



## Rapport de l'étude pilote :

# Procédure d'évaluation par les pairs (PEP)

## pour l'évaluation de la qualité des expertises médicales de l'AI

Etabli par le secrétariat spécialisé COQEM et approuvé par la COQEM lors de la séance du 25 octobre 2023.

### Résumé

L'objectif premier de l'étude pilote était de valider le catalogue de questions de contrôle de la procédure d'évaluation par les pairs (PEP). L'étude a été menée avec succès, la formation des évaluateurs a été efficace et la mise en œuvre de la formation a été bonne. La plateforme électronique en ligne utilisée s'est révélée fiable au cours de l'étude et aucun problème technique important n'a été rencontré.

Les évaluateurs ont jugé que les définitions des critères et du degré de la gravité des lacunes (définition des ancrages) des questions de contrôle étaient compréhensibles et importantes, ce qui a confirmé l'utilité de la formation, puisque seules des précisions minimales ont été nécessaires. Cela souligne la qualité d'ores et déjà atteinte par le support de formation. Le temps moyen consacré à l'examen d'une expertise partielle a été de 95 minutes, tandis que les expertises lacunaires ont nécessité un temps significativement plus élevé (107 minutes). Dans ce contexte, il apparaît clairement que la qualité de l'expertise a également une influence considérable sur le temps consacré à la PEP.

Un examen du catalogue de questions en vue d'éventuelles réductions a montré que la suppression des questions de contrôle formelles dans la version complète PEP n'entraînerait pas de gain de temps substantiel. Cela suggère que toutes les questions de contrôle actuelles sont essentielles pour une évaluation approfondie des expertises et devraient être conservées. En revanche, trois questions de contrôle supplémentaires ont dû être ajoutées à la version courte PEP afin d'améliorer la saisie des aspects critiques de la qualité. Parallèlement, il a été proposé de supprimer deux des questions de contrôle les moins pertinentes de la version courte afin d'améliorer l'efficacité de la procédure et de se concentrer davantage sur les éléments essentiels de l'évaluation.

# Procédure d'évaluation par les pairs (PEP-CH)

M. Braun<sup>a</sup>, M. Hermelink<sup>a</sup>, R. Jaquier<sup>c</sup>, M. Liebrez<sup>c,d</sup> & R. Schleifer<sup>a</sup>

a Secrétariat spécialisé de la Commission fédérale d'assurance qualité des expertises médicales  
b Office fédéral des assurances sociales  
c Commission fédérale d'assurance qualité des expertises médicales  
d Forensisch-Psychiatrischer Dienst der Universität Bern

## Introduction

Jusqu'à présent, les services médicaux régionaux de l'assurance-invalidité en Suisse ont évalué la qualité des expertises de manière non systématique et non uniforme, au cas par cas<sup>1</sup>.

Un groupe de travail de l'Office fédéral des assurances sociales a développé la procédure suisse d'évaluation par les pairs (PEP-CH) pour l'évaluation des expertises médicales. La Commission fédérale pour l'assurance qualité dans le domaine de l'expertise médicale, créée par le législateur en 2022, a mis en œuvre la PEP-CH sous forme d'outil en ligne. Cet outil tient compte de la procédure établie par l'assurance pension allemande<sup>2,3</sup>.

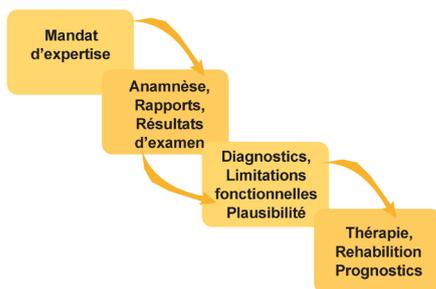
## Description de l'instrument pour l'évaluation des expertises médicales

L'instrument contient 24 questions pour mesurer 6 critères.

Chaque question est formulée de manière détaillée et contient des définitions de gravité pour l'évaluation sur une échelle de 4 points.

Critères de qualité	
A:	Structure formelle
B:	Clarté
C:	Transparence
D:	Exhaustivité
E:	Compétences techniques
F:	Économité

### Compréhensibilité de l'argumentation



## Application future de l'instrument

- Assurance qualité générale, non liée à un cas particulier (échantillon représentatif, base de référence)
- Examen périodique de l'état actuel de la qualité de l'expertise (mesure du changement)
- Audits et examens d'accréditation des organismes d'expertise (déficits : écart par rapport à la ligne de base)
- Formation et perfectionnement des évaluateurs (retour par un service en ligne en double aveugle)

## Exemple d'une question de la PEP-CH

Question D3	Critère D : exhaustivité
Dans quelle mesure les divergences significatives entre l'évaluation médico-assurantielle et d'autres rapports antérieurs concernant les capacités fonctionnelles sont-elles décrites et évaluées ?	

Il s'agit ici de conclusions sur les capacités fonctionnelles qui divergeraient des conclusions de rapports antérieurs et pourraient ainsi aboutir à une autre décision juridique. Toute divergence en termes de quantité et de qualité des capacités fonctionnelles, ou concernant le début et la durée de la perte de capacité ou du besoin de traitement, doit être décrite et évaluée.

Par **rapports antérieurs**, on entend entre autres : rapport de sortie d'un centre de réadaptation, expertise de l'organisme d'assurance sociale, expertise judiciaire, expertise d'autres assureurs, rapports de professionnels impliqués dans le traitement ou d'hôpitaux.

Si l'expertise ne mentionne pas de rapports antérieurs divergents, cocher « aucune lacune ».

<input type="checkbox"/>	Aucune lacune	
<input type="checkbox"/>	Lacunes mineures	Il existe des divergences significatives entre l'évaluation médico-assurantielle et d'autres rapports antérieurs concernant les capacités fonctionnelles. L'expertise les met en évidence, mais l'évaluation n'en explique pas certains aspects.
<input type="checkbox"/>	Lacunes manifestes	Il existe des divergences significatives entre l'évaluation médico-assurantielle et d'autres rapports antérieurs concernant les capacités fonctionnelles. L'expertise les met en évidence, mais l'évaluation ne les explique pas.
<input type="checkbox"/>	Lacunes graves	L'expertise cite des rapports antérieurs, mais ne dit rien des observations qu'ils contiennent sur les capacités fonctionnelles. OU Il existe des divergences significatives entre l'évaluation médico-assurantielle et d'autres rapports antérieurs concernant les capacités fonctionnelles, mais elles ne sont pas démontrées dans l'évaluation.

## Validation de l'instrument

- Mise en œuvre du projet de mai à octobre 2023
- Formation de 8 experts indépendants (2 rhumatologues, 2 médecins généralistes, 2 neurologues et 2 psychiatres)
- 10 échantillons anonymes de rapports d'experts avec 3-5 disciplines
- Méthode en double aveugle (outil en ligne)
- Évaluation de l'importance des 24 items pour la décision sur la qualité
  - Réduction du nombre des questions
  - Validation de l'instrument

## Conclusion

La procédure suisse d'évaluation par les pairs (PEP-CH) offre un moyen d'améliorer l'évaluation médicale dans le cadre de la sécurité sociale, ce qui permet d'obtenir de meilleurs résultats plus équitables pour les personnes qui dépendent de ces prestations.

### Littérature:

<sup>1</sup> Schleifer, R., & Liebrez, M. (2023). Introducing the Federal Commission for Quality Assurance in Medical Assessment. *Therapeutische Umschau. Revue thérapeutique*, 80(2), 85-88.

<sup>2</sup> Deutsche Rentenversicherung (2021). Qualitätssicherung der Begutachtung - Manual zum Peer Review [Quality Assurance of Review - Manual for Peer Review]. [https://www.deutsche-rentenversicherung.de/SharedDocs/Downloads/DE/Experten/infos\\_fuer\\_aerzte/begutachtung/manual\\_peer\\_review.html](https://www.deutsche-rentenversicherung.de/SharedDocs/Downloads/DE/Experten/infos_fuer_aerzte/begutachtung/manual_peer_review.html) [accessed 15th September, 2023].

<sup>3</sup> Strahl A, Gerlich C, Alpers GW, et al (2018). Development and evaluation of a standardized peer-training in the context of peer review for quality assurance in work capacity evaluation. *BMC Med Educ*. 18(1):135.

## Table des matières

1. Situation initiale .....	4
2. Objectifs de l'enquête .....	6
3. Conception de l'étude pilote.....	7
4. Evaluation des données et présentation des résultats .....	9
4.1 Gravité des lacunes .....	9
4.2 Appréciation de l'importance des différentes questions de contrôle .....	11
4.3 Temps nécessaire : .....	13
4.4 Différences entre les 23 questions de contrôle selon les disciplines .....	14
4.5 Analyse factorielle des 23 questions de contrôle.....	16
4.6 Analyse de régression des 23 questions de contrôle comme prédicteurs de l'évaluation globale .....	17
5. Discussion des résultats .....	18
5.1 Acquisition des évaluateurs qualifiés (faisabilité) .....	18
5.2 Développement et mise en œuvre d'une première formation pour les évaluateurs .....	18
5.3 Test de la forme de saisie électronique pour la version complète PEP .....	18
5.4 Saisie plus précise du temps nécessaire à la version complète PEP .....	19
5.5 Identification des questions de contrôle pertinentes dans la version courte PEP .....	20
5.6 Identification des questions de contrôle selon la gravité des lacunes .....	21
5.7 Décider si et comment la version complète PEP doit être raccourcie .....	21
5.8 Limitations de l'étude pilote PEP .....	22
6. Conclusions et recommandations .....	23

## Liste des tableaux

Tableau 1: Aperçu des questions des versions complète et courte PEP .....	5
Tableau 2: Conception de l'étude pilote version complète PEP .....	7
Tableau 3: Fréquences des lacunes par question de l'évaluation consensuelle .....	9
Tableau 4: Fréquences des lacunes par question des expertises partielles.....	10
Tableau 5: Appréciation de l'importance des questions de l'évaluation consensuelle .....	11
Tableau 6: Appréciation de l'importance des questions des expertises partielles .....	12
Tableau 7: Temps nécessaire en fonction de l'évaluation globale .....	13
Tableau 8: Temps passé en minutes par discipline .....	14
Tableau 9: Différences d'évaluation des lacunes selon les disciplines.....	14
Tableau 10: Différences d'évaluation de l'importance selon les disciplines.....	15
Tableau 11: Aperçu des questions des versions complète et courte PEP.....	24

## 1. Situation initiale

Un groupe d'experts mandaté par l'assurance-invalidité fédérale a élaboré en 2020 un catalogue de questions sur la procédure d'évaluation par les pairs (PEP) sous la direction du Dr Hermelink, en se basant sur le manuel de contrôle de la *Deutsche Rentenversicherung* (édition 2018). La Commission fédérale d'assurance qualité des expertises médicales (COQEM) a repris ces travaux préliminaires depuis le début de son activité en 2022, en a discutés en détail et a décidé de valider le catalogue de questions de cette procédure dans le cadre d'une étude pilote afin d'obtenir une base de décision pour l'utilisation d'une version courte PEP par rapport à la version complète PEP.

La COQEM a l'intention d'utiliser la version complète PEP pour un échantillon aléatoire représentatif d'au moins 100 des 11'000 expertises réalisées chaque année en Suisse dans le cadre d'une procédure en double aveugle et de manière anonyme, avec des évaluateurs indépendants, expérimentés et formés à la PEP à partir de 2024, afin de mesurer le niveau de qualité des expertises médicales (mesures de l'état actuel)<sup>1</sup> et de suivre l'évolution de la qualité au cours des prochaines années, ainsi que d'identifier d'éventuels lacunes de qualité lors d'audits de centres d'expertises et de contribuer à une amélioration dans le cadre d'un dialogue sur la qualité avec les personnes concernées. La PEP envisagée par la COQEM ne doit pas seulement être indépendante des cas, mais aussi du mandant (assurance) et de l'OFAS (partenaire contractuel des centres d'expertises pluridisciplinaires) et se concentrer fortement, outre les exigences formelles, sur les aspects professionnels de la qualité des expertises.

D'autre part, il est prévu que les questions de contrôle de la version courte PEP soient utilisées de manière non anonyme par des médecins SMR (Service médical régional) formés dans le cadre de l'assurance qualité liée au cas de toutes les expertises AI et qu'ainsi, en tant que procédure uniforme de contrôle de la qualité, elles minimisent les pratiques parfois différentes et garantissent ainsi l'égalité de droit. Les résultats seront classés par cas dans des dossiers d'assurés, envoyés aux experts respectifs pour un feedback et enregistrés dans une banque de données<sup>2</sup> qui sera à l'avenir mise à la disposition de la Commission fédérale d'assurance qualité des expertises médicales (COQEM). La version courte PEP se concentrera sur les aspects formels des expertises et sur leur utilisabilité dans le cadre de la procédure AI.

Alors que la version courte PEP est réalisée par le médecin en charge du cas (souvent un médecin généraliste polyvalent) dans le cadre des contrôles systématiques des médecins des SMR pour toutes les disciplines (dans le sens d'un « polyvalent » médico-assurantiel), les expertises réalisées dans le cadre de la version complète PEP doivent être évaluées à chaque

---

<sup>1</sup> Dans un premier temps, il s'agit ici d'expertises de l'Assurance-Invalidité. Pour d'éventuelles applications dans le domaine de la LAA, des adaptations de la procédure seraient nécessaires.

<sup>2</sup> Cette base de données sera basée sur la plate-forme "SuisseMed@p". Ainsi, outre les résultats de la PEP, des données supplémentaires seront disponibles.

fois par un médecin spécialiste correspondant et les évaluations consensuelles en cas de pluridisciplinarité doivent être réalisées par un médecin généraliste de médecine interne, afin de garantir le meilleur contrôle professionnel possible.

**Tableau 1: Aperçu des questions des versions complète et courte PEP**

Critère	No.	Version complète PEP			Version courte PEP		
		Expertise monodisc.	Expertise pluridisc.	Consensuelle	Expertise monodisc.	Expertise pluridisc.	Consensuelle
Forme	A1	A1	A1	A1			
	A2	A2		A2	A2		A2
Compréhension	B1	B1	B1	B1			
	B2	B2	B2	B2			
	B3	B3	B3	B3			
Transparence	C1	C1	C1		C1	C1	
	C2	C2	C2				
	C3	C3	C3				
	C4	C4	C4				
	C5	C5	C5				
Complétude	D1	D1	D1				
	D2	D2	D2		D2	D2	
	D3	D3	D3	D3	D3	D3	D3
	D4	D4	D4	D4	D4	D4	D4
	D5			D5			D5
Qualité professionnelle	E1	E1	E1				
	E2	E2	E2				
	E3	E3	E3		E3	E3	
	E4	E4	E4		E4	E4	
	E5	E5	E5		E5	E5	
	E6	E6	E6	E6	E6	E6	E6
	E7	E7	E7	E7			
Economicité	F1	F1	F1				
	F2	F2	F2				
Argumentation							

Remarque : les questions surlignées en couleur saumon ont été nouvellement développées par des experts CH. Les questions non colorées ont été reprises de la version DRV (DE).

Par rapport à la version courte PEP, la réalisation de la version complète PEP est un peu plus étendue (23 contre 10 questions pour la version destinée aux expertises monodisciplinaires, et 11 contre 6 questions pour l'évaluation consensuelle)<sup>3</sup>. D'autre part, il faut s'attendre à un investissement personnel plus important en termes relatifs.

<sup>3</sup> Nous estimons le temps supplémentaire nécessaire pour la version complète PEP par rapport à la version courte PEP à 10-15 min par expertise partielle.

## 2. Objectifs de l'enquête

L'objectif de cette étude pilote était de vérifier la faisabilité pratique du projet. Il s'agissait en particulier de développer et de tester le concept de formation des évaluateurs. De plus, la version en ligne de la version complète PEP devait être mise à disposition et testée en pratique par les utilisateurs. Il était également prévu d'étudier les éventuels avantages (valeur ajoutée) de la version complète (si existants) par rapport à la version courte. Les résultats de l'étude devaient par exemple servir à raccourcir la version complète ou à opter pour la version courte<sup>4</sup>.

Les objectifs de l'étude :

- 1) Acquisition d'évaluateurs qualifiés (faisabilité).
- 2) Développement et réalisation d'une première formation pour les évaluateurs.
- 3) Tester la plateforme de saisie en ligne pour la version complète PEP (voir annexe IV).
- 4) Saisie plus précise du temps nécessaire à la version complète PEP pour calculer les différences (Le temps nécessaire à la version courte PEP issue du projet pilote SMR est connu).
- 5) Y avait-t-il des questions de contrôle pertinentes dans les 13 questions supprimées entre la version complète PEP (24) et celle de la version courte PEP (10), respectivement dans les 5 questions supprimées entre l'évaluation consensuelle de la version complète PEP (11) et celle de la version courte PEP (6) des expertises bi- et pluridisciplinaires ?
- 6) Quelles questions de contrôle présentaient des lacunes, à quelle fréquence et avec quelle gravité<sup>5</sup> ? Existait-t-il éventuellement des questions de contrôle « décisives » ? Existait-t-il, du point de vue des évaluateurs, des questions de contrôle particulièrement importantes pour l'évaluation de la qualité ? Ces questions de contrôle étaient-elles incluses dans la version courte PEP ?
- 7) Décider si et comment la version complète PEP aurait pu être raccourcie.

Pour cette étude pilote, nous avons délibérément renoncé à déterminer la fiabilité entre les évaluateurs car les 10 expertises évaluées représentaient un nombre de cas relativement faible. En particulier, en cas de faible concordance, il n'aurait pas été possible de déterminer à ce stade si ces résultats étaient dus à un manque de formation, à une expérience trop limitée de l'instrument par les évaluateurs ou à d'éventuelles lacunes réelles de la procédure.

---

<sup>4</sup> Selon un entretien oral avec l'un des développeurs de la version complète PEP, le professeur R. Stieglitz, le 9 novembre 2022, une réduction de la version complète PEP était tout à fait réalisable. Il était toutefois important de garantir la comparabilité de la version courte PEP et de la version complète PEP.

<sup>5</sup> Gravité des lacunes (aucune, légère, manifeste, grave).

### 3. Conception de l'étude pilote

Un pool de 10 expertises pluridisciplinaires (3 à 5 disciplines) a été constitué. L'étude portait sur les expertises spécialisées et les expertises consensuelles. Les disciplines les plus fréquentes des expertises pluridisciplinaires en Suisse ont été la psychiatrie (24%), la rhumatologie (13%), la neurologie (14%) et la médecine générale (26%). En outre, la neuropsychologie (6%) est également prise en compte lors de la sélection des expertises, bien que celles-ci n'aient pas été explicitement évaluées par des neuropsychologues dans le cadre de cette étude pilote PEP<sup>6</sup>.

Pour cette étude pilote, nous n'avons volontairement pas utilisé un échantillon représentatif d'expertises, mais nous avons procédé à une présélection d'expertises lacunaires de l'année 2022<sup>7</sup> afin d'obtenir une dispersion suffisante des possibilités de réponses aux questions de contrôle. Si les expertises ne présentaient pas ou peu de lacunes, l'étude ne permettrait pas de tirer des conclusions sur l'instrument de la version complète PEP ou de la version courte PEP. La présélection des expertises « de haute qualité » et des expertises « lacunaires » s'est basée sur l'évaluation « a priori » du secrétariat spécialisé COQEM et de quelques membres de la commission.

Les expertises (voir tableau 2) ont été évaluées respectivement par 2 psychiatres, 2 rhumatologues, 2 neurologues et 2 médecins généralistes, l'évaluation consensuelle a été effectuée par les 2 médecins généralistes. L'évaluation a été réalisée sur la base d'expertises anonymisées<sup>8</sup>.

Tableau 2: Conception de l'étude pilote version complète PEP

Expertises	Disciplines	Disciplines des évaluations et consensuelle				
		A (23)	P (23)	R (23)	N (23)	C (11)
PGA-23002	A, P, R, N, NP	A1, A2	P1, P2	R1, R2	N1, N2	A1, A2
PGA-23003	A, P, R	A1, A2	P1, P2	R1, R2		A1, A2
PGA-23010	A, P, R, N	A1, A2	P1, P2	R1, R2	N1, N2	A1, A2
PGA-23011	A, P, R, N	A1, A2	P1, P2	R1, R2	N1, N2	A1, A2
PGA-23012	A, P, R, N	A1, A2	P1, P2	R1, R2	N1, N2	A1, A2
PGA-23017	A, P, R, N	A1, A2	P1, P2	R1, R2	N1, N2	A1, A2
PGA-23018	A, P, R, N	A1, A2	P1, P2	R1, R2	N1, N2	A1, A2
PGA-23019	A, P, N, NP	A1, A2	P1, P2		N1, N2	A1, A2

<sup>6</sup> Les expertises spécifiques neuropsychologiques doivent absolument être prises en compte, car une question du PEP se réfère explicitement à l'utilisation/la prise en compte d'examen complémentaires et ne peut pas être évaluée autrement.

<sup>7</sup> Les exigences formelles actuelles de l'OFAS ont été prises en compte.

<sup>8</sup> L'anonymisation a été effectuée électroniquement par le secrétariat spécialisé de la COQEM avec Adobe Pro.

PGA-23020	A, P, N ortho, cardio	A1, A2	P1, P2	R1, R2	N1, N2	A1, A2
PGA-23023	A, P, ortho, N	A1, A2	P1, P2	R1, R2	N1, N2	A1, A2

PGA : expertise pluridisciplinaire, A : médecine générale, P : psychiatrie, R : rhumatologie, N : neurologie, NP : neuropsychologie, C : partie consensuelle (Le nombre entre parenthèses correspond au nombre de questions de contrôle). L'expertise partielle de l'orthopédie a été évaluée par deux rhumatologues.

En plus de l'évaluation de la version complète PEP, les expertises ont été évaluées par les évaluateurs d'une manière globale (insuffisant, moyen, bon) après avoir terminé le catalogue des questions de contrôle. Ensuite, les réponses individuelles ont été passées en revue une nouvelle fois et l'importance de chaque question de contrôle pour l'évaluation globale a été évaluée à l'aide d'une échelle de 3 appréciations (importante, moyenne, sans importance).

Ce concept a permis d'examiner si la version courte PEP, qui ne comportait que 10 questions de contrôle (une partie de la version complète) permettait déjà de saisir les aspects importants ou quels contenus ressortaient avec les 14 autres questions de contrôle de la version complète. Pour l'évaluation consensuelle C11, seules 6 questions de contrôle ont été utilisées pour la version courte PEP. Ici aussi, la question a été posée de savoir si les 5 questions de contrôle supplémentaires de la version complète offraient une valeur ajoutée.

Sur la base des résultats descriptifs (tendances), les membres de la commission ont évalué le type de feedback qu'ils jugeaient particulièrement utile et celui qu'ils jugeaient inutile pour le développement de la qualité des expertises médicales.

Pour cette étude pilote, des contrats ont été établis en février 2023 avec 8 médecins spécialistes (2 psychiatres, 2 rhumatologues, 2 médecins généralistes et 2 neurologues), qui ont été formés d'avril à mai 2023 sur tous les aspects de la version complète PEP et sur l'utilisation de la plateforme en ligne à l'aide de deux expertises anonymisées.

De début juin à fin septembre 2023, les évaluations des 10 expertises pluridisciplinaires anonymisées ont été effectuées par les 8 évaluateurs.

Les données d'évaluation des 10 expertises par les 8 évaluateurs ont été exportées de la plateforme en ligne et mises à disposition sous forme de fichier Excel. Cela a permis une évaluation ultérieure à l'aide du logiciel de statistiques SAS (Version 8.3 Deutsch SAS Enterprise Guide)

#### 4. Evaluation des données et présentation des résultats

Tout d'abord, des méthodes statistiques descriptives ont été utilisées afin de décrire les données de manière compréhensible : Les fréquences et les pourcentages des lacunes de chaque question de contrôle, la médiane et les valeurs moyennes du degré de gravité des lacunes et de l'importance de chaque question de contrôle pour l'évaluation globale des expertises ont été déterminés.

Dans un deuxième temps, des méthodes d'analyse inférentielle de la statistique ont été utilisées pour mettre en évidence des différences significatives des valeurs moyennes du degré de gravité des lacunes et de l'importance selon les critères ou les disciplines (au moyen d'analyses de variance) et pour établir des corrélations significatives (au moyen d'analyses de régression, d'analyses factorielles) et pour fournir des informations utiles qui ont servi à éliminer d'éventuelles questions de contrôle.

##### 4.1 Gravité des lacunes

Dans un premier temps, les pourcentages de lacunes ont été déterminés et listés pour chaque question de contrôle : il a ainsi été possible de voir rapidement pour quelles questions de contrôle la plupart des lacunes graves ont été constatées, respectivement pour quelles questions aucune lacune n'a été observée dans la majorité des cas. En outre, le degré de gravité des lacunes a également pu être calculé à l'aide de la médiane ou de la moyenne et de l'écart-type et utilisé pour les analyses différentielles ultérieures des questions de contrôle. La formulation des questions de contrôle est présentée dans l'annexe III.

Tableau 3: Fréquences des lacunes par question de l'évaluation consensuelle

	Nombre d'évaluations				Proportions en pourcentage				Médiane	Valeur moyenne
	Pas de lacune =1	Légères lacunes =2	Lacunes manifestes =3	Graves lacunes =4	Pas de lacune =1	Légères lacunes =2	Lacunes manifestes =3	Graves lacunes =4		
<b>A1</b>	13	1	1	5	65%	5%	5%	25%	1	1.9
<b>A2</b>	10	0	2	8	50%	0%	10%	40%	1	2.4
<b>B1</b>	12	6	2	0	60%	30%	10%	0%	1	1.5
<b>B2</b>	6	2	5	7	30%	10%	25%	35%	3	2.7
<b>B3</b>	17	3	0	0	85%	15%	0%	0%	1	1.2
<b>D3</b>	13	2	2	3	65%	10%	10%	15%	1	1.8
<b>D4</b>	10	2	3	5	50%	10%	15%	25%	1	2.2
<b>D5</b>	7	2	0	11	35%	10%	0%	55%	4	2.8
<b>E6</b>	9	2	6	3	45%	10%	30%	15%	2	2.2
<b>E7</b>	12	2	2	4	60%	10%	10%	20%	1	1.9
	<b>Vert</b>	<b>Jaune</b>	<b>Rouge</b>		<b>Vert</b>	<b>Jaune</b>	<b>Rouge</b>		<b>Médiane</b>	<b>Valeur moyenne</b>
<b>G1</b>	3	8	9		15%	40%	45%		2	2.3

Les questions de contrôle de la version courte PEP sont colorées en bleu clair.

La question de contrôle *B2* « Dans quelle mesure les expressions techniques et les abréviations utiles à la compréhension de l'expertise sont-elles expliquées dans un langage courant ? » du critère de « clarté » n'est pas prévue pour la version courte PEP, mais a montré dans cette étude pilote une proportion supérieure à la moyenne de 25% de lacunes manifestes et de 35% de lacunes graves lors de l'évaluation consensuelle.

La question de contrôle *E6* « L'évaluation de la cohérence et de la plausibilité est-elle argumentée par l'expert de manière compréhensible ? » du critère « qualité professionnelle » a montré un pourcentage élevé de 30% de lacunes manifestes et 15% de lacunes graves dans l'évaluation consensuelle et pourrait être recommandée pour la version courte PEP.

**Tableau 4: Fréquences des lacunes par question des expertises partielles**

	Nombre d'évaluations				Proportions en pourcentage				Médiane	Valeur moyenne
	Pas de lacune =1	Légères lacunes =2	Lacunes manifestes =3	Graves lacunes =4	Pas de lacune =1	Légères lacunes =2	Lacunes manifestes =3	Graves lacunes =4		
<b>A1</b>	64	2	4	8	82%	3%	5%	10%	1	1.4
<b>B1</b>	64	8	5	1	82%	10%	5%	1%	1	1.3
<b>B2</b>	57	9	9	3	73%	12%	12%	4%	1	1.5
<b>B3</b>	68	6	2	2	87%	6%	2%	2%	1	1.2
<b>C1</b>	73	5			94%	6%			1	1.1
<b>C2</b>	44	9	12	13	56%	12%	15%	17%	1	1.9
<b>C3</b>	53	12	10	3	68%	15%	13%	4%	1	1.5
<b>C4</b>	60	6	8	4	77%	8%	10%	5%	1	1.4
<b>C5</b>	38	21	8	11	49%	27%	10%	14%	1	1.9
<b>D1</b>	37	8	15	18	47%	10%	19%	23%	2	2.2
<b>D2</b>	50	14	2	12	64%	18%	3%	15%	1	1.7
<b>D3</b>	53	6	6	13	68%	8%	8%	17%	1	1.7
<b>D4</b>	49	8	5	16	63%	10%	6%	21%	1	1.8
<b>E1</b>	53	3	13	9	68%	4%	17%	12%	1	1.7
<b>E2</b>	69	6	2	1	88%	8%	3%	1%	1	1.2
<b>E3</b>	13	7	3	20	30%	16%	7%	47%	3	2.7
<b>E4</b>	31	14	19	14	40%	18%	24%	18%	2	2.2
<b>E5</b>	40	10	22	6	51%	13%	28%	8%	1	1.9
<b>E6</b>	35	14	18	11	45%	18%	23%	14%	2	2.1
<b>E7</b>	38	12	15	13	49%	15%	19%	17%	2	2.0
<b>F1</b>	71	5	0	2	91%	6%	0%	3%	1	1.1
<b>F2</b>	47	14	9	8	60%	18%	12%	10%	1	1.7
	<b>Vert</b>	<b>Jaune</b>	<b>Rouge</b>		<b>Vert</b>	<b>Jaune</b>	<b>Rouge</b>		<b>Médiane</b>	<b>Valeur moyenne</b>
<b>G1</b>	24	24	30		31%	31%	38%		2	2.1

Les questions de contrôle de la version courte PEP sont colorées en bleu clair.

La question de contrôle *D1* « Dans quelle mesure l'entretien et l'examen clinique sont-ils menés de manière exhaustive par rapport à la complexité de la problématique ? » du critère « exhaustivité » a montré dans l'étude pilote un pourcentage élevé de 19% de lacunes manifestes et 23% de lacunes graves dans les expertises partielles et pourrait être utilisée pour la version courte PEP.

La question de contrôle *E1* « Dans quelle mesure les connaissances médicales généralement acceptées sont-elles prises en compte conformément aux directives en vigueur ? » du critère « qualité professionnelle » a montré un pourcentage élevé de lacunes manifestes (17%) et de lacunes graves (12%) dans les expertises partielles. Les résultats de cette question de contrôle ont parlé en faveur de la haute compétence professionnelle des évaluateurs et la question s'est posée de savoir si cette question de contrôle pourrait également être recommandée pour la version courte PEP des médecins généralistes polyvalents des SMR, étant donné que ceux-ci devraient déjà tenir compte des directives en vigueur des différentes sociétés de discipline médicale.

La question de contrôle *E7* « Dans quelle mesure l'expertise décrit-elle le traitement suivi jusque-là et se prononce-t-elle sur les autres options thérapeutiques ? » du critère « qualité professionnelle » a montré un pourcentage élevé de 19% de lacunes manifestes et 17% de lacunes graves dans les expertises partielles et pourrait également être recommandée pour la version courte PEP.

En revanche, la question de contrôle *E2* « Les médecins spécialistes procèdent-ils à des évaluations en dehors de leur discipline ? » ne présente que des lacunes mineures. Il a donc été décidé de ne pas utiliser cette question pour la version courte PEP.

#### 4.2 Appréciation de l'importance des différentes questions de contrôle

Ensuite, l'importance des différentes questions de contrôle pour l'évaluation globale des expertises a été décrite (1 : important, 2 : ni l'un ni l'autre, 3 : sans importance), ainsi que l'évaluation globale (1 : insuffisant, 2 : ni l'un ni l'autre, 3 : bon), qui ont été représentées sous forme de pourcentages, respectivement calculées sous forme de médiane et de valeurs moyennes et utilisées pour les autres analyses différentielles des questions de contrôle.

**Tableau 5: Appréciation de l'importance des questions de l'évaluation consensuelle**

	Nombre d'évaluations			Proportions en pourcentage			Médiane	Valeur moyenne
	Important =1	Ni l'un ni l'autre =2	Sans importance =3	Important =1	Ni l'un ni l'autre =2	Sans importance =3		
<b>A1</b>	18	2	0	90%	10%	0%	1	1.1
<b>A2</b>	12	8	0	60%	40%	0%	1	1.4
<b>B1</b>	18	2	0	90%	0%	0%	1	1.1
<b>B2</b>	16	4	0	80%	20%	0%	1	1.2
<b>B3</b>	17	3	0	85%	15%	0%	1	1.2
<b>D3</b>	20	0	0	100%	0%	0%	1	1.0
<b>D4</b>	20	0	0	100%	0%	0%	1	1.0
<b>D5</b>	20	0	0	100%	0%	0%	1	1.0
<b>E6</b>	20	0	0	100%	0%	0%	1	1.0
<b>E7</b>	19	1	0	95%	5%	0%	1	1.1
<b>G1</b>	20	0	0	100%	0%	0%	1	1.0

Les questions de contrôle de la version courte PEP sont colorées en bleu clair.

Les questions de contrôle *E6* et *E7* sont également considérées comme importantes par tous les pairs et pourraient être recommandées pour la version courte PEP dans l'évaluation consensuelle.

Tableau 6: Appréciation de l'importance des questions des expertises partielles

	Nombre d'évaluations			Proportions en pourcentage			Médiane	Valeur moyenne
	Important =1	Ni l'un ni l'autre =2	Sans importance =3	Important =1	Ni l'un ni l'autre =2	Sans importance =3		
<b>A1</b>	36	15	25	47%	20%	33%	2	1.9
<b>B1</b>	55	12	9	72%	16%	12%	1	1.4
<b>B2</b>	43	22	11	57%	29%	14%	2	1.6
<b>B3</b>	50	15	11	66%	20%	14%	1	1.5
<b>C1</b>	35	15	26	46%	20%	34%	2	1.9
<b>C2</b>	61	11	4	80%	14%	5%	1	1.3
<b>C3</b>	55	13	8	72%	17%	11%	1	1.4
<b>C4</b>	60	9	7	79%	12%	9%	1	1.3
<b>C5</b>	67	4	5	88%	5%	7%	1	1.2
<b>D1</b>	64	9	3	84%	12%	4%	1	1.2
<b>D2</b>	33	18	25	43%	24%	33%	2	1.9
<b>D3</b>	63	9	4	83%	12%	5%	1	1.2
<b>D4</b>	61	12	3	80%	16%	4%	1	1.2
<b>E1</b>	63	10	3	83%	13%	4%	1	1.2
<b>E2</b>	49	16	11	64%	21%	14%	1	1.5
<b>E3</b>	63	6	7	83%	8%	9%	1	1.3
<b>E4</b>	68	6	2	89%	8%	3%	1	1.1
<b>E5</b>	59	14	3	78%	18%	4%	1	1.3
<b>E6</b>	66	8	2	87%	11%	3%	1	1.2
<b>E7</b>	63	11	2	83%	14%	3%	1	1.2
<b>F1</b>	52	18	6	68%	24%	8%	1	1.4
<b>F2</b>	60	10	6	79%	13%	8%	1	1.3
<b>G1</b>	75	1	0	99%	1%	0%	1	1.0

Les questions de contrôle de la version courte PEP sont colorées en bleu clair.

Les questions de contrôle C2 à C5 du critère « transparence » sont considérées par les évaluateurs comme plus importantes que la question C1 et pourraient être davantage recommandées pour la version courte PEP.

La question de contrôle D1 du critère « exhaustivité » est jugée plus importante par les évaluateurs que la question D2 et pourrait être davantage recommandée pour la version courte PEP.

Les évaluateurs considèrent que les questions de contrôle E1 et E7 du critère « qualité professionnelle » sont plus importantes que la question de contrôle E2 et pourraient être davantage recommandées pour la version courte PEP.

### 4.3 Temps nécessaire :

La troisième étape a consisté à calculer les valeurs moyennes du temps total consacré à l'évaluation de l'expertise. Il existait ici deux indications de temps différentes : d'une part pour la discipline spécialisée avec 23 questions de contrôle et d'autre part pour l'évaluation consensuelle avec 11 questions de contrôle. Il fallait s'attendre à ce que le temps consacré aux expertises très lacunaires soit beaucoup plus important que celui consacré aux très bonnes expertises. C'est pourquoi il convient de calculer le temps nécessaire pour les trois évaluations globales (1 : médiocre, 2 : ni l'un ni l'autre, 3 : bon).

La moyenne arithmétique du temps consacré à toutes les expertises partielles était de 94,9 minutes et de 94,5 minutes pour les évaluations consensuelles. Bien que les évaluations consensuelles aient été beaucoup plus courtes, il a souvent été nécessaire de consulter les différentes expertises partielles, ce qui explique le temps relativement élevé consacré à ces dernières.

Une fois l'évaluation PEP d'une expertise terminée, les évaluateurs ont été interrogés non seulement sur le temps consacré à l'expertise, mais aussi sur l'évaluation globale de l'expertise en tant que "médiocre", "bonne" ou "ni l'une ni l'autre". Le temps consacré a alors été mis en relation avec cette évaluation globale et a montré que le temps consacré à l'évaluation des expertises lacunaires était nettement plus élevé, comme le montre le tableau 7.

**Tableau 7: Temps nécessaire en fonction de l'évaluation globale**

	<b>Nombre d'expertises</b>	<b>Moyenne Temps passé en minutes</b>	<b>Écart-type en minutes</b>
Expertises partielles lacunaires	43	107.4	44.0
Expertises partielles « ni lacunaires ni bonnes »	18	76.0	26.0
« Bonnes » expertises partielles	16	82.8	25.8
Évaluations consensuelles insuffisantes	13	100.7	35.0
Évaluations consensuelles « ni lacunaires ni bonnes »	6	83	18.6
« Bonnes » évaluations consensuelles	1	80	-

L'analyse de variance a montré que le temps consacré à l'évaluation des expertises partielles différait significativement selon les trois évaluations globales (1 : médiocre, 2 : ni l'un ni l'autre, 3 : bon) présentait une différence significative ( $F=5.71$ ,  $p<0.005$ ).

Pour les quelques évaluations consensuelles, l'analyse de variance n'a pas montré de différence significative ( $F=0.76$ ,  $p=0.48$ ).

Le temps consacré différait significativement ( $F=2,67$ ,  $p<0,05$ ) selon la discipline, comme le montre le tableau 8.

Tableau 8: Temps passé en minutes par discipline

Discipline	Nbre	Moyenne	SD	Minimum	Maximum
Médecine générale / interne	20	95	29	35	150
Neurologie	18	75	28	35	120
Psychiatrie	20	95	33	45	150
Rhumatologie	16	113	59	60	265
Orthopédie	2	135	21	120	150
Total	76	95	40	35	265

#### 4.4 Différences entre les 23 questions de contrôle selon les disciplines

L'analyse de variance a montré que les moyennes des disciplines différaient significativement en ce qui concerne le degré de gravité des lacunes des questions de contrôle A1, D1, D2, D3, E1, E7 et F1. Il n'y a pas eu de classification claire des lacunes par discipline. Comme seules 10 expertises ont été évaluées dans le cadre de cette étude pilote, les raisons de ces différences significatives n'ont pas pu être clairement attribuées aux disciplines ou aux médecins spécialistes.

Tableau 9: Différences d'évaluation des lacunes selon les disciplines

No	Médecine générale / interne		Neurologie		Psychiatrie		Rhumatologie		Valeur F	p
	M	SD	M	SD	M	SD	M	SD		
A1	2.2	1.4	1.3	0.8	1.1	0.2	1.0	0.0	6.44***	0.0002
B1	1.1	0.3	1.7	0.5	1.4	0.8	1.5	1.0	1.68	0.1647
B2	1.7	1.0	1.7	0.8	1.4	0.8	1.0	0.0	2.18	0.0801
B3	1.2	0.3	1.2	0.5	1.4	0.9	1.0	0.0	1.15	0.3424
C1	1.0	0.0	1.0	0.0	1.2	0.4	1.3	0.3	1.53	0.2023
C2	2.2	1.4	2.0	1.1	1.8	1.0	1.8	1.2	0.44	0.7818
C3	1.3	0.7	1.5	0.7	2.0	0.9	1.5	1.1	1.77	0.1452
C4	1.4	0.9	1.2	0.5	1.9	1.1	1.2	0.8	1.96	0.1096
C5	1.8	1.1	1.8	1.0	2.0	1.2	2.1	1.2	0.23	0.9199
D1	3.0	1.2	1.8	1.1	1.9	1.1	1.9	1.2	3.83**	0.0071
D2	1.6	0.9	1.8	1.2	1.5	0.5	1.9	1.3	4.64**	0.0022

D3	1.2	0.7	1.5	1.0	<b>2.5</b>	<b>1.4</b>	1.6	1.0	4.06**	0.0051
D4	1.7	1.2	1.6	1.2	2.1	1.3	2.1	1.2	0.74	0.5646
E1	1.6	1.2	1.6	1.1	<b>2.8</b>	<b>1.0</b>	1.0	0.0	9.47***	<.0001
E2	1.2	0.5	1.3	0.8	1.2	0.5	1.1	0.3	0.42	0.7953
E3_a	1.5	0.5	1.5	0.5	1.6	0.5	1.4	0.5	0.59	0.6707
E3	2.6	1.4	3.4	0.9	3.1	1.2	1.8	1.3		
E4	1.8	1.2	2.1	1.1	2.6	1.1	2.5	1.1	1.57	0.1924
E5	2.0	1.2	1.7	1.0	2.2	1.0	1.9	1.1	0.51	0.7279
E6	1.7	1.0	2.2	1.3	2.5	1.0	2.1	1.2	1.55	0.1985
E7	1.3	0.7	2.0	1.1	<b>2.4</b>	<b>1.1</b>	<b>2.5</b>	<b>1.3</b>	3.73**	0.0082
F1	1.0	0.0	1.5	1.0	1.0	0.0	1.1	0.3	3.43*	0.0127
F2	1.8	1.2	1.7	1.1	1.7	0.9	1.8	1.1	0.05	0.9946
G	2.0	0.8	2.0	0.9	2.4	0.8	1.9	0.9	1.04	0.3943

\*\*\* Niveau de signification avec une probabilité d'erreur de  $p < 0.001$

\*\* Niveau de signification avec une probabilité d'erreur de  $p < 0.01$

\* Niveau de signification avec une probabilité d'erreur de  $p < 0.05$

L'analyse de variance a montré que les valeurs moyennes des disciplines différaient de manière significative en ce qui concernait l'évaluation de l'importance des questions de contrôle A1, B3, C1, C2, C3, D1, D2, E1, E2, E5 et E7. Les médecins généralistes, en particulier, ont estimé que toutes les questions de contrôle étaient importantes.

Tableau 10: Différences d'évaluation de l'importance selon les disciplines

No	Médecine générale / interne		Neurologie		Psychiatrie		Rhumatologie		Valeur F	p
	M	SD	M	SD	M	SD	M	SD		
A1	1.2	0.4	<b>2.3</b>	<b>0.8</b>	<b>2.0</b>	<b>0.9</b>	<b>2.1</b>	<b>1.0</b>	6.26***	0.0002
B1	1.1	0.2	1.5	0.6	1.7	0.9	1.5	0.8	2.42	0.0566
B2	1.3	0.5	1.9	0.7	1.8	0.9	1.5	0.8	2.38	0.0599
B3	1.1	0.2	1.8	0.7	1.8	0.9	1.4	0.8	4.31**	0.0035
C1	1.2	0.4	<b>2.5</b>	<b>0.5</b>	1.7	0.8	<b>2.3</b>	<b>1.0</b>	11.3***	<.0001
C2	1.0	0.0	1.4	0.5	1.6	0.8	1.1	0.3	3.52*	0.0112
C3	1.0	0.0	1.5	0.6	1.7	0.9	1.4	0.7	3.21*	0.0177
C4	1.0	0.0	1.4	0.6	1.5	0.8	1.4	0.7	1.87	0.1260
C5	1.0	0.0	1.8	1.1	1.3	0.6	1.4	0.8	1.78	0.1417
D1	1.0	0.0	1.8	1.4	1.0	0.0	1.4	0.7	4.04**	0.0052
D2	1.2	0.4	<b>2.5</b>	<b>0.5</b>	1.7	0.8	<b>2.3</b>	<b>0.9</b>	14.2***	<.0001

D3	1.0	0.0	1.3	0.6	1.2	0.5	1.5	0.7	2.25	0.0720
D4	1.0	0.0	1.4	0.6	1.3	0.6	1.3	0.6	2.13	0.0862
E1	1.0	0.0	1.4	0.5	1.1	0.2	1.5	0.8	3.97**	0.0058
E2	1.0	0.0	1.7	0.7	1.7	0.8	1.8	1.0	4.57**	0.0024
E3_a	1.0	0.0	1.3	0.7	1.4	0.8	1.5	0.8	0.59	0.6707
E3	1.0	0.0	1.3	0.7	1.2	0.6	1.5	0.8	1.75	0.1497
E4	1.0	0.0	1.2	0.4	1.2	0.4	1.3	0.3	0.92	0.4590
E5	1.0	0.0	1.5	0.6	1.2	0.4	1.3	0.6	3.05*	0.0224
E6	1.0	0.0	1.1	0.3	1.4	0.6	1.3	0.6	1.93	0.1150
E7	1.0	0.0	<b>1.6</b>	<b>0.6</b>	1.1	0.2	<b>1.3</b>	<b>0.6</b>	5.22***	0.0010
F1	1.2	0.4	1.6	0.6	1.5	0.7	1.5	0.8	1.50	0.2123
F2	1.1	0.2	1.2	0.4	1.5	0.7	1.5	0.9	1.86	0.1270
G	1.0	0.0	1.0	0.2	1.0	0.0	1.0	0.0	0.80	0.5312

\*\*\* Niveau de signification avec une probabilité d'erreur de  $p < 0.001$

\*\* Niveau de signification avec une probabilité d'erreur de  $p < 0.01$

\* Niveau de signification avec une probabilité d'erreur de  $p < 0.05$

#### 4.5 Analyse factorielle des 23 questions de contrôle

Quatre analyses factorielles ont été effectuées, pour les questions de contrôle des expertises partielles et pour les questions de contrôle des évaluations consensuelles, ainsi que pour les évaluations des lacunes et les évaluations de l'importance. Les quatre analyses factorielles des analyses en composantes principales ont montré que toutes les questions constituaient un seul et même facteur et qu'aucune question ne se « démarquait » et ne devait être éliminée, car elle mesurait une construction nettement différente. Même la question de contrôle A1, de nature très formelle, avait une charge encore relativement haute sur le premier facteur, ce qui a indiqué qu'aucune question de contrôle ne constituait un propre facteur et ne mesurait quelque chose de totalement différent.

#### **4.6 Analyse de régression des 23 questions de contrôle comme prédicteurs de l'évaluation globale**

L'évaluation de la gravité des lacunes à l'aide des 23 questions de contrôle des expertises partielles et des 10 questions de contrôle des évaluations consensuelles était plus ou moins fortement liée aux 97 évaluations globales des expertises : L'analyse de régression a calculé l'intensité de ces liens et a permis de voir quelles questions devaient absolument être conservées comme prédicteurs de l'évaluation globale : Cela n'a été le cas de manière hautement significative que pour la question de contrôle *G* de la reproductibilité de la chaîne d'argumentation. La question de contrôle *E3* « L'expert médical compétent tient-il compte des diagnostics complémentaires dans son évaluation ? » dans les expertises partielles et la question de contrôle *D4* « Dans quelle mesure les réponses aux questions médico-assurantielles posées par le mandant, y compris aux questions complémentaires, sont-elles complètes ? » dans l'évaluation consensuelle ont été légèrement significatives avec une probabilité d'erreur inférieure à 5%.

Toutes les autres questions de contrôle n'ont pas été significatives dans l'analyse de régression et n'expliquaient pas beaucoup plus de variance que celles-ci pour les questions de contrôle. Le critère majeur « Compréhensibilité de l'expertise » était un critère de qualité dont la réalisation était considérée comme décisive et absolument nécessaire. Dans le critère de qualité majeur, l'ensemble de l'expertise a été considéré dans son contexte et la compréhensibilité de l'expertise a été évaluée à l'aide de la chaîne d'argumentation. Cela s'est à présent révélé empiriquement.

## **5. Discussion des résultats**

Les objectifs de l'étude pilote ont été formulés au chapitre 2 et sont maintenant discutés en fonction des résultats :

### **5.1 Acquisition des évaluateurs qualifiés (faisabilité)**

Pour cette étude pilote, 8 médecins spécialistes expérimentés ont pu être trouvés conformément aux exigences de l'appel d'offres (voir annexe I) et qui sont prêts à continuer à travailler en tant qu'évaluateurs indépendants. L'étude pilote a été réalisée uniquement en allemand. En vue de la réalisation de la version complète PEP en 2024 sur un échantillon représentatif de toutes les expertises réalisées en 2023, il faudrait également recruter des évaluateurs francophones ou bilingues. Environ 5% de l'échantillon d'expertises sera en italien. Presque tous les médecins spécialistes tessinois ont fait leurs études dans une université francophone ou germanophone et sont généralement parfaitement bilingues.

### **5.2 Développement et mise en œuvre d'une première formation pour les évaluateurs**

La formation des 8 médecins spécialistes s'est déroulée entre le 4 avril, avec une introduction en ligne, et fin mai, avec la réunion finale selon le plan de l'annexe II. La formation en présentiel s'est avérée très utile, car elle a permis d'aborder concrètement l'évaluation PEP d'une expertise préalablement anonymisée et remise. Elle a permis de clarifier la compréhension de certaines questions de contrôle et leur délimitation par rapport aux autres questions de contrôle, ainsi que les définitions des critères et du degré de gravité des lacunes (ancrages) respectives. Lors de l'analyse d'une deuxième expertise anonymisée qui a été réalisée en ligne, les différences d'appréciation de certaines questions de contrôle de l'évaluation PEP ont été clarifiées avec les évaluateurs. Les descriptions du contenu, tant des critères que des définitions des ancrages se sont révélées très utiles et suffisamment précises. Les évaluateurs n'ont pas demandé de reformuler ou de compléter les critères ou les définitions des ancrages.

### **5.3 Test de la forme de saisie électronique pour la version complète PEP**

Dès l'introduction en ligne de la formation du 4 avril, l'outil en ligne pour la saisie électronique de la version complète PEP (voir annexe IV) a été présenté, puis l'accès a été rendu possible par e-mail avec un mot de passe personnel. Seule une des huit personnes a rencontré des difficultés d'accès en ligne, qui ont toutefois pu être résolues avec l'administrateur.

La plateforme de saisie électronique en ligne de la version complète PEP a fonctionné sans problème pendant toute l'étude pilote, à l'entière satisfaction de tous les participants.

Les expertises à évaluer ont été anonymisées par le secrétariat spécialisé de la COQEM avec Adobe Pro et transmises via la connexion sécurisée de l'Office fédéral de l'informatique via le service de transfert de fichiers. Ce processus a également fonctionné de manière irréprochable.

#### **5.4 Saisie plus précise du temps nécessaire à la version complète PEP**

Dans le cadre de l'étude pilote SMR réalisée en 2021 par le Dr Hermelink, le temps moyen consacré à la version courte PEP avec dix questions de contrôle était de 91 minutes pour une expertise spécialisée, respectivement une expertise partielle, la seule lecture de l'expertise étant en moyenne de 48 minutes. Selon le chapitre 4.3, la moyenne arithmétique du temps consacré à l'ensemble des expertises partielles était de 94,9 minutes dans cette étude pilote et n'était donc pas beaucoup plus élevée avec 24 questions de contrôle qu'avec la version courte PEP de 10 questions de contrôle.

Bien que l'expérience acquise avec l'évaluation PEP puisse permettre une réduction potentielle du temps nécessaire, l'effort reste considérable, en particulier pour les expertises lacunaires pour lesquelles les lacunes doivent être formulées de manière relativement détaillées et compréhensibles. L'étude pilote a montré que la durée moyenne de l'évaluation PEP était de 107 minutes pour les expertises partielles lacunaires et de 100 minutes pour les évaluations consensuelles, contre 80 minutes pour les bonnes expertises.

Pour une expertise pluridisciplinaire avec quatre disciplines, il faut compter en moyenne environ 450 minutes, soit 7,5 heures, ce qui correspond à une journée de travail complète, si la qualité du contenu est analysée de manière sérieuse et professionnelle et si les experts doivent recevoir un feed-back qualifié.

Il semble plausible que les conditions générales puissent avoir une influence significative sur la réalisation des évaluations PEP, peut-être même plus importante que le nombre de questions de contrôle. Si les médecins disposent de suffisamment de temps et que les pourcentages de postes sont adaptés de manière appropriée lors de l'introduction de la version courte PEP à l'échelle nationale pour une évaluation uniforme de toutes les expertises, cela pourrait favoriser un examen approfondi et complet du contenu des expertises externes.

Afin d'évaluer les 11'000 expertises annuelles, dont environ 5'000 expertises pluridisciplinaires, 2'000 expertises bidisciplinaires et 4'000 expertises monodisciplinaires, il faudrait estimer le temps de travail d'environ 34 médecins à plein temps par années. Comme les médecins SMR doivent déjà lire et évaluer les expertises reçues, nous estimons que l'introduction systématique de la PEP dans tous les SMR des offices AI en Suisse nécessitera une augmentation moyenne du personnel par environ 8 postes équivalent plein temps, afin d'assurer la réalisation des évaluations PEP dans de bonnes conditions.

### **5.5 Identification des questions de contrôle pertinentes dans la version courte PEP**

Le chapitre 4.1 a présenté les évaluations de la gravité des lacunes pour chaque question de contrôle (voir tableaux 3 et 4) et le chapitre 4.2 est consacré à l'importance accordée par les évaluateurs à chaque question de contrôle pour l'évaluation globale de la qualité d'une expertise (voir tableaux 5 et 6). Dans ces quatre tableaux, les questions de contrôle de la version courte PEP sont mises en évidence en bleu clair afin d'identifier rapidement s'il existe des questions de contrôle pertinentes dans les 14 questions de contrôle omises entre la version complète PEP (24) et la version courte PEP (10) ainsi que dans les cinq questions de contrôle omises entre l'évaluation consensuelle de la version complète PEP (11) et l'évaluation consensuelle de la version courte PEP (6).

Il est ressorti de cette étude pilote que les questions de contrôle *D1*, *E1* et *E7*, jusqu'ici omises dans la version courte PEP, ont permis d'identifier davantage de lacunes graves que les questions *D2* et *E2* des expertises partielles. La formulation exacte des questions de contrôle est présentée à l'annexe III.

Lors de l'évaluation consensuelle, la question de contrôle *B2* a permis d'identifier nettement plus de lacunes graves (valeur moyenne de 2.7) que dans les expertises partielles de médecine générale par les mêmes évaluateurs (valeur moyenne de 1.7), ce qui a indiqué que les exigences en matière de clarté étaient beaucoup plus élevées lors de l'évaluation consensuelle que lors des expertises partielles.

De même, en ce qui concerne l'évaluation de l'importance, les questions de contrôle *D1*, *E1* et *E7* ont été jugées comme plus importantes que la question de contrôle *E2*. De plus, les évaluateurs ont considéré que les questions de contrôle *C2* à *C5* du critère « transparence » comme plus importantes que la question de contrôle *C1*.

Lors de sa séance du 25 octobre 2023, la commission a discuté de ce rapport. En ce qui concerne la question de contrôle E1, il a été remarqué que les médecins des SMR ne couvrent souvent pas toutes les disciplines, mais sont souvent des médecins généralistes, raison pour laquelle la question de contrôle E1 ne doit pas être incluse dans la version courte PEP. La commission s'est prononcée en faveur du maintien de la question de contrôle D2 dans la version courte PEP, car elle a considéré que le codage des diagnostics selon le DSM ou la CIM était un critère de qualité important. Enfin, la commission a décidé d'ajouter les questions de contrôle D1 et E7 et de supprimer E2 pour la version courte PEP.

### **5.6 Identification des questions de contrôle selon la gravité des lacunes**

Au début de l'étude pilote, la question suivante a également été soulevée : « Y a-t-il éventuellement des questions de contrôle décisives ? Y a-t-il, du point de vue des évaluateurs, des questions qui sont particulièrement importantes pour l'évaluation de la qualité ? » En outre, il a été examiné si ces questions de contrôle étaient incluses dans la version courte PEP. Comme mentionné ci-dessus, les lacunes de toutes les questions de contrôle ont été présentées en détail dans les tableaux 3 et 4. La question de contrôle D5 « Dans les expertises interdisciplinaires, existe-t-il une véritable appréciation intégrative ? » est particulièrement frappante : 55% des évaluations consensuelle ont révélé de graves lacunes ; cette question est déjà incluse dans la version courte actuelle PEP. En ce qui concerne les expertises partielles, la question de contrôle E3 « L'expert médical compétent tient-il compte des diagnostics complémentaires dans son évaluation ? » s'est distingué par 47% de graves lacunes ; elle figure déjà dans la version courte actuelle PEP. Du point de vue des évaluateurs, presque toutes les questions de contrôle ont été jugées comme importantes, mais les questions de nature formelle des critères A, B et C ont semblé un peu moins importantes.

### **5.7 Décider si et comment la version complète PEP doit être raccourcie**

Du point de vue des évaluateurs, aucune question de contrôle n'a été considérée comme peu importante pour l'évaluation de la qualité des expertises médicales.

Il y avait cependant trois questions de contrôle plutôt de nature formelle A1, C1 et D2, qui n'ont été considérées comme « importantes » qu'à 30%, voire sans importance à 30% (voir tableau 6) :

A1 « Dans quelle mesure les prescriptions structurelles, y compris celles sur l'évaluation consensuelle, sont-elles respectées ? » du critère « structure formelle » est considérée comme « importante » à 36% et « sans importance » à 33%.

C1 « Le temps d'examen (date, en heures de - à) est-il documenté ? » du critère « transparence » est considérée comme « important » à 35% et « sans importance » à 34%.

*D2* « Dans quelle mesure les diagnostics sont-ils basés sur un système reconnu et codés selon ce système ? » du critère « exhaustivité » est considérée comme « importante » à 43% et « sans importance » à 33%.

L'argument en faveur du maintien de ces trois questions de contrôle moins importantes a été mentionné lors des discussions de la commission en 2022, à savoir que cette évaluation formelle peut être effectuée très rapidement et facilement et qu'elle ne représente pas une grande réduction de la charge de travail. Etant donné que cette évaluation formelle ne nécessite pas de connaissances spécifiques, il a également été proposé que l'évaluation de ces critères formels puisse être déléguée à un professionnel formé.

En ce qui concerne la décision de la commission de juin 2023 d'introduire l'indicateur de qualité « La durée de l'examen est-elle appropriée ? », la condition préalable est bien entendu que la durée de l'examen soit indiquée et lue par l'évaluateur. Dans ce contexte, la question de contrôle *C1* peut être maintenue. Cependant, l'évaluation de l'adéquation de la durée de l'examen requiert des compétences du médecin spécialiste et devrait être reformulée dans le cadre du critère *E*. D'autre part, un temps d'examen excessivement court révèle des lacunes dans l'exhaustivité de l'anamnèse et la clarification des résultats.

Lors de sa séance du 25 octobre 2023, la COQEM a décidé d'utiliser toutes les questions de contrôle pour la version complète PEP et de ne pas réduire le catalogue de questions.

## **5.8 Limitations de l'étude pilote PEP**

Dans cette étude pilote, dix expertises ont été évaluées et que, par conséquent, les différences significatives de l'évaluation des lacunes et de l'appréciation de l'importance ne pouvaient pas être clairement attribuées aux expertises partielles ou aux médecins spécialistes. De plus, on a délibérément renoncé à calculer la fiabilité entre les évaluateurs afin de mettre l'accent initial sur les aspects qualitatifs mentionnés.

Malgré ces limites, l'étude offre de précieuses conclusions : L'utilisation de la version complète PEP et de la version courte PEP a montré que l'instrument fonctionne en principe, même si l'investissement en temps et en argent pour sa mise en œuvre est considérable. Grâce à une formation à grande échelle, on peut toutefois s'attendre à une augmentation de l'efficacité dans les applications futures.

Un autre aspect est l'identification des goulots d'étranglement chez les évaluateurs (ici aussi en ce qui concerne les différentes régions linguistiques), en particulier chez les psychiatres, qui peuvent entraîner des retards dans le traitement de la PEP en raison de leur charge de travail élevée, ce qui ralentit (globalement) le processus.

## 6. Conclusions et recommandations

L'étude pilote a été menée avec succès, avec une mise en œuvre efficace de la formation et une réalisation réussie. La plateforme électronique en ligne a fait ses preuves et a fonctionné sans difficultés techniques notables. Les évaluateurs ont considéré les définitions des ancrages des questions de contrôle claires et utiles, avec seulement quelques ajustements mineurs nécessaires dans le cadre de la formation.

En ce qui concerne l'adaptation du catalogue de questions, bien que peu de questions de contrôle formelles aient été identifiées, ne permettant pas une réduction significative du temps de traitement, la commission a décidé, lors de sa séance du 25 octobre 2023, de conserver toutes les questions de contrôle. En outre, la version courte PEP est enrichie par l'ajout des questions de contrôle *D1* et *E7* et rationalisée par la suppression de la question de contrôle *D2*. Concernant la mise en place future de la PEP : le temps nécessaire à l'évaluation d'une seule expertise pluridisciplinaire reste considérable. Pour l'évaluation complète d'un échantillon aléatoire de 100 expertises pluridisciplinaires, il faut prévoir un coût total d'environ 200'000 CHF. En outre, le coût de la réalisation d'une PEP supplémentaire d'environ 50 expertises pluridisciplinaires dans le cadre d'audits coûterait environ 100'000 CHF.

Suite à la décision de la commission de collecter six indicateurs de qualité, il est prévu d'adapter les questions de contrôle correspondantes A2, C1, D3, E5 et E6 de la PEP afin d'y inclure la formation des définitions d'ancrage du degré de gravité des lacunes. Ces cinq questions de contrôle doivent en outre être relevées dans la version courte PEP.

Par conséquent, la commission a prévu les étapes suivantes :

- La publication du rapport pilote et du catalogue de questions de la version complète PEP, afin de garantir la transparence et l'accessibilité des résultats ;
- La recommandation d'utiliser la version complète PEP pour l'assurance qualité interne des organismes d'évaluation et à des fins de formation des nouveaux évaluateurs, afin de promouvoir en permanence les compétences dans ce domaine ;
- La recommandation de l'utilisation de la version courte PEP à l'attention de l'Office fédéral des assurances sociales (OFAS) et des Offices AI, afin de permettre une évaluation standardisée et efficace
- L'engagement de évaluateurs supplémentaires par la COQEM afin de pouvoir évaluer de manière adéquate les expertises en français et en italien, ce qui augmente la couverture linguistique ;
- L'évaluation d'un échantillon aléatoire représentatif de 100 expertises par an, afin de garantir un contrôle et une amélioration continus de la qualité ;
- La réalisation d'audits dans le cadre de procédures d'autorisation ou en cas d'indications de lacunes graves, afin de garantir l'intégrité des processus d'expertise.

Tableau 11: Aperçu des questions des versions complète et courte PEP

Critère	No.	Version complète PEP			Version courte PEP		
		Expertise monodisc.	Expertise pluridisc.	Consensuelle	Expertise monodisc.	Expertise pluridisc.	Consensuelle
Forme	A1	A1	A1	A1			
	A2	A2		A2	A2		A2
Compréhension	B1	B1	B1	B1			
	B2	B2	B2	B2			
	B3	B3	B3	B3			
Transparence	C1	C1	C1		C1	C1	
	C2	C2	C2				
	C3	C3	C3				
	C4	C4	C4				
	C5	C5	C5				
Complétude	D1	D1	D1		D1	D1	
	D2	D2	D2				
	D3	D3	D3	D3	D3	D3	D3
	D4	D4	D4	D4	D4	D4	D4
	D5			D5			D5
Qualité professionnelle	E1	E1	E1				
	E2	E2	E2				
	E3	E3	E3		E3	E3	
	E4	E4	E4		E4	E4	
	E5	E5	E5		E5	E5	
	E6	E6	E6	E6	E6	E6	E6
	E7	E7	E7	E7	E7	E7	
Economicité	F1	F1	F1				
	F2	F2	F2				
Argumentation							

Remarque : les questions surlignées en couleur saumon ont été nouvellement développées par des experts CH. Les questions non colorées ont été reprises de la version DRV (DE).

## **Annexe I : Profil d'exigences des évaluateurs**

### Evaluateurs I (évaluateur spécialiste en médecine interne), procédure PEP

- Exigences selon l'art. 7m OPGA, al. 1, let. a, b, c :
  - a) Titre de formation postgrade selon l'art. 2, al. 1, let. b et c (ordonnance du 27 juin 2007 sur les professions médicales) ;
  - b) Inscription au registre MedReg conformément à l'article 51, paragraphe 1 (loi sur les professions médicales du 23 juin 2006) ;
  - c) une autorisation de pratiquer valable en Suisse.
- Au moins 10 ans d'expérience en médecine des assurances et/ou en expertise.
- Médecin spécialiste en médecine interne générale ou formation comparable.
- Certificat de l'Association Médecine d'Assurance Suisse (Swiss Insurance Medicine, SIM) ou médecins chefs et médecins dirigeants dans des hôpitaux universitaires et des hôpitaux de catégorie A.
- Très bonnes connaissances d'au moins deux langues nationales.
- Disposition à collaborer pendant au moins trois ans (si possible jusqu'à 12 ans) en tant qu'évaluateurs PEP sur mandat de la COQEM.
- Temps de travail de 30 à 60 heures par an, organisation flexible du temps avec des compensations financières conformes au marché.

### Evaluateurs II (évaluateur spécialisé), procédure PEP

- Exigences selon l'art. 7m OPGA, al. 1, let. a, b, c :
  - d) Titre de formation postgrade selon l'art. 2, al. 1, let. b et c (Ordonnance sur les professions médicales du 27 juin 2007) ;
  - e) Inscription au registre conformément à l'article 51, paragraphe 1 (loi sur les professions médicales du 23 juin 2006) ;
  - f) Une autorisation de pratiquer valable en Suisse.
- Au moins 10 ans d'expérience en médecine des assurances et/ou en expertise.
- Médecins spécialistes en psychiatrie et psychothérapie, neurologie, rhumatologie.
- Certificat de l'Association Médecine d'Assurance Suisse (Swiss Insurance Medicine, SIM) ou médecins chefs et médecins dirigeants dans des hôpitaux universitaires et des hôpitaux de catégorie A.
- Très bonnes connaissances d'au moins deux langues nationales.
- Disposition à collaborer pendant au moins trois ans (si possible jusqu'à 12 ans) en tant qu'évaluateurs PEP sur mandat de la COQEM.
- Temps de travail de 30 à 60 heures par an, organisation flexible du temps avec des compensations financières conformes au marché.

## Annexe II : Mise en œuvre de la formation des évaluateurs

*Participants : 8 / Responsable du cours : 1 (+1 assistant)*

- 1) Envoi de la documentation (après réception du contrat signé) : étude du manuel PEP et de catalogue des questions de contrôle (environ 2 heures).
- 2) Formation en ligne (4 avril 2023) sur le concept PEP, le manuel, les questions de contrôle PEP, ainsi que l'utilisation de l'outil en ligne (environ 3 heures).
- 3) Envoi d'un rapport d'expertise pluridisciplinaire anonymisé par Inca-Mail : étude préalable du rapport (env. 2 heures).
- 4) Sur place (17 avril 2023 à Berne\* et 22 mai 2023 à Zurich) : Discussion des lacunes de l'expertise et remplissage des questions de contrôle par équipes de deux et discussion de certaines questions de contrôle sur la base des lacunes constatées avec le responsable du cours (env. 3 heures).
- 5) Envoi d'une expertise pluridisciplinaire anonymisée par Inca-Mail, création de l'accès en ligne et réalisation de l'évaluation PEP de l'expertise pluridisciplinaire via l'outil en ligne (env. 4 heures).
- 6) Feedback écrit (liste Excel) des 8 évaluations avec la liste de toutes les lacunes constatées : discussion en ligne des divergences entre les différentes questions de contrôle par sous-discipline et l'évaluation consensuelle par les médecins généralistes (env. 2 heures) : Pratiquement toutes les différences ont pu être clarifiées et résolues. Les entretiens de clarification ont permis de mieux comprendre le contenu des questions de contrôle et les définitions des ancrages. Les spécialistes n'ont pas demandé de reformulations ou de précisions supplémentaires.

Temps nécessaire environ 16 heures

Estimation des coûts : 28'160 CHF

Le projet pilote est prévu en allemand avec 8 participants (4 disciplines).

## Annexe III : catalogue de questions PEP pour l'étude pilote 2023

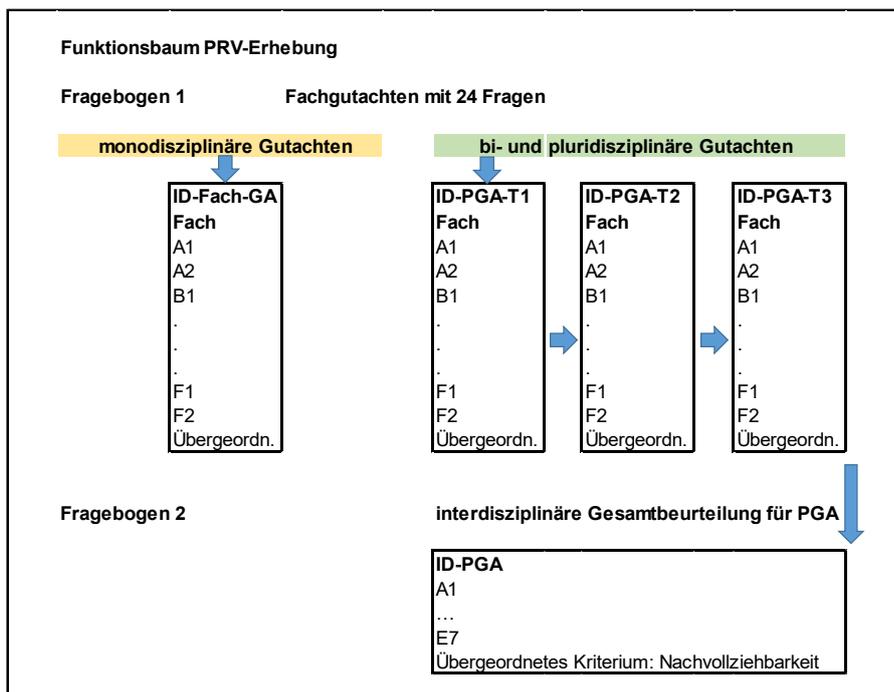
<b>Critère A : <i>structure formelle</i></b>	
Question A1	Dans quelle mesure les prescriptions structurelles, y compris celles sur l'évaluation consensuelle, sont-elles respectées ?
Question A2	Les délais entre l'examen de l'assuré et la réception de l'expertise par l'AI sont-ils appropriés ?
<b>Critère B : <i>clarté</i></b>	
Question B1	L'expertise est-elle rédigée correctement sur le plan linguistique et de façon univoque ?
Question B2	Dans quelle mesure les expressions techniques et les abréviations utiles à la compréhension de l'expertise sont-elles expliquées dans un langage courant ?
Question B3	Dans quelle mesure la terminologie médico-assurantielle est-elle correctement employée ?
<b>Critère C : <i>transparence</i></b>	
Question C1	Le temps d'examen (date, en heures de ... à ...) est-il documenté ?
Question C2	Dans quelle mesure la source, la date et les observations clés relatives aux informations médicales sont-elles identifiables ?
Question C3	Dans quelle mesure les méthodes ayant servi à recueillir les résultats médicaux sont-elles mentionnées ?
Question C4	Dans quelle mesure les sources d'information utilisées par l'expert pour recueillir ses propres résultats d'examen se distinguent-elles ?
Question C5	Dans quelle mesure la base sur laquelle repose l'évaluation des résultats d'examen est-elle présentée ?
<b>Critère D : <i>exhaustivité</i></b>	
Question D1	Dans quelle mesure l'entretien et l'examen clinique sont-ils menés de manière exhaustive par rapport à la complexité de la problématique ?
Question D2	Dans quelle mesure les diagnostics sont-ils basés sur un système reconnu et codés selon ce système ?
Question D3	Dans quelle mesure les divergences significatives entre l'évaluation médico-assurantielle et d'autres rapports antérieurs concernant les capacités fonctionnelles sont-elles décrites et évaluées ?
Question D4	Dans quelle mesure les réponses aux questions médico-assurantielles posées par le mandant, y compris aux questions complémentaires, sont-elles complètes ?
<b>Critère E : <i>qualité professionnelle</i></b>	
Question E1	Dans quelle mesure les connaissances médicales généralement acceptées sont-elles prises en compte conformément aux directives en vigueur ?
Question E2	Les médecins spécialistes procèdent-ils à des évaluations en dehors de leur discipline ?

Question E3	L'expert médical compétent tient-il compte des diagnostics complémentaires dans son évaluation ?
Question E4	Les déficits fonctionnels et les résultats médicaux documentés sont-ils mis en lien avec les diagnostics posés ?
Question E5	Les facteurs personnels et psychosociaux sont-ils pris en compte ?
Question E6	L'évaluation de la cohérence et de la plausibilité est-elle argumentée par l'expert de manière compréhensible ?
Question E7	Dans quelle mesure l'expertise décrit-elle le traitement suivi jusque-là et se prononce-t-elle sur les autres options thérapeutiques ?
<b>Critère F : <i>économicit�</i></b>	
Question F1	Dans quelle mesure les m�thodes de diagnostic utilis�es sont-elles appropri�es et n�cessaires pour r�pondre � la question m�dico-assurantielle pos�e ?
Question F2	Dans quelle mesure les m�thodes de diagnostic utilis�es sont-elles suffisantes pour r�pondre � la question m�dico-assurantielle pos�e ?
<b>Crit�re majeur (feux tricolores) : <i>compr�hensibilit� de l'expertise</i></b>	
Crit�re majeur	Veillez, � l'aide du sch�ma suivant, v�rifier la compr�hensibilit� de l'expertise sur le plan argumentatif.

# Annexe IV : Configuration requise : Plateforme en ligne pour la PEP

## 1. Exigences sommaires

Le questionnaire « expertise spécialisée » doit être rempli par tous les experts en indiquant la discipline concernée. Le questionnaire « évaluation globale interdisciplinaire » doit être rempli uniquement par le généraliste de médecine interne générale mandaté pour les expertises pluridisciplinaires. En cas d'expertise bidisciplinaire, c'est l'un des deux spécialistes qui procède à l'évaluation globale.



Il existe une liste d'évaluateurs formés qui sont responsables des expertises spécifiques bi- ou pluridisciplinaires. Les médecins spécialistes en médecine interne générale peuvent remplir aussi bien leur expertise spécialisée que l'« évaluation globale interdisciplinaire » lors des expertises bi- ou pluridisciplinaires en tant que généralistes.

## 2. Description fonctionnelle succincte

Login avec mot de passe par évaluateur (CH-Login).

L'outil de saisie en ligne sert aux évaluateurs externes à évaluer de manière sûre les expertises pluri-, bi- et monodisciplinaires anonymisées. Une expertise monodisciplinaire est réalisée dans une discipline spécialisée. Les expertises pluri- et bidisciplinaires se composent de plusieurs expertises spécifiques de différentes disciplines et d'une évaluation consensuelle interdisciplinaire.

Il existe deux types de catalogues de questions de contrôle : un catalogue pour les expertises monodisciplinaires avec 24 questions ou les expertises partielles avec 23 questions et un catalogue pour l'évaluation interdisciplinaire consensuelle avec 11 questions.

Les questions se composent d'instructions détaillées (environ 2'000 caractères) et de réponses sur une échelle à 4 niveaux (1 = Pas de lacune, 2 = Lacunes légères, 3 = Lacunes manifestes et 4 = Graves lacunes) et d'une saisie de texte intégral pour les niveaux 2 à 4 par question (description des lacunes).

Les réponses sont enregistrées successivement dans une banque de données à laquelle l'évaluateur concerné peut accéder à tout moment (read and write concernant ses évaluations).

Le secrétariat spécialisé de la COQEM a accès à la banque de données et peut extraire et évaluer les résultats des évaluateurs.

L'outil a d'abord été créé en allemand et devra également être développé en français et en italien pour les expertises et les utilisateurs francophones et italophones.

### **3. Etapes / règles de traitement**

1) Saisie et contrôle d'accès d'un évaluateur externe.

2) L'évaluateur ne peut voir et modifier que ses évaluations (read and write), mais pas celles des autres évaluateurs. Un évaluateur peut évaluer plusieurs expertises.

3) Chaque expertise a un numéro d'identification et doit être saisie et enregistrée en premier lieu. Chaque réponse doit être enregistrée successivement dans une base de données et l'évaluateur doit pouvoir lire et modifier ses réponses précédentes à tout moment.

4) En règle générale, une expertise, qu'elle soit pluridisciplinaire ou monodisciplinaire, est évaluée à chaque fois par un évaluateur. Dans certains cas, les différentes expertises partielles spécifiques à une discipline peuvent être évaluées par un évaluateur de la discipline correspondante.

5) L'évaluation d'une expertise ne peut être achevée que lorsque toutes les questions ont été saisies sur une échelle à 4 niveaux (1. Pas de lacune, 2. Lacunes légères, 3. Lacunes manifestes et 4. Graves lacunes) et avec une saisie minimale du texte intégral (description des lacunes) pour les niveaux 2 à 4 par question : dernière étape de contrôle.

6) Dans une première phase, le système LOGIN est basé sur le nom d'utilisateur de l'adresse e-mail et un mot de passe. Dans une deuxième phase, le processus de LOGIN doit pouvoir se faire très facilement par 2 facteurs (e-mail et SMS) via CH-LOGIN, dès que l'OFIT de la société Equalizer aura créé l'interface.

7) Dans le cas des expertises pluridisciplinaires, il y a un nombre variable de sous-expertises et une évaluation consensuelle interdisciplinaire. C'est pourquoi il est plus simple que le nombre d'expertises partielles soit contrôlé par l'évaluateur en tant qu'utilisateur : L'évaluateur sait combien d'expertises partielles doivent d'abord être revues en tant qu'expertises spécialisées avant que l'évaluation consensuelle interdisciplinaire ne soit remplie : Au début, l'évaluateur décide toujours s'il veut remplir le questionnaire spécifique à la discipline ou l'évaluation consensuelle interdisciplinaire.

- **Input** : Les réponses des évaluations en ligne sont écrites dans une base de données et enregistrées par évaluateur et par expertise.
- **Output** : La banque de données contenant les évaluations PEP des évaluateurs.
- **Critères d'acceptation** : Fiabilité, sécurité des données, facilité d'utilisation.