-1012

- 2 L'hypothèse initiale est pratiquement éliminée
- 1 L'hypothèse initiale est moins probable
- 0 L'hypothèse initiale n'est pas influencée par la donnée supplémentaire
- 1 L'hypothèse initiale est plus probable
- 2 L'hypothèse initiale est pratiquement certaine

### Questionnaire d'investigation

Si vous pensez faire :	Et que l'on vous fournit :	L 'examen initialement prévu est :
Un examen para clinique	Une nouvelle information clinique ou le résultat d'un examen complémentaire	-2-1012

- 2 L'examen initialement prévu est absolument contre-indiqué
- 1 L'examen initialement prévu est moins justifié
- 0 L'examen initialement prévu n'est pas influencé par la donnée supplémentaire
- 1 L'examen initialement prévu est plus justifié
- 2 L'examen initialement prévu est indispensable

### Questionnaire thérapeutique

Si vous pensez prescrire ou faire :	Et que l'on vous fournit :	La prescription ou le geste envisagé(e) initialement est :
Un traitement	Une nouvelle information clinique ou le résultat d'un examen complémentaire	-2-1012

- 2 La prescription ou le geste envisagé(e) est absolument contre-indiqué
- 1 La prescription ou le geste envisagé(e) est moins justifié
- 0 La prescription ou le geste envisagé(e) n'est pas influencé par la donnée supplémentaire
- 1 La prescription ou le geste envisagé(e) est plus justifié
- 2 La prescription ou le geste envisagé(e) est indispensable

### TCS Vignette Clinique N°1

Aziz, nouveau-né à terme eutrophique, est admis à la 12<sup>ème</sup> heure de vie pour prise en charge d'un ictère précoce à bilirubine non conjuguée. Il est issu d'une mère primipare, primigeste de GS O<sup>-</sup> négatif. Le bilan systématique au sang du cordon avait montré un GS du nouveau-né A<sup>+</sup> positif et un Test de Coombs direct positif.

Si vous pensez à :	Et que l'on vous fournit :	L 'hypothèse diagnostique initiale est :
Ictère hémolytique par incompatibilité OA	Une recherche d'agglutinines irrégulières (RAI) négative au cours du premier trimestre de la grossesse	-2 -1 0 1 2

- 2 L'hypothèse initiale est pratiquement éliminée
- 1 L'hypothèse initiale est moins probable
- 0 L'hypothèse initiale n'est pas influencée par la donnée supplémentaire
- 1 L'hypothèse initiale est plus probable
- 2 L'hypothèse initiale est pratiquement certaine

### TCS Vignette Clinique N°2

Lina, nouveau-né à terme eutrophique, née par voie basse sans dystocie. Elle est admise à la 16<sup>ème</sup> heure de vie pour prise en charge d'un ictère précoce à bilirubine non conjuguée avec une anémie régénérative. Elle est issue d'une mère G1P1, de GS A<sup>+</sup> positif. Le GS du nouveau-né O<sup>+</sup> positif et le Test de Coombs Direct est négatif.

Si vous pensez à :	Et que l'on vous fournit :	L 'hypothèse diagnostique initiale est :
Ictère par hémolyse constitutionnelle	Un antécédent de splénectomie chez la mère	-2 -1 0 1 2

- 2 L'hypothèse initiale est pratiquement éliminée
- 1 L'hypothèse initiale est moins probable
- 0 L'hypothèse initiale n'est pas influencée par la donnée supplémentaire
- 1 L'hypothèse initiale est plus probable
- 2 L'hypothèse initiale est pratiquement certaine

### TCS Vignette Clinique N°3

Rayene, nouveau-né à terme eutrophique, est admis en unité de soins intensifs devant des convulsions à type de clonies du membre supérieur droit, notées à H10 de vie. Il est né à terme, l'accouchement était laborieux avec l'application de forceps pour défaut d'expulsion. Le score d'Apgar était à 2-7-8 à 1-3 et 5 mn de vie.

Si vous pensez faire :	Et que l'on vous fournit :	L'examen initialement prévu est :
Une tomодensitométrie cérébrale	Une échographie transfontanellaire normale	-2 -1 0 1 2

- 2 L'examen initialement prévu est absolument contre-indiqué
- 1 L'examen initialement prévu est moins justifié
- 0 L'examen initialement prévu n'est pas influencé par la donnée supplémentaire
- 1 L'examen initialement prévu est plus justifié
- 2 L'examen initialement prévu est indispensable

### TCS Vignette Clinique N°4

Youssef, nouveau-né à terme eutrophique, avec une bonne adaptation à la vie extra-utérine. Il est admis à J11 de vie pour exploration d'une cholestase avec des urines foncées, des selles décolorées et une hépatomégalie.

Si vous pensez à :	Et que l'on vous fournit :	L'hypothèse diagnostique initiale est :
Atrésie des voies biliaires	Le résultat de l'échographie hépatobiliaire : hépatomégalie homogène vésicule biliaire visualisée, voies biliaires intra hépatiques et extra hépatiques non dilatées	-2 -1 0 1 2

- 2 L'hypothèse initiale est pratiquement éliminée
- 1 L'hypothèse initiale est moins probable
- 0 L'hypothèse initiale n'est pas influencée par la donnée supplémentaire
- 1 L'hypothèse initiale est plus probable
- 2 L'hypothèse initiale est pratiquement certaine

### TCS Vignette Clinique N°5

Amenallah, nouveau-né à terme eutrophique, né à 39 semaines d'aménorrhées, sorti à domicile à J2 de vie. Il est admis des urgences à J5 de vie pour refus de téter, dyspnée et troubles hémodynamiques.

Si vous pensez à :	Et que vous trouvez à l'examen clinique :	L'hypothèse diagnostique initiale est :
Infection néonatale	Des pouls fémoraux abolis, une différentielle des pressions artérielles entre membres supérieurs et inférieurs.	-2-1012

-2 L'hypothèse initiale est pratiquement éliminée

-1 L'hypothèse initiale est moins probable

0 L'hypothèse initiale n'est pas influencée par la donnée supplémentaire

1 L'hypothèse initiale est plus probable

2 L'hypothèse initiale est pratiquement certaine

### TCS Vignette Clinique N°6

Ahmed est né par césarienne en urgence pour souffrance fœtale aigue. A la naissance, il avait une fréquence cardiaque à 80 battements/mn et une respiration irrégulière. Vous êtes appelé à le réanimer.

Si vous pensez faire :	Et que l'on vous informe que :	La prescription ou le geste envisagé(e) initialement est :
Une ventilation en pression positive au masque	Le liquide amniotique était méconial	-2-1012

-2 La prescription ou le geste envisagé(e) est absolument contre-indiqué

-1 La prescription ou le geste envisagé(e) est moins justifié

0 La prescription ou le geste envisagé(e) n'est pas influencé par la donnée supplémentaire

1 La prescription ou le geste envisagé(e) est plus justifié

2 La prescription ou le geste envisagé(e) est indispensable

### TCS Vignette Clinique N°7

Hiba, née à terme par voie basse à 39 semaines d'aménorrhées. A l'examen systématique, vous notez une hypertrophie clitoridienne, sans autres anomalies cliniques associées.

Si vous pensez à une :	Et que l'on vous dit :	L'hypothèse diagnostique initiale est :
Hyperplasie congénitale des surrénales	Que le frère aîné de Hiba est décédé la deuxième semaine de vie dans un tableau de déshydratation.	-2-1012

- 2 L'hypothèse initiale est pratiquement éliminée
- 1 L'hypothèse initiale est moins probable
- 0 L'hypothèse initiale n'est pas influencée par la donnée supplémentaire
- 1 L'hypothèse initiale est plus probable
- 2 L'hypothèse initiale est pratiquement certaine

### TCS Vignette Clinique N°8

Ines, nouveau-né eutrophique, née à terme par voie basse. A H28 de vie, vous êtes amené à l'examiner avant la sortie, l'examen clinique était sans particularités.

Si vous pensez	Et que l'on vous dit :	La prescription ou le geste envisagé(e) initialement est :
Autoriser sa sortie	que le méconium n'est pas encore émis	-2-1012

- 2 La prescription ou le geste envisagé(e) est absolument contre-indiqué
- 1 La prescription ou le geste envisagé(e) est moins justifié
- 0 La prescription ou le geste envisagé(e) n'est pas influencé par la donnée supplémentaire
- 1 La prescription ou le geste envisagé(e) est plus justifié
- 2 La prescription ou le geste envisagé(e) est indispensable

### TCS Vignette Clinique N°9

Hichem, nouveau-né eutrophique, né à terme par voie basse. A H36 de vie, vous êtes amené à l'examiner avant la sortie, l'examen clinique était sans particularités.

Si vous pensez	Et que l'on vous dit :	La prescription ou le geste envisagé(e) initialement est :
Autoriser sa sortie	qu'il n'a pas émis ses urines	-2-1012

-2 La prescription ou le geste envisagé(e) est absolument contre-indiqué

-1 La prescription ou le geste envisagé(e) est moins justifié

0 La prescription ou le geste envisagé(e) n'est pas influencé par la donnée supplémentaire

1 La prescription ou le geste envisagé(e) est plus justifié

2 La prescription ou le geste envisagé(e) est indispensable

### TCS Vignette Clinique N°10

Yassine, issu d'une mère primipare, est revu en consultation externe à l'âge de deux semaines. Vous avez objectivé une prise pondérale insuffisante que vous avez rattachée à une mauvaise conduite de l'allaitement maternel.

Si vous pensez prescrire :	Et que l'on vous dit :	La prescription ou le geste envisagé(e) initialement est :
La poursuite de l'allaitement maternel exclusif	que la mère prend des antalgiques à type de paracétamol pour des arthralgies.	-2-1012

-2 La prescription ou le geste envisagé(e) est absolument contre-indiqué

-1 La prescription ou le geste envisagé(e) est moins justifié

0 La prescription ou le geste envisagé(e) n'est pas influencé par la donnée supplémentaire

1 La prescription ou le geste envisagé(e) est plus justifié

2 La prescription ou le geste envisagé(e) est indispensable

### TCS Vignette Clinique N°11

Adem, est amené aux urgences à J 10 de vie pour hématomé. L'examen a noté une légère pâleur sans autres anomalies cliniques. Le bilan d'hémostase était normal. La numération formule sanguine a montré une anémie à 11 g/dl.

Si vous pensez demander :	Et que la mère vous dit :	L'examen initialement prévu est :
Une fibroscopie oeso-gastro-duodénale	qu'elle a des mamelons crevassés.	-2-1012

- 2 L'examen initialement prévu est absolument contre-indiqué
- 1 L'examen initialement prévu est moins justifié
- 0 L'examen initialement prévu n'est pas influencé par la donnée supplémentaire
- 1 L'examen initialement prévu est plus justifié
- 2 L'examen initialement prévu est indispensable

### TCS Vignette Clinique N°12

Ayoub, nouveau-né eutrophique à terme, né par voie basse avec une bonne adaptation à la vie extra-utérine. Il était gardé auprès de sa mère. A H 10 de vie, vous avez objectivé une hyporéactivité et un geignement. La glycémie au talon était normale. Le nouveau-né était admis en unité de soins intensifs.

Si vous pensez à une :	Et que l'on vous fournit :	L'hypothèse diagnostique initiale est :
Infection materno-fœtale	Le résultat de la numération formule sanguine et de la CRP : NFS sans anomalies, CRP<6mg/l	-2-1012

- 2 L'hypothèse initiale est pratiquement éliminée
- 1 L'hypothèse initiale est moins probable
- 0 L'hypothèse initiale n'est pas influencée par la donnée supplémentaire
- 1 L'hypothèse initiale est plus probable
- 2 L'hypothèse initiale est pratiquement certaine

### TCS Vignette Clinique N°13

Yassine et yasmine, nouveau-nés issus d'une grossesse gémellaire compliquée d'une prématurité spontanée à 36 semaines d'aménorrhées, sans autre anamnèse infectieuse à l'accouchement. Les deux nouveau-nés étaient asymptomatiques à la naissance.

Si vous pensez à :	Et que l'on vous dit:	L 'hypothèse diagnostique initiale est :
La gémellité comme cause de prématurité	La mère a développé une fièvre à 39 °C, 4 heures après l'accouchement.	-2-1012

- 2 L'hypothèse initiale est pratiquement éliminée
- 1 L'hypothèse initiale est moins probable
- 0 L'hypothèse initiale n'est pas influencée par la donnée supplémentaire
- 1 L'hypothèse initiale est plus probable
- 2 L'hypothèse initiale est pratiquement certaine

### TCS Vignette Clinique N°14

Sana, nouveau-né à terme, né par césarienne pour une souffrance fœtale aigue, à 39 semaines d'aménorrhées. Elle était admise à la naissance pour prise en charge d'une hypotrophie sévère harmonieuse (poids=1800g, Taille=42cm, périmètre crânien=32 cm).

Elle est issue d'une mère âgée de 34 ans, G1P1, grossesse compliquée d'une toxémie gravidique, découverte au deuxième trimestre.

Si vous pensez que:	Et que l'on vous dit que:	L 'hypothèse diagnostique initiale est :
Un syndrome vasculo-rénal est à l'origine de cette hypotrophie	L'échographie transfontanellaire a montré : des calcifications parenchymateuses avec une ventriculomégalie bilatérale	-2-1012

- 2 L'hypothèse initiale est pratiquement éliminée
- 1 L'hypothèse initiale est moins probable
- 0 L'hypothèse initiale n'est pas influencée par la donnée supplémentaire
- 1 L'hypothèse initiale est plus probable
- 2 L'hypothèse initiale est pratiquement certaine



### TCS Vignette Clinique N°15

Yosra, nouveau-né à terme, née par voie basse dans un contexte de choriomniotites (fièvre maternelle à 39°C avec ouverture prolongée de la poche des eaux). Elle a développé à la naissance une détresse respiratoire indiquant son admission en unité de soins intensifs.

Si vous pensez prescrire	Et que l'on vous fournit :	La prescription ou le geste envisagé(e) initialement est :
Une antibiothérapie probabiliste triple : ampicilline, céphalosporine de 3 <sup>ème</sup> génération et aminosides	A l'examen bactériologique direct du liquide gastrique : présence de cocci Gram positif en chainettes	-2-1012

-2 La prescription ou le geste envisagé(e) est absolument contre-indiqué

-1 La prescription ou le geste envisagé(e) est moins justifié

0 La prescription ou le geste envisagé(e) n'est pas influencé par la donnée supplémentaire

1 La prescription ou le geste envisagé(e) est plus justifié

2 La prescription ou le geste envisagé(e) est indispensable

### TCS Vignette Clinique N°16

Amenallah, nouveau-né à terme eutrophique, admis à la naissance pour suspicion de trisomie 21. Il a développé à H4 de vie des vomissements bilieux sans ballonnement abdominal ni autres anomalies cliniques.

Si vous pensez à une :	Et que l'on vous dit :	L'hypothèse diagnostique initiale est :
Atrésie duodénale	qu'il a émis son méconium dès la première heure de vie	-2-1012

-2 L'hypothèse initiale est pratiquement éliminée

-1 L'hypothèse initiale est moins probable

0 L'hypothèse initiale n'est pas influencée par la donnée supplémentaire

1 L'hypothèse initiale est plus probable

2 L'hypothèse initiale est pratiquement certaine

### TCS Vignette Clinique N°17

Raouf, nouveau-né à terme eutrophique, issu d'une grossesse menée à 39 semaines d'aménorrhées, compliquée d'une rupture prématurée des membranes de 48 heures. L'accouchement a eu lieu par césarienne pour échec de déclenchement. Il a développé une détresse respiratoire immédiate indiquant son admission en unité de soins intensifs néonataux.

Si vous pensez à une :	Et que l'on vous fournit :	L'hypothèse diagnostique initiale est :
Infection materno-fœtale	Une radiographie de thorax normale	-2-1012

- 2 L'hypothèse initiale est pratiquement éliminée
- 1 L'hypothèse initiale est moins probable
- 0 L'hypothèse initiale n'est pas influencée par la donnée supplémentaire
- 1 L'hypothèse initiale est plus probable
- 2 L'hypothèse initiale est pratiquement certaine

### TCS Vignette Clinique N°18

Aziz, nouveau-né à terme eutrophique, issu d'une grossesse compliquée de diabète gestationnel équilibré sous régime et d'un hydramnios. L'accouchement a eu lieu par voie basse sans anamnèse infectieuse. Le nouveau-né était asymptomatique à la naissance.

Il a développé une détresse respiratoire secondaire à H1 de vie avec battement des ailes du nez, un tirage intercostal et des mousses aux lèvres.

Si vous pensez à une :	Et que l'on vous dit :	L'hypothèse diagnostique initiale est :
Atrésie de l'œsophage	que l'épreuve à la seringue faite à la naissance était positive.	-2-1012

- 2 L'hypothèse initiale est pratiquement éliminée
- 1 L'hypothèse initiale est moins probable
- 0 L'hypothèse initiale n'est pas influencée par la donnée supplémentaire
- 1 L'hypothèse initiale est plus probable
- 2 L'hypothèse initiale est pratiquement certaine

### TCS Vignette Clinique N°19

Rafik, nouveau-né à terme eutrophique, issu d'une grossesse menée à 41 semaines d'aménorrhées avec un accouchement par césarienne pour défaut d'engagement. Il était admis à J 1 de vie en néonatalogie pour exploration d'une hydrocéphalie.

Si vous pensez à une :	Et que l'on vous fournit :	L'hypothèse diagnostique initiale est :
Toxoplasmose congénitale	Deux sérologies de toxoplasmose maternelles négatives, faites à 12 et à 34 semaines d'aménorrhées.	-2-1012

- 2 L'hypothèse initiale est pratiquement éliminée
- 1 L'hypothèse initiale est moins probable
- 0 L'hypothèse initiale n'est pas influencée par la donnée supplémentaire
- 1 L'hypothèse initiale est plus probable
- 2 L'hypothèse initiale est pratiquement certaine

### TCS Vignette Clinique N°20

Siwar, nouveau-né à terme eutrophique, est amenée aux urgences au troisième jour de vie pour fièvre chiffrée à 39°C avec un refus de téter. L'examen a objectivé une hyporéactivité avec une température à 37,5°C.

Si vous pensez prescrire :	Et que l'on vous dit :	La prescription ou le geste envisagé(e) initialement est :
une ponction lombaire	Que le nouveau-né a présenté des mouvements anormaux à type de fixité de regard.	-2-1012

- 2 La prescription ou le geste envisagé(e) est absolument contre-indiqué
- 1 La prescription ou le geste envisagé(e) est moins justifié
- 0 La prescription ou le geste envisagé(e) n'est pas influencé par la donnée supplémentaire
- 1 La prescription ou le geste envisagé(e) est plus justifié
- 2 La prescription ou le geste envisagé(e) est indispensable

**Annexe 3**

<b>Vignette</b>	<b>Oui</b>	<b>Non</b>	<b>À réviser</b>
Décrit une situation problématique, même pour un expert (pas assez de données pour conclure, données ambiguës ou conflictuelles)			
Décrit une situation crédible et appropriée au niveau de formation des personnes évaluées			
La vignette est nécessaire pour comprendre la question et pour mettre la personne évaluée dans un contexte clinique spécifique			
S'agit-il d'une présentation clinique usuelle (la situation ne doit pas être un piège)			
La vignette est correctement rédigée (incorections, longueur, confusion)			
<b>Options (Diagnostiques, d'Investigation, Thérapeutiques)</b>	<b>Oui</b>	<b>Non</b>	<b>À réviser</b>
Pertinentes à la situation clinique, du point de vue d'un expert			
La même option ne se retrouve pas dans deux questions consécutives			
<b>L'information complémentaire dans la 2e colonne</b>	<b>Oui</b>	<b>Non</b>	<b>À réviser</b>
Permet de tester la force de l'association entre l'option et l'information nouvelle, dans le contexte clinique décrit par la vignette			
Les items sont construits de façon à répartir les réponses de façon équilibrée entre: -2, -1, 0, +1, +2			
Les items construits correspondent à des questions que l'on se pose réellement en situation clinique			
<b>La composition du panel de référence</b>	<b>Oui</b>	<b>Non</b>	<b>À réviser</b>
Le nombre est suffisamment important (10 ou plus si l'enjeu est considérable)			
Chaque membre du panel a passé le test individuellement, dans des conditions similaires à celles des personnes testées			
Les membres choisis peuvent légitimement faire partie d'un panel de référence, compte tenu des caractéristiques des personnes examinées.			

### Annexe 4

Réponses des experts au test de concordance de script et scores attribués																				
Echelle Likert	Vignette 1		Vignette 2		Vignette 3		Vignette 4		Vignette 5		Vignette 6		Vignette 7		Vignette 8		Vignette 9		Vignette 10	
	R	S	R	S	R	S	R	S	R	S	R	S	R	S	R	S	R	S	R	S
-2	0	0	0	0	0	0	1	0.16	2	0.28	9	1	0	0	8	1	6	1	0	0
-1	0	0	0	0	1	0.2	6	1	7	1	1	0.11	0	0	2	0.25	2	0.33	1	0.16
0	4	1	0	0	1	0.2	2	0.33	1	0.14	0	0	0	0	0	0	2	0.33	6	1
1	4	1	5	1	3	0.6	1	0.16	0	0	0	0	5	1	0	0	0	0	0	0
2	2	0.5	5	1	5	1	0	0	0	0	0	0	5	1	0	0	0	0	3	0.5
Echelle Likert	Vignette 11		Vignette 12		Vignette 13		Vignette 14		Vignette 15		Vignette 16		Vignette 17		Vignette 18		Vignette 19		Vignette 20	
	R	S	R	S	R	S	R	S	R	S	R	S	R	S	R	S	R	S	R	S
-2	1	0.2 5	0	0	4	0.6 6	7	1	0	0	0	0	0	0	3	0.6	1	0.2	1	0.16
-1	4	1	3	0.42	6	1	3	0.42	7	1	0	0	1	0.12	2	0.4	5	1	0	0
0	3	0.7 5	7	1	0	0	0	0	2	0.28	10	1	8	1	5	1	4	0.8	1	0.16
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0.33
2	2	0.5	0	0	0	0	0	0	1	0.14	0	0	1	0.12	0	0	0	0	6	1

## Résumé

**Introduction :** Le test de concordance de script évalue le raisonnement clinique en contexte d'incertitude. Il vise à comparer l'organisation des connaissances, élément essentiel de la compétence clinique.

**L'objectif de ce travail** était de construire un test de concordance de script et de démontrer sa pertinence dans l'évaluation des compétences en Néonatalogie des étudiants du deuxième cycle et des internes de la faculté de médecine de Tunis.

**Méthodes :** Un test de concordance de script comportant 20 vignettes à un item chacune, a été administré à 52 étudiants de deuxième cycle d'études médicales et à 11 internes de médecine générale. Le test d'une durée de 30 minutes, a été pratiqué en fin de stage.

**Résultats :** Les experts ont eu un score moyen, rapporté sur 100, de  $82.52 \pm 7.35$  IC<sub>95%</sub> [77.26-87.78], plus élevé que les externes de  $58.52 \pm 9.72$  IC<sub>95%</sub> [55.82-61.23], ( $p < 0.0001$ ) ; et les internes de  $63.17 \pm 11.36$  IC<sub>95%</sub> [55.53-70.81] ( $p < 0.0001$ ). La différence des scores entre étudiants et internes n'était pas statistiquement significative,

**Conclusion :** Notre travail a démontré la faisabilité du TCS en stage de néonatalogie. Il a permis de différencier les experts des internes et des étudiants. Le recours au TCS serait plus pertinent pour l'évaluation des internes et des résidents, que celle des étudiants de deuxième cycle qui doivent encore acquérir des repères avant de manier l'incertitude.

**Mots clés :** Test de concordance de script, Evaluation, Compétence, Raisonnement clinique, Néonatalogie.

### Abstract

**Background :** Script Concordance Test evaluates clinical reasoning in context of uncertainty. It aims to evaluate knowledge organization, which represents an essential component of the clinical competence.

**Aim :** To build a script concordance test and demonstrate its relevance in the evaluation of Neonatology skills of fourth year medical students and interns of the Medicine Faculty of Tunis.

**Methods :** A Script Concordance Test including 20 vignettes and 20 items, was provided to 52 fourth year medical students and 11 family medicine interns. The 30 minutes test was completed at the end of the training.

**Results :** Script Concordance Test scores obtained by experts were higher than those obtained by students and family medicine interns. The scores (out of 100) were  $82.52 \pm 7.35$  CI<sub>95%</sub> [77.26-87.78] for the experts,  $58.52 \pm 9.72$  CI<sub>95%</sub>[55.82-61.23] ( $p < 0.0001$ ) for the students and  $63.17 \pm 11.36$  CI<sub>95%</sub> [55.53-70.81] ( $p < 0.0001$ ) for the residents. The difference in scores between students and interns was not statistically significant.

**Conclusion:** Our data suggests that script concordance test could be used to assess the acquisition of clinical reasoning among fourth year medical students in neonatology. It differentiates experts and students. The use of script concordance test would be more relevant for the evaluation of residents than students who have yet to acquire marks before handling uncertainty.

**Keywords :** Script concordance test, Evaluation, skills, clinical reasoning, Neonatology.