LIÈGE université Library

https://lib.uliege.be https://matheo.uliege.be

Mémoire Introduction d'un agent conversationnel basé sur l'IA pour l'apprentissage au raisonnement clinique dans la formation des infirmiers en pratique avancée, une étude pilote

Auteur : Vos, Julien
Promoteur(s) : GILBERT, Allison
Faculté : Faculté de Médecine
Diplôme : Master en sciences infirmières, à finalité spécialisée en pratiques avancées
Année académique : 2025-2026
URI/URL : http://hdl.handle.net/2268.2/25494

Avertissement à l'attention des usagers :

Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.

Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.

LIÈGE université
Médecine
Sciences de la santé publique

Introduction d'un agent conversationnel basé sur l'IA pour l'apprentissage au raisonnement clinique dans la formation des infirmiers en pratique avancée, une étude pilote

Mémoire présenté par Julien Vos

en vue de l'obtention du grade de

Master en Sciences Infirmières

Année académique 2025 -2026

Introduction d'un agent conversationnel basé sur l'IA pour l'apprentissage au raisonnement clinique dans la formation des infirmiers en pratique avancée, une étude pilote

Mémoire présenté par **Julien Vos**

en vue de l'obtention du grade de

Master en Sciences Infirmières

Année académique 2025 -2026

Promoteur : Dr Allison Gilbert

Co- promoteur : Pr Giovanni Briganti

Nombre de mots : 6715

Engagement de non plagiat

LIÈGE université Médecine Sciences de la santé publique

Engagement de non plagiat.

Je soussigné(e) NOM Prénom…Vos Julien……
Matricule étudiant : …S185513……

Déclare avoir pris connaissance de la charte anti-plagiat de l'Université de Liège et des dispositions du Règlement général des études et des évaluations. Je suis pleinement conscient(e) que la copie intégrale ou d'extraits de documents publiés sous quelque forme que ce soit (ouvrages, publications, rapports d'étudiant, internet, etc…) sans citation (i.e. mise en évidence de la citation par des guillemets) ni référence bibliographique précise est un plagiat qui constitue une violation des droits d'auteur relatifs aux documents originaux copiés indûment ainsi qu'une fraude. En conséquence, je m'engage à citer, selon les standards en vigueur dans ma discipline, toutes les sources que j'ai utilisées pour produire et écrire le mémoire que je dépose.

Fait le …20/02/2026

Signature

Annexe 7 du règlement des mémoires du MSSP

Déclaration d'utilisation de l'intelligence artificielle générative

Dans le cadre de ce mémoire, l'intelligence artificielle a été utilisée comme outil d'aide à la reformulation, à la clarification de certains passages, à la vérification de cohérence entre sections et à l'orientation ponctuelle dans la recherche d'informations. Elle a également été utilisée pour générer une figure de synthèse du déroulement de l'étude, à partir des éléments méthodologiques fournis par l'auteur.

Les contenus proposés par l'IA ont été relus, vérifiés, modifiés et validés par l'auteur. L'IA n'a pas été utilisée pour se substituer au travail personnel de recherche, d'analyse, d'interprétation ou de discussion des résultats. L'auteur demeure seul responsable du contenu final du mémoire.

Cet usage s'inscrit dans le cadre de la charte ULiège d'utilisation des intelligences artificielles génératives dans les travaux universitaires, qui rappelle la nécessité d'un usage responsable, critique et transparent de ces outils.

Remerciements

Je souhaite adresser mes sincères remerciements à Madame Allison Gilbert pour son encadrement tout au long de ce travail. La qualité de son accompagnement, la rigueur de ses remarques ainsi que la pertinence de ses orientations méthodologiques ont largement contribué à structurer cette recherche et à en renforcer la cohérence scientifique. Sa disponibilité et son exigence ont constitué des repères constants dans l'avancement de ce mémoire.

Je remercie également Amandine Halin pour son soutien et son implication au cours de ce projet. Ses échanges, ses retours et sa disponibilité ont été précieux à différentes étapes de ce travail et ont participé à en faciliter la réalisation.

Enfin, je tiens à exprimer ma profonde gratitude à ma mère pour son soutien constant tout au long de ce parcours. Sa présence, sa patience et ses encouragements ont constitué un appui essentiel dans la réalisation de ce mémoire. Je remercie également mes amis pour leur soutien et leur présence, qui ont contribué à maintenir un équilibre indispensable durant cette période de travail.

Table des matières

ENGAGEMENT DE NON PLAGIAT **2**

DÉCLARATION D'UTILISATION DE L'INTELLIGENCE ARTIFICIELLE GÉNÉRATIVE **3**

REMERCIEMENTS **4**

RÉSUMÉ **7**

ABSTRACT **8**

LISTE DES ABRÉVIATIONS **9**

PRÉAMBULE **10**

INTRODUCTION **11**

L'émergence d'un nouveau rôle infirmier 11

IPA et acquisition d'un nouveau domaine d'expertise 12

Compétences des IPA et défis pédagogiques 12

Impact potentiel de l'intelligence artificielle générative sur l'apprentissage du raisonnement clinique en formation 13

Hypothèses de recherche 15

Hypothèse principale *15*

Question de recherche *15*

Objectif principal *15*

Objectifs secondaires *15*

MÉTHODES **16**

Type d'étude et objectif général 16

Population et recrutement 16

Sélection du modèle pédagogique et organisation 16

Préparation de l'agent conversationnel pour le modèle pédagogique *17*

Intervention et déroulement 17

Contexte *17*

Organisation de la séance et déroulement de l'intervention *18*

Mesures et recueil des données 20

Analyses 21

Aspects éthiques et protection des données 23

RÉSULTATS : **24**

Résultats quantitatifs : 24

Caractéristiques des participants *24*

Complétude du dispositif *26*

Perceptions vis-à-vis de l'expérience *26*
Utilisabilité perçue de l'outil *28*
Performance au raisonnement clinique *29*
RÉSULTATS QUALITATIFS 30
1. L'IA comme guidage réflexif perçu *30*
2. Coût cognitif et temporalité *30*
3. Maintien de l'autonomie professionnelle *31*
4. Expérience pédagogique vécue *31*
5. Positionnement expérientiel face à l'outil *32*
DISCUSSION 33
1. UN SIGNAL FAVORABLE SUR LE RAISONNEMENT CLINIQUE 33
2. UN EFFET ASSOCIÉ À UN GUIDAGE RÉFLEXIF 33
3. MISE EN PERSPECTIVE AVEC LA LITTÉRATURE RÉCENTE 33
4. UN POTENTIEL RÉEL, MAIS UNE APPROPRIATION ENCORE INCOMPLÈTE 35
5. LIMITES DE L'ÉTUDE 36
6. PERSPECTIVES PÉDAGOGIQUES 36
CONCLUSION 37
RÉFÉRENCES : 38
ANNEXES 42

Résumé

Introduction : Le raisonnement clinique est une compétence centrale dans la formation des infirmiers de pratique avancée. Son apprentissage reste complexe, car il mobilise des processus cognitifs dynamiques, contextualisés et en partie implicites. Les agents conversationnels basés sur l'intelligence artificielle générative pourraient soutenir cet apprentissage s'ils sont intégrés dans un cadre pédagogique structuré.

Cette étude pilote visait à explorer l'apport potentiel d'un agent conversationnel basé sur l'IA dans l'apprentissage du raisonnement clinique d'étudiants en master en sciences infirmières, ainsi que son utilisabilité perçue et l'expérience vécue par les participants.

Méthodes : Une étude pilote combinant des approches quantitative et qualitative a été menée auprès de 21 étudiants inscrits en Master 1 en Sciences infirmières à l'Université de Liège. Les participants ont réalisé deux vignettes cliniques complexes selon un plan croisé, l'une sans agent conversationnel et l'autre avec l'agent AURIA, développé dans Microsoft Copilot selon une posture de tuteur socratique. Les productions écrites ont été évaluées à l'aide d'une grille dérivée du modèle de Tanner et de la *Lasater Clinical Judgment Rubric*. Les scores appariés ont été comparés par test de Wilcoxon. L'utilisabilité a été évaluée par le *System Usability Scale*. Un débriefing collectif a fait l'objet d'une analyse thématique inductive.

Résultats : Onze participants disposaient de données appariées complètes pour le critère principal. Les scores médians de raisonnement clinique étaient plus élevés avec IA dans les deux séries de correction : 24 [22–25] contre 29 [26–30] pour l'évaluateur 1 ($p = 0,0057$) et 23 [23–30] contre 39 [33,5–40] pour l'évaluateur 2 ($p = 0,0058$). Le score médian d'utilisabilité était de 62,5/100 [53,1–65,0]. L'analyse qualitative indiquait que l'agent était perçu comme un support de relance et de structuration du raisonnement, mais son usage restait associé à une charge cognitive importante, une contrainte temporelle et une appropriation variable.

Conclusion : Cette étude pilote met en évidence un signal favorable, mais exploratoire, concernant l'usage d'un agent conversationnel basé sur l'IA pour soutenir l'apprentissage du raisonnement clinique. L'intérêt pédagogique de l'outil semble dépendre de son intégration dans un cadre explicite, progressif et critique.

Mots-clés : intelligence artificielle générative ; agent conversationnel ; raisonnement clinique ; pratique infirmière avancée ; formation infirmière.

Abstract

Introduction: Clinical reasoning is a central competency in the training of advanced practice nurses. Its development remains complex, as it involves dynamic, contextualized and partly implicit cognitive processes. Conversational agents based on generative artificial intelligence could support this learning process if integrated within a structured pedagogical framework. This pilot study aimed to explore the potential contribution of an AI-based conversational agent to the development of clinical reasoning among master's students in nursing sciences, as well as its perceived usability and the participants' lived experience.

Methods: A pilot study combining quantitative and qualitative approaches was conducted with 21 students enrolled in the first year of the Master's programme in Nursing Sciences at the University of Liège. Participants completed two complex clinical vignettes according to a crossover design, one without a conversational agent and the other with AURIA, an agent developed in Microsoft Copilot and designed according to a Socratic tutor approach. The written productions were assessed using a grid derived from Tanner's model and the Lasater Clinical Judgment Rubric. Paired scores were compared using the Wilcoxon test. Usability was assessed using the System Usability Scale. A collective debriefing was analysed through inductive thematic analysis.

Results: Eleven participants had complete paired data for the primary outcome. Median clinical reasoning scores were higher with AI in both assessment series: 24 [22–25] versus 29 [26–30] for evaluator 1 ($p = 0.0057$), and 23 [23–30] versus 39 [33.5–40] for evaluator 2 ($p = 0.0058$). The median usability score was 62.5/100 [53.1–65.0]. The qualitative analysis indicated that the agent was perceived as a support for prompting and structuring reasoning, although its use remained associated with substantial cognitive load, time constraints and variable appropriation.

Conclusion: This pilot study highlights a favourable but exploratory signal regarding the use of an AI-based conversational agent to support the learning of clinical reasoning. The educational value of the tool appears to depend on its integration within an explicit, progressive and critical framework.

Keywords: generative artificial intelligence; conversational agent; clinical reasoning; advanced nursing practice; nursing education.

Liste des abréviations

IA : Intelligence artificielle

GenIA : Intelligence artificielle générative

IPA : Infirmier de pratique avancée

SUS : System Usability Scale

LCJR : Lasater Clinical Judgment Rubric

CII : Conseil international des infirmières

OMS : Organisation mondiale de la santé

KCE : Centre fédéral d'expertise des soins de santé

ICN : International Council of Nurses

LLM : Large Language Model

AB/BA : Séquences de randomisation du plan croisé

RGPD : Règlement général sur la protection des données

Auria : « *Assistant for Understanding and Reasoning in Nursing Analysis* », agent conversationnel pédagogique développé dans l'environnement Microsoft Copilot

Copilot : Environnement d'intelligence artificielle de Microsoft utilisé pour le déploiement de l'agent conversationnel

Préambule

Ce mémoire s'inscrit dans le champ des sciences infirmières et porte sur l'apprentissage du raisonnement clinique chez les étudiants en pratique avancée. Ce choix part d'un constat issu de mon parcours infirmier : dans les situations de soins complexes, le raisonnement clinique ne repose pas uniquement sur des connaissances théoriques, mais aussi sur la capacité à organiser les informations, formuler des hypothèses, prioriser les problèmes et justifier ses décisions.

L'intérêt de cette recherche est d'explorer la place possible d'un agent conversationnel basé sur l'intelligence artificielle générative comme support pédagogique. L'objectif n'est pas de considérer l'IA comme une réponse autonome, ni comme un substitut au jugement clinique, mais comme un outil pouvant aider à relancer, structurer et expliciter la réflexion.

Ce travail répond donc à une double motivation : mieux comprendre comment soutenir l'apprentissage d'une compétence centrale en pratique avancée, et interroger de manière critique l'intégration de nouveaux outils numériques dans la formation infirmière.

Introduction

L'émergence d'un nouveau rôle infirmier

Depuis la loi coordonnée du 10 mai 2015 relative à l'exercice des professions des soins de santé, modifiée à plusieurs reprises, la Belgique encadre juridiquement le rôle des infirmiers de pratique avancée (IPA), marquant une évolution importante du champ infirmier et répondant à plusieurs défis du système de santé belge [1,2].

Parmi les définitions de l'IPA, nous retenons celle du Conseil international des infirmières, rapportée par Roulin, Borrero et Iori : « Une infirmière qui exerce en pratique avancée est une infirmière diplômée qui a acquis des connaissances théoriques, le savoir-faire nécessaire aux prises de décisions complexes, de même que les compétences cliniques indispensables à la pratique avancée de sa profession. Les caractéristiques de cette pratique avancée sont déterminées par le contexte dans lequel l'infirmière sera autorisée à exercer. Une formation de base de niveau maîtrise est recommandée » [3]. Cette définition peut être complétée par celle de Hamric et al., qui décrivent la pratique avancée infirmière comme une application centrée sur le patient d'un éventail élargi de compétences visant à améliorer les indicateurs de santé des patients et des populations dans un domaine clinique spécialisé [4].

L'émergence du rôle IPA s'inscrit dans un contexte marqué par le vieillissement de la population, l'augmentation des maladies chroniques, la complexification des trajectoires de soins et les tensions sur les ressources médicales [2,5-8]. Dans ce cadre, l'intégration des IPA peut contribuer à améliorer l'accès aux soins, la continuité des prises en charge, la coordination interprofessionnelle et la qualité des soins pour des patients présentant des situations aiguës, complexes ou chroniques. À l'échelle internationale, les synthèses récentes montrent que les rôles infirmiers de pratique avancée sont associés à des effets favorables sur l'accès aux soins, certains résultats cliniques, la satisfaction des patients et, selon les contextes, l'efficience organisationnelle [9-14]. Cette évolution ne relève donc pas uniquement d'une logique de substitution ou de répartition des tâches, mais de la reconnaissance d'une expertise clinique infirmière avancée contribuant à la qualité, à la continuité et à la sécurité des soins [9-14].

IPA et acquisition d'un nouveau domaine d'expertise

La formation en pratique avancée représente une évolution qualitative du champ infirmier, marquée par l'acquisition de compétences cliniques, décisionnelles et organisationnelles approfondies. Dans leur domaine de spécialisation, les IPA exercent une autonomie clinique élargie leur permettant d'évaluer des situations complexes, de poser des hypothèses diagnostiques, de proposer des interventions adaptées et d'assurer le suivi des patients dans une perspective collaborative [15,16]. Cette pratique se déploie en complémentarité avec les autres professionnels de santé, notamment les médecins, au sein de conventions de collaboration interprofessionnelle qui encadrent les responsabilités partagées, les actes spécifiques et les modalités de coordination [16].

Le rôle de l'IPA ne se limite donc pas à une extension technique du champ infirmier. Il implique une capacité à intégrer des données cliniques multiples, à interpréter des situations instables, à prioriser des problèmes concurrents et à ajuster les décisions dans l'incertitude. Le profil belge de l'IPA insiste explicitement sur le raisonnement clinique, la réflexion critique, l'usage des données probantes, la coordination interprofessionnelle et la contribution à l'amélioration des pratiques [15,17]. La pratique avancée suppose ainsi un déplacement du cœur de compétence vers des activités cognitives de haut niveau, ce qui rend leur apprentissage particulièrement central en formation.

Compétences des IPA et défis pédagogiques

Pour s'intégrer dans ces nouvelles pratiques, les IPA doivent acquérir des compétences cliniques, décisionnelles et communicationnelles spécifiques, leur permettant d'assumer leurs fonctions dans des contextes de soins complexes et multidisciplinaires. Ces compétences incluent notamment la maîtrise du raisonnement clinique, la prise de décision autonome fondée sur les données probantes, la coordination interprofessionnelle et une capacité réflexive orientée vers l'amélioration continue des pratiques [17-20].

Le raisonnement clinique constitue une compétence centrale pour les IPA. Il conditionne la qualité du jugement, la pertinence des décisions et la sécurité des interventions. Il repose sur la capacité à analyser des situations instables, hiérarchiser les priorités, formuler des hypothèses et agir dans l'incertitude tout en maintenant une réflexion éthique et contextualisée [18,19].

Sur le plan pédagogique, plusieurs défis persistent dans la formation au raisonnement clinique [20]. L'hétérogénéité des parcours professionnels rend la progression variable d'un étudiant à l'autre. À cela s'ajoutent la difficulté d'enseigner des processus cognitifs complexes, souvent implicites et liés à l'expérience, le manque d'outils interactifs permettant de soutenir et d'évaluer la prise de décision, ainsi que la nécessité de dépasser une évaluation centrée uniquement sur les compétences techniques. Ces constats soulignent le besoin de méthodes pédagogiques innovantes, incluant la simulation clinique, la rétroaction formative et les outils numériques émergents [21-23].

Les données récentes suggèrent que certaines approches pédagogiques, comme la simulation, les environnements numériques interactifs, les applications d'apprentissage et les stratégies de rétroaction structurée, peuvent soutenir le raisonnement clinique, la prise de décision et la motivation des étudiants infirmiers [21-23]. Toutefois, ces effets dépendent de la qualité du scénario, du niveau d'accompagnement pédagogique et de l'articulation entre activité réflexive et transfert vers des situations cliniques plausibles. Le problème n'est donc pas seulement de disposer d'un outil innovant, mais de comprendre dans quelles conditions cet outil soutient réellement l'apprentissage du raisonnement.

Impact potentiel de l'intelligence artificielle générative sur l'apprentissage du raisonnement clinique en formation

Dans la formation infirmière, cette évolution pédagogique s'inscrit dans une numérisation croissante des dispositifs d'apprentissage. Les outils numériques ne servent plus seulement à transmettre des contenus ; ils peuvent aussi soutenir l'analyse de situations cliniques, favoriser l'engagement actif de l'étudiant et offrir une rétroaction plus rapide. Dans ce contexte, l'intérêt d'un agent conversationnel repose sur sa capacité à interagir avec l'apprenant, à reformuler, à relancer le questionnement et à simuler une forme de dialogue pédagogique. Cette caractéristique est particulièrement pertinente pour le raisonnement clinique, qui se construit moins par mémorisation de réponses que par confrontation structurée à l'incertitude, aux hypothèses et aux justifications [24,25].

Les grands modèles de langage, dont ChatGPT constitue l'exemple le plus diffusé, suscitent un intérêt croissant dans les formations en santé en raison de leur capacité à générer des explications, proposer des cas, reformuler des contenus et soutenir des interactions individualisées. Dans les revues récentes portant sur la formation infirmière, ChatGPT apparaît

comme un outil potentiellement utile pour la préparation de matériel pédagogique, l'analyse de situations cliniques, le soutien à l'apprentissage personnalisé et certaines formes de reformulation pédagogique. Néanmoins, cette littérature reste prudente, car elle demeure largement dominée par des éditoriaux, des commentaires et des études descriptives ou exploratoires, limitant la solidité des conclusions sur son efficacité pédagogique réelle [26,27]. L'intérêt pédagogique de l'intelligence artificielle générative ne peut pas être dissocié de ses limites. Ces modèles peuvent produire des réponses inexactes mais plausibles, reproduire des biais, poser des problèmes de confidentialité et favoriser une dépendance excessive à l'outil si son usage n'est pas encadré pédagogiquement [26,27]. Plus largement, dans les formations en santé, les synthèses récentes montrent que la qualité des preuves concernant les effets de l'IA sur les résultats éducatifs reste faible, hétérogène et méthodologiquement limitée [28,29].

Dans ce cadre, ce mémoire explore la faisabilité d'une méthode pédagogique mobilisant un agent conversationnel basé sur l'IA générative pour soutenir l'apprentissage du raisonnement clinique chez des étudiants en sciences infirmières engagés dans une formation de pratique avancée. Il vise également à documenter les conditions d'usage de l'outil, son acceptabilité, sa capacité à soutenir une réflexion clinique plus élaborée et les tensions qu'il peut introduire dans l'apprentissage. Ce positionnement est cohérent avec l'état actuel de la littérature, qui suggère un potentiel pédagogique réel tout en soulignant la nécessité d'évaluer plus rigoureusement ses effets sur les apprentissages et l'activité cognitive des étudiants [25,28,29].

Hypothèses de recherche

Hypothèse principale

L'hypothèse principale est que l'utilisation d'un agent conversationnel basé sur l'IA comme outil d'accompagnement pédagogique améliore la qualité du raisonnement clinique chez les étudiants infirmiers en pratique avancée, par rapport à une formation sans recours à cet outil.

Question de recherche

L'utilisation d'un agent conversationnel basé sur l'IA peut-elle soutenir l'apprentissage du raisonnement clinique dans la formation des infirmiers en pratique avancée ?

Objectif principal

Évaluer l'effet de l'utilisation d'un agent conversationnel basé sur l'IA sur la performance au raisonnement clinique d'étudiants en master en sciences infirmières.

Objectifs secondaires

Explorer l'acceptabilité et la facilité d'utilisation de l'agent conversationnel par les étudiants.

Identifier les limites et les opportunités perçues de l'intégration d'un agent conversationnel basé sur l'IA dans la formation clinique.

Méthodes

Type d'étude et objectif général

Cette étude pilote a adopté un plan expérimental croisé (cross-over) comparant deux modalités d'apprentissage du raisonnement clinique : une condition sans recours à un agent conversationnel et une condition avec recours à un agent conversationnel.

Population et recrutement

La population étudiée a été constituée des étudiants inscrits en Master 1 en Sciences infirmières, finalité pratique avancée, à l'Université de Liège. Les critères d'inclusion utilisés ont été l'inscription au cursus et la signature d'un consentement éclairé. Les étudiants ne relevant pas de cette finalité, ne consentant pas à l'étude ou présentant une expertise avancée en intelligence artificielle ont été exclus. L'échantillon visé a été préalablement calculé à un objectif de 20 participants, correspondant à la cohorte disponible. Le recrutement a reposé sur le volontariat après une séance d'information.

Sélection du modèle pédagogique et organisation

Le tuteur socratique

Préalablement à l'expérimentation, les étudiants ont reçu des consignes communes quant à l'usage attendu de l'agent conversationnel. Il leur a été demandé de l'utiliser comme un appui à leur réflexion clinique, en interaction avec leur propre raisonnement, et non comme un substitut à celui-ci. L'objectif n'était pas d'obtenir une réponse immédiatement exploitable, mais de soutenir l'analyse du cas, la hiérarchisation des données, l'exploration d'hypothèses et la justification des décisions envisagées. Une illustration de l'interface et du cadre d'utilisation de l'agent est présentée en annexe 5.

Préparation de l'agent conversationnel pour le modèle pédagogique

L'agent conversationnel utilisé dans le cadre de l'étude était AURIA (Assistant for Understanding and Reasoning In Nursing Analysis), intégré à l'environnement Microsoft Copilot et partagé avec les étudiants dans le cadre du cours. Il ne s'agissait pas d'un usage libre de Copilot, mais d'un agent dédié, créé en amont pour les besoins de l'exercice pédagogique et selon le choix de la méthodologie pédagogique initialement sélectionnée. Sa préparation a été réalisée par le promoteur (AG) et co-promoteur (GB) de ce projet. Dans cette perspective, un prompt générique a été élaboré afin d'orienter l'agent vers une posture de relance réflexive, de questionnement, de structuration du raisonnement et d'exploration d'hypothèses, plutôt que vers la simple fourniture de réponses directes ou de solutions toutes faites (annexe 5).

Des tests préalables ont été réalisés avant la séance, avec la participation d'une enseignante du master en sciences infirmières (AH) et de l'étudiant chercheur (JV), afin d'évaluer la pertinence des réponses produites par l'agent et d'ajuster le cadrage de son utilisation.

Intervention et déroulement

Le processus complet de l'intervention est schématisé à la Figure 1.

Contexte

L'expérimentation s'est déroulée le 11 décembre 2025, de 9h00 à 12h00, dans le cadre d'un cours de démarche clinique du Master 1 en Sciences infirmières, finalité pratique avancée, à l'Université de Liège. La séance a été organisée en distanciel synchrone, les étudiants participant au cours via l'interface Microsoft Teams.

La répartition des participants a été réalisée par l'équipe de recherche avant le début de l'expérimentation, afin d'équilibrer les groupes selon l'expérience clinique déclarée. Les participants ont été stratifiés selon deux catégories d'expérience (< 5 ans / ≥ 5 ans), puis répartis de façon aléatoire vers le groupe 1 ou 2.

Organisation de la séance et déroulement de l'intervention

L'étude a été réalisée selon un déroulement structuré comprenant un briefing initial, deux phases successives de résolution de vignettes cliniques selon le plan croisé de l'étude, la complétion de questionnaires après chaque phase, puis un échange collectif final enregistré. Les étudiants recevaient des consignes communes concernant l'utilisation de l'agent conversationnel, présenté comme un support à la réflexion clinique et non comme un substitut au raisonnement personnel.
Deux vignettes cliniques complexes de difficulté comparable ont été utilisées : l'une centrée sur une situation aiguë, l'autre sur une situation chronique nécessitant coordination et priorisation clinique. Leur description détaillée figure en annexe 6 et 7.

Lors de la première phase, un groupe réalisait une vignette clinique (vignette A) sans recours à l'IA tandis que l'autre réalisait l'autre vignette (vignette B) dans la même condition. Lors de la seconde phase, les participants changeaient de vignette et travaillaient cette fois avec recours à l'IA. Pour chaque cas, une feuille d'exercice structurée était remise aux participants. Celle-ci comprenait cinq objectifs de travail écrits : résumer le cas complexe identifié, repérer et justifier les éléments nécessitant une attention particulière, identifier les problèmes de santé, établir un plan d'action sur base des risques et attentes pluridisciplinaires, puis réaliser une synthèse destinée à l'enseignant. Les feuilles d'exercice des cas A et B sont présentées en annexes (Annexes 6 et 7).

À l'issue de la séance, un échange collectif post-exercice a été enregistré afin de recueillir le ressenti à chaud des participants concernant l'outil, son apport dans la réflexion clinique et le vécu du dispositif pédagogique. Des questions ouvertes guidaient ce recueil.

Le processus est schématisé dans la figure suivante.

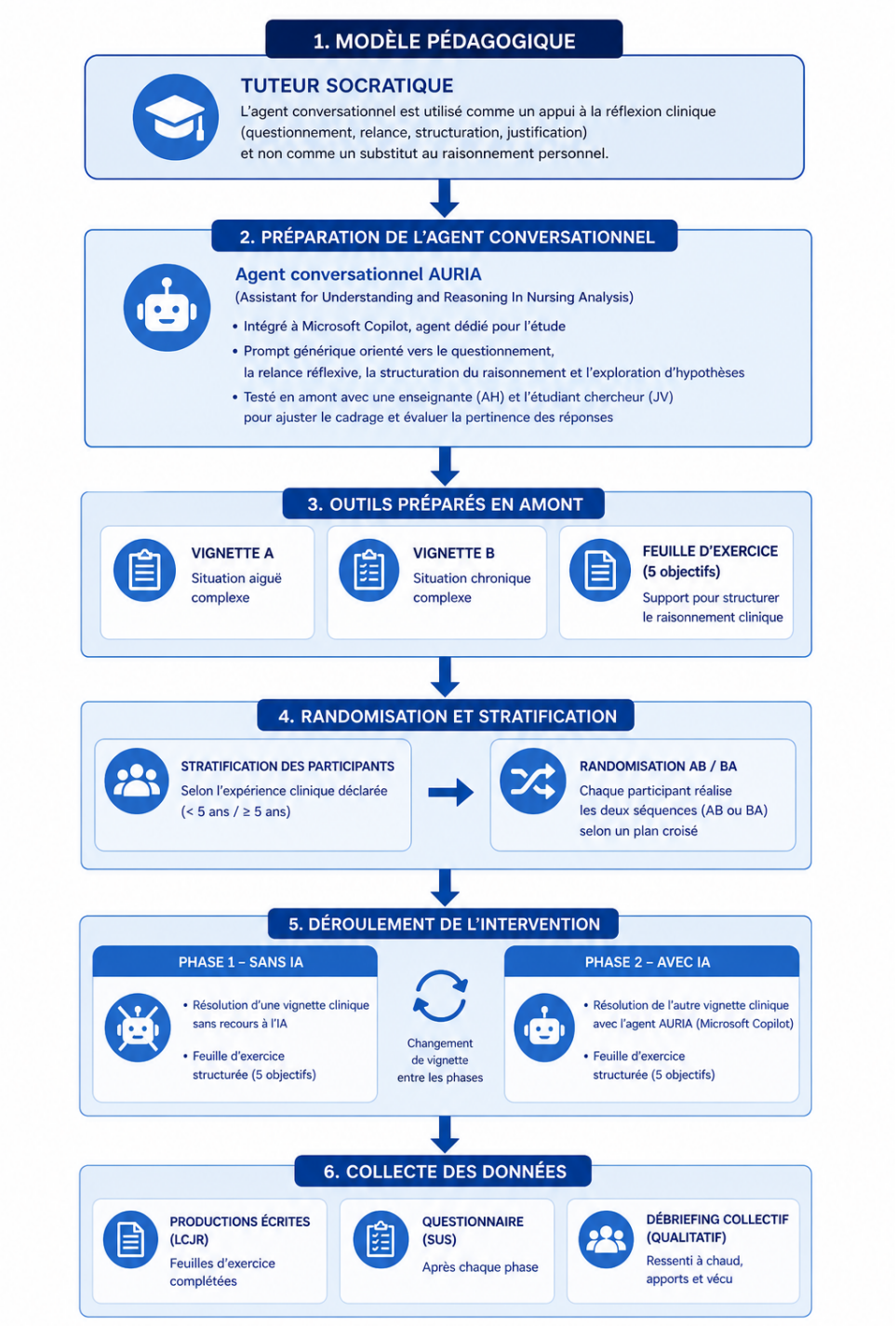


Figure 1. Processus de l'étude explicitant les différentes étapes de l'intervention pédagogique, depuis le choix du modèle pédagogique jusqu'au recueil des données. Figure générée avec ChatGPT à partir des informations méthodologiques fournies par l'auteur, version du 15 mai 2026.

Mesures et recueil des données

L'approche quantitative

Les données récoltées pour cette approche comprenaient :

- les caractéristiques sociodémographiques des participants ;
- la satisfaction globale cotée sur 10 ;
- la caractérisation de la satisfaction sur base de différents items ;
- les réponses à l'échelle « System Usability Scale » (SUS) ;
- les scores de raisonnement clinique obtenus à partir des productions écrites.

En ce qui concernent les données sociodémographiques, la satisfaction globale et ses caractéristiques, celles-ci ont été récoltées par le biais d'un questionnaire en ligne (Microsoft Forms) transmis aux participants en fin de session. L'utilisabilité perçue a été évaluée à l'aide de la version française du « System Usability Scale » (SUS) dont le score global a été calculé conformément à la procédure de cotation décrite pour la version française validée par Gronier et Baudet (2021). L'échelle SUS et ses modalités détaillées de calcul sont présentées en annexe 2. Cette échelle était également incluse dans le questionnaire en ligne.
Le questionnaire utilisé est présenté en annexe 3.

En ce qui concerne les scores de raisonnement clinique, ceux-ci ont été étudiés à l'aide d'une grille simplifiée dérivée du modèle de Tanner et de la « Lasater Clinical Judgment Rubric » (annexe 1), couvrant les dimensions remarquer, interpréter, répondre et réfléchir. Un des indicateurs de la dimension « répondre » (calme, maîtrise émotionnelle, posture professionnelle) n'a pas été évalué en raison du format distanciel ne permettant pas l'observation directe des comportements. La grille utilisée comprenait ainsi 10 indicateurs côtés de 1 à 4, pour un score total sur 40. Deux évaluateurs ont corrigé les copies à l'aveugle.

L'approche qualitative

Le matériau qualitatif reposait sur l'enregistrement vidéo du débriefing collectif post-exercice, réalisé à l'issue de la séance. Cet échange collectif semi-structuré était guidé par un canevas préétabli portant sur trois dimensions principales : la perception de l'agent conversationnel, son apport dans la réflexion clinique et le ressenti global vis-à-vis du dispositif pédagogique. Le guide utilisé est présenté en annexe 4.

La transcription initiale du débriefing a été générée automatiquement via Microsoft Teams, puis relue, corrigée et anonymisée avant analyse. Les éléments directement identifiants ont été remplacés par des identifiants anonymes. L'annexe 8 présente les extraits mobilisés, les codes initiaux, les regroupements successifs et la construction des thèmes.

L'analyse qualitative a été conduite selon une approche thématique inductive inspirée de Braun et Clarke. Le processus d'analyse a été mené de manière itérative, à partir d'une lecture répétée du matériau, d'un codage initial, puis d'un regroupement progressif des codes en catégories et en thèmes analytiques. Les thèmes ont ensuite été révisés, définis et nommés afin de produire une structuration finale des résultats qualitatifs. Le processus analytique détaillé, incluant la piste d'audit du codage, est présenté en annexe 8.

Analyses

Analyses statistiques

Les données ont été encodées et structurées dans une base de données Excel, puis analysées à l'aide du logiciel R via l'interface R Commander (Rcmdr).

Les variables catégorielles ont été décrites par des effectifs (n) et des pourcentages (%). Les items de satisfaction évalués par échelle de Likert ont également été décrits par effectifs et pourcentages. Les variables quantitatives ont été explorées à l'aide d'indicateurs de tendance centrale et de dispersion. La normalité des distributions a été explorée à l'aide du test de Shapiro-Wilk. Compte tenu du faible effectif de l'étude, les résultats ont été principalement présentés sous forme de médiane et d'intervalle interquartile.

La comparaison des scores de raisonnement clinique entre les conditions sans IA et avec IA a été réalisée chez les participants disposant de données appariées complètes pour les deux conditions. La normalité des différences appariées n'était pas rejetée par le test de Shapiro-Wilk. Toutefois, compte tenu du faible effectif, du caractère pilote de l'étude et de la nature

des scores issus d'une grille de cotation, un test de Wilcoxon pour données appariées a été retenu. Les données manquantes n'ont pas été imputées.
Le score du System Usability Scale a été calculé selon la procédure standardisée : les items impairs ont été recodés en soustrayant 1 au score brut, les items pairs ont été recodés en calculant 5 moins le score brut, puis la somme des scores recodés a été multipliée par 2,5 afin d'obtenir un score total sur 100. Le score SUS n'a été calculé que pour les participants ayant répondu aux dix items.

Aucune analyse spécifique des effets de séquence, de période ou d'accord inter-évaluateurs n'a été réalisée.

Analyse qualitative

Afin d'analyser les données qualitatives, une approche thématique inductive inspirée de Braun et Clarke a été retenue.

La transcription initiale du débrief a été générée automatiquement via Microsoft Teams, puis relue, corrigée et pseudonymisée avant analyse. Les éléments identifiants ont été remplacés par des identifiants anonymes. Le matériau a ensuite fait l'objet d'une lecture répétée, d'un codage initial, puis d'un regroupement progressif en thèmes analytiques. Les étapes de révision, de définition et de rédaction des thèmes ont permis d'aboutir à une structuration finale des résultats qualitatifs. Le détail du processus analytique, incluant les tableaux de codage et les regroupements successifs, est présenté en annexe 8.

Aspects éthiques et protection des données

L'étude a fait l'objet d'un avis auprès du Comité d'Éthique Hospitalo-Facultaire Universitaire de Liège. Dans son avis référencé 2025/624, le Comité a constaté que l'étude n'entrait pas dans le cadre de la loi du 7 mai 2004 relative aux expérimentations sur la personne humaine et n'a pas émis d'objection éthique à sa réalisation.

La participation reposait sur un consentement éclairé. Les données recueillies ont été anonymisées avant analyse. Aucune donnée clinique identifiable n'a été introduite dans l'agent conversationnel. Les cas complexes ont été élaborés sur la base des connaissances des investigateurs et pouvaient correspondre à des situations vécues, sans reprendre aucune donnée identifiable concernant les patients. Le recueil des informations personnelles des participants et des copies des exercices a été réalisé via la plateforme sécurisée Microsoft Teams de l'ULiège. Les enregistrements et transcriptions utilisés pour l'analyse qualitative ont été stockés sur les espaces sécurisés institutionnels, et utilisés uniquement dans le cadre de l'étude par les chercheurs identifiés dans le projet. Les matériaux, grilles et instruments complets sont présentés en annexe.

Résultats :

Résultats quantitatifs :

Caractéristiques des participants

La cohorte recrutée comprend 21 participants dont 16 femmes (76,2%) et 5 hommes (23,8%). Parmi les participants, 3 ont moins de 25 ans (14,3%), 8 sont entre 25 et 34 ans (38,1%), 7 entre 35 et 44 ans (33,3%) et 3 ont entre 45 et 54 ans (14,3%).

Les différents résultats sont repris à la Table 1.

Tableau 1. Caractéristiques sociodémographiques des participants inclus dans l'étude

Caractéristiques	**N**	**%**
Genre		
Homme	5	23,8
Femme	16	76,2
Age		
<25 ans	3	14,3
25-34 ans	8	38,1
35-44 ans	7	33,3
45-54 ans	3	14,3
Statut professionnel		
Etudiant	1	4,8
Employé	18	85,7
Indépendant	2	9,5
Expérience professionnelle		
< 2ans	3	14,3
2-5 ans	4	19
6-10 ans	6	28,6
> 10 ans	8	38,1
Formation de base		
Bachelier	8	38,1
Bachelier de spécialisation	13	61,9

Domaine clinique		
Soins à domicile	3	14,3
Urgences	5	23,8
Soins Intensifs	4	19,0
Psychiatrie	1	4,8
Hospitalier	6	28,6
Sang	1	4,8
Polycliniques	1	4,8

Les participants ont été répartis en deux groupes : 10 participants dans le groupe 1 (47,6 %) et 11 dans le groupe 2 (52,4 %). La comparaison des caractéristiques des deux groupes est disponible à la Table 2.

Tableau 2. Comparaison des caractéristiques sociodémographiques et professionnelles selon le groupe d'affectation

Caractéristiques	**Groupe 1 N(%)**	**Groupe 2 N(%)**
Genre		
Homme	3 (30,0)	2 (18,2)
Femme	7 (70,0)	9 (81,8)
Age		
<25 ans	2 (20,0)	1 (9,1)
25-34 ans	4 (40,00)	4 (36,4)
35-44 ans	3 (30,0)	4 (36,4)
45-54 ans	1 (10,0)	2 (18,2)
Données manquantes	0 (0)	0 (0)
Statut professionnel		
Étudiant	1 (10,0)	0 (0)
Employé	8 (80,0)	10 (90,9)
Indépendant	1 (10,0)	1 (9,1)
Expérience clinique		

< 2ans	2 (20,0)	1 (9,1)
2-5 ans	2 (20,0)	2 (18,2)
6-10 ans	4 (40,0)	2 (18,2)
> 10 ans	2 (20,0)	6 (54,5)
Formation de base		
Bachelier	4 (40,0)	4 (36,4)
Bachelier de spécialisation	6 (60,0)	7 (63,6)
Domaine clinique		
Soins à domicile	1 (10,0)	2 (18,2)
Urgences	4 (40,0)	1 (9,1)
Soins Intensifs	0 (0)	4 (36,4)
Psychiatrie	0 (0)	1 (9,1)
Hospitalier	3 (30,0)	3 (27,3)
Sang	1 (10)	0 (0)
Polycliniques	1 (10)	0 (0)

Complétude du dispositif

Parmi les 21 participants, 11 (52,4 %) ont rendu les deux cas attendus, 6 (28,6 %) ont fourni une réponse partielle et 4 (19,0 %) n'ont pas complété l'exercice. Parmi les participants ayant rendu les deux cas (n=11), 5 (45,5%) faisaient partie du groupe 2 et 6 (54,5%) du groupe 1.

Perceptions vis-à-vis de l'expérience

La satisfaction globale vis-à-vis de la méthode pédagogique a été évaluée sur une échelle de 1 à 10. Celle-ci a été renseignée par 16 participants sur 21 (76,2 %). La médiane observée était de 8 [7–9].

Satisfaction par rapport à l'expérience

La satisfaction par rapport à l'expérience a été évaluée par 16 participants sur une échelle de Likert à 5 niveaux, allant de 1 « pas du tout d'accord » à 5 « tout à fait d'accord ». Parmi les répondants, 1 participant (6,3 %) a répondu 3 « indifférent », 10 participants (62,5 %) ont répondu 4 « d'accord » et 5 participants (31,3 %) ont répondu 5 « tout à fait d'accord ».

Apport du dispositif dans la réflexion clinique

L'apport du dispositif dans la réflexion clinique a été évalué par 15 participants sur une échelle de Likert à 5 niveaux, allant de 1 « pas du tout d'accord » à 5 « tout à fait d'accord ». Parmi les répondants, 3 participants (20,0 %) ont répondu 2 « pas d'accord », 1 participant (6,7 %) a répondu 3 « indifférent », 8 participants (53,3 %) ont répondu 4 « d'accord » et 3 participants (20,0 %) ont répondu 5 « tout à fait d'accord ».

Correspondance de l'activité aux attentes

La correspondance de l'activité aux attentes des participants a été évaluée par 15 participants sur une échelle de Likert à 5 niveaux, allant de 1 « pas du tout d'accord » à 5 « tout à fait d'accord ». Parmi les répondants, 6 participants (40,0 %) ont répondu 3 « indifférent », 6 participants (40,0 %) ont répondu 4 « d'accord » et 3 participants (20,0 %) ont répondu 5 « tout à fait d'accord ».

Participation future à une expérience similaire

L'envie de participer à nouveau à une telle expérience a été évaluée par 15 participants sur une échelle de Likert à 5 niveaux, allant de 1 « pas du tout d'accord » à 5 « tout à fait d'accord ». Parmi les répondants, 9 participants (60,0 %) ont répondu 4 « d'accord » et 6 participants (40,0 %) ont répondu 5 « tout à fait d'accord ».

Les résultats sont repris à la Table 3.

Tableau 3. Distribution des réponses aux items de satisfaction évalués par échelle de Likert

Question	1 = Pas du tout d'accord n(%)	2 : Pas d'accord n(%)	3 = Indifférent n(%)	4 = D'accord n(%)	5 = Tout à fait d'accord n(%)
Je suis satisfait de l'expérience	0 (0)	0 (0)	1 (6,3)	10 (62,5)	5 (31,3)
Le dispositif m'a permis de réfléchir à	0 (0)	3 (20,0)	1 (6,7)	8 (53,3)	3 (20,0)

ma pratique ou à mon raisonnement clinique.					
L'activité correspondait à mes attentes.	0 (0)	0 (0)	6 (40,0)	6 (40,0)	3 (20,0)
Je participerais volontiers à une expérience similaire.	0 (0)	0 (0)	0 (0)	9 (60,0)	6 (40,0)

Utilisabilité perçue de l'outil

Le score global de l'échelle « System Usability Scale » a été calculé pour 14 participants sur 21 (66,7 %) disposant de réponses complètes aux 10 items. Le score médian observé était de 62,5 [53,1–65,0].

Les résultats sont repris à la Table 4.

Tableau 4. Répartition des scores d'utilisabilité perçue selon les catégories d'interprétation du System Usability Scale

Catégories de niveaux de scores	n(%)
Pire à mauvaise (25 à 35,6)	2 (14,3)
Mauvaise à acceptable (35,7 à 50,9)	1 (7,1)
Acceptable à bonne (50,9 à 71,4)	10 (71,4)
Bonne à excellent (71,4 à 85,5)	1 (7,1)

La distribution des scores SUS est présentée à la Figure 2.

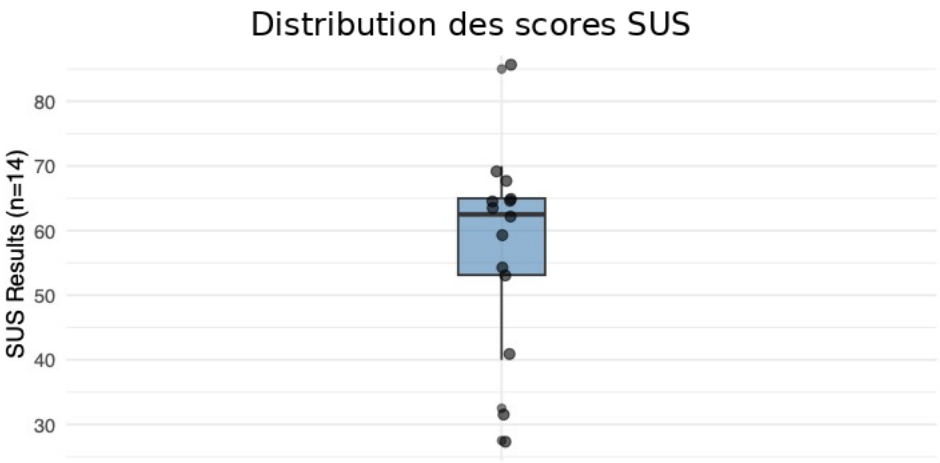


Figure 2. Distribution des scores d'utilisabilité perçue mesurés par le System Usability Scale (SUS) chez les participants disposant de réponses complètes aux dix items (n = 14).

Performance au raisonnement clinique

Pour l'évaluateur 1, le score médian sans IA était de 24 [22–25], contre 29 [26–30] avec IA. La médiane des différences individuelles était de +6 points. Le test de Wilcoxon mettait en évidence une différence significative en faveur de la condition avec IA (V = 64,5 ; p = 0,0057). Pour l'évaluateur 2, le score médian sans IA était de 23 [23–30], contre 39 [33,5–40] avec IA. La médiane des différences individuelles était de +9 points. Le test de Wilcoxon mettait également en évidence une différence significative en faveur de la condition avec IA (V = 55 ; p = 0,0058).

Les deux séries de correction allaient donc dans le même sens, avec des scores médians plus élevés en condition avec IA qu'en condition sans IA. Les résultats sont présentés séparément pour chaque évaluateur.

Résultats qualitatifs

Le recueil qualitatif post-séance portait sur la perception de l'outil d'IA, son apport dans la réflexion clinique et le ressenti global vis-à-vis du dispositif pédagogique. L'analyse thématique inductive a permis d'identifier cinq thèmes : l'IA comme guidage réflexif perçu, le coût cognitif et la temporalité, le maintien de l'autonomie professionnelle, l'expérience pédagogique vécue et le positionnement expérientiel face à l'outil. Le détail du codage, des verbatims et des regroupements thématiques est présenté en annexe 8.

1. L'IA comme guidage réflexif perçu

Les participants ne décrivaient pas l'agent conversationnel comme un fournisseur direct de réponses, mais comme un support au processus réflexif. L'outil était perçu comme aidant à structurer le raisonnement, à relancer l'analyse et à explorer certaines dimensions du problème qui n'auraient pas toujours été mobilisées spontanément.

« c'est un super bon outil parce qu'il pousse vraiment à la réflexion, à aller plus loin, à vraiment ne rien négliger, rien oublier » (ID2)

Cet apport apparaissait toutefois conditionnel. Plusieurs participants indiquaient que l'agent confirmait davantage un raisonnement déjà présent qu'il ne modifiait réellement l'hypothèse initiale.

« l'IA m'a juste confirmé que mon idée de base était bonne » (ID2)

D'autres participants soulignaient les limites de cet apport, en décrivant un outil parfois centré sur la reformulation ou le recentrage, avec peu de propositions réellement nouvelles. L'utilité perçue semblait également dépendre du niveau de connaissances préalables et du type de situation clinique. L'agent était jugé plus pertinent dans des situations complexes, incertaines ou inhabituelles, et moins utile dans des situations routinières ou déjà maîtrisées.

« il faut déjà nous-même savoir plus ou moins de quoi on parle » (ID3)

2. Coût cognitif et temporalité

L'utilisation de l'agent conversationnel était associée à une charge cognitive importante. Les participants évoquaient un effort mental accru, lié à la réflexion supplémentaire, à la vérification des réponses et au temps nécessaire pour exploiter l'outil. L'IA n'était donc pas décrite comme un facteur de gain de temps, mais comme un moyen d'approfondir le traitement de l'information.

« Ça pousse à la réflexion (…) mais effectivement, il faut avoir du temps et on n'a pas toujours le temps » (ID13)

Cette charge cognitive entrait directement en tension avec la contrainte temporelle de l'exercice. Plusieurs participants rapportaient qu'en situation de temps court, l'usage de l'outil devenait frustrant ou difficilement soutenable.

« j'ai fini par tout bâcler parce que c'était trop long pour le quart d'heure » (ID8)

Les verbatims faisaient aussi apparaître une réflexion située selon le contexte clinique. L'outil semblait davantage compatible avec des situations chroniques ou complexes qu'avec des prises en charge aiguës, où la rapidité décisionnelle reste prioritaire.

« pour des trucs chroniques qui suivent, c'est super intéressant » (ID2)

3. Maintien de l'autonomie professionnelle

Les discours recueillis faisaient apparaître une régulation active de la place de l'outil dans le raisonnement. Les participants insistaient sur la nécessité de conserver la décision finale et de ne pas laisser l'agent prendre la place du jugement professionnel. L'outil semblait acceptable à condition de rester sous contrôle.

« J'ai compris que c'est me poussait beaucoup plus dans la réflexion (…) donc d'avoir toujours le volant quoi » (ID17)

À l'inverse, certains verbatims décrivaient une expérience dans laquelle le participant s'était laissé guider par l'outil de manière trop importante, ce qui conduisait ensuite à reformuler un usage plus complémentaire.

« Je me suis laissé manger par l'IA » (ID12)

Plusieurs participants associaient également l'usage pertinent de l'outil à l'existence de compétences préalables ou à la nécessité d'une formation spécifique. Les données ne traduisaient donc pas un rejet global de l'IA, mais une volonté de contrôler sa place dans le processus décisionnel et de préserver l'autonomie professionnelle.

4. Expérience pédagogique vécue

Le dispositif était fréquemment décrit comme une expérience pédagogique spécifique, et non comme un simple usage technique d'un outil. Plusieurs participants établissaient une analogie avec une situation d'encadrement ou de supervision, dans la mesure où l'agent relançait la réflexion au lieu de fournir directement une solution.

« j'avais l'impression d'être en stage quand la MFP nous pose des questions. Et pourquoi et pourquoi ? Pourquoi ? » (ID13)

Cette dimension pédagogique était associée à un raisonnement plus approfondi, mais aussi à certaines limites liées au format proposé, notamment au temps imparti et à l'intensité cognitive de l'exercice. Le travail en groupe était également perçu comme un soutien à l'apprentissage, notamment par confrontation des points de vue et comparaison des démarches.

« ça me permet de savoir ce que les autres auraient fait (…) et de peaufiner mes actions » (ID13)

L'expérience apparaissait ainsi comme une modalité pédagogique particulière, articulant relance réflexive, contraintes de format et apprentissage collectif.

5. Positionnement expérientiel face à l'outil

Le vécu global associait intérêt, curiosité, prudence et parfois frustration. Les participants décrivaient une expérience nouvelle, souvent jugée intéressante, mais aussi marquée par une conscience des limites de l'outil et des conditions nécessaires à son usage.

« perso je n'avais jamais utilisé l'IA donc du coup j'ai trouvé ça assez intéressant comme système » (ID21)

D'autres verbatims faisaient apparaître une appropriation plus ambivalente, marquée par une déstabilisation initiale ou des limites techniques.

« je m'attendais pas du tout à ça donc ça m'a un petit peu déstabilisée. Mais c'était mieux, du coup, c'était mieux comme outil. » (ID4)

Certains participants exprimaient enfin une ouverture prudente à un réusage futur de l'outil, sous réserve de conditions plus favorables. L'expérience ne semblait donc ni conduire à une adhésion naïve ni à un rejet de l'IA, mais plutôt à un positionnement critique, dépendant du temps disponible, du niveau de familiarité avec l'outil et de la capacité à maintenir une posture professionnelle autonome.

Discussion

1. Un signal favorable sur le raisonnement clinique

Cette étude pilote suggère qu'un agent conversationnel basé sur l'IA, intégré dans un dispositif pédagogique structuré, pourrait soutenir l'apprentissage du raisonnement clinique chez des étudiants en pratique avancée. Les deux séries de correction montrent des scores médians plus élevés en condition avec IA qu'en condition sans IA, avec une différence statistiquement significative malgré le faible effectif analysable.

Ce résultat constitue un signal quantitatif favorable et cohérent avec l'hypothèse principale. Il doit toutefois rester interprété avec prudence, car seuls 11 participants disposaient de données appariées complètes pour le critère principal. L'effectif réduit limite donc la portée statistique du résultat et la généralisation des tendances observées.

2. Un effet associé à un guidage réflexif

Les participants ne décrivaient pas l'agent conversationnel comme une source autonome de réponses, mais comme un support de relance et de structuration du raisonnement. Les verbatims suggèrent que l'outil incite à approfondir l'analyse, à ne pas négliger certains éléments et à mieux organiser la réflexion clinique. Cette lecture est cohérente avec l'amélioration des performances observées dans les productions écrites.

L'effet de l'outil semble ainsi moins relever d'une substitution du jugement clinique que d'un guidage réflexif facilitant l'explicitation du raisonnement. Cet effet dépend toutefois de conditions contextuelles, notamment le temps disponible, la complexité des situations, le niveau de connaissances des étudiants et le maintien d'une responsabilité professionnelle active.

3. Mise en perspective avec la littérature récente

Les résultats de cette étude suggèrent que l'agent conversationnel a principalement été perçu comme un support de relance et de structuration du raisonnement, plutôt que comme une source autonome de réponses. Cette observation rejoint les travaux récents sur l'intégration de ChatGPT et des agents conversationnels en formation infirmière, qui décrivent ces outils comme des soutiens possibles à l'apprentissage individualisé, à la reformulation, à la génération ou à l'analyse de situations cliniques et à l'organisation des connaissances [24-27,30-32]. Dans cette perspective, l'intérêt pédagogique de l'agent ne semble pas résider

uniquement dans la qualité des réponses produites, mais dans sa capacité à susciter un dialogue réflexif autour de la situation clinique.

Les données qualitatives montrent toutefois que cet apport reste conditionnel. Plusieurs participants ont rapporté une charge cognitive importante, une contrainte temporelle marquée et une difficulté à exploiter l'outil sans familiarité préalable avec l'IA. Ces éléments rejoignent les limites décrites dans la littérature récente, notamment le risque d'informations inexactes, la nécessité de vérifier les réponses, les enjeux d'intégrité académique, la dépendance potentielle à l'outil et le besoin d'un encadrement pédagogique explicite [26,27,31,32]. La revue narrative d'Abujaber et al. présente ainsi ChatGPT comme une opportunité pour la formation infirmière, mais insiste sur la nécessité d'un usage supervisé, critique et intégré dans un cadre pédagogique clair [31]. De même, la revue PAGER d'Amankwaa et al. souligne que les usages de ChatGPT en formation infirmière restent encore hétérogènes et insuffisamment stabilisés sur le plan méthodologique [32].

Les résultats qualitatifs de cette étude convergent également avec les travaux décrivant l'importance de la familiarité préalable et de la littératie IA. Les participants qui avaient déjà expérimenté l'IA semblaient mieux identifier les conditions d'usage de l'agent, alors que ceux qui découvraient l'outil rapportaient davantage une appropriation difficile. Cette observation rejoint les études qualitatives menées auprès d'étudiants ou d'enseignants en soins infirmiers, qui montrent que l'intérêt perçu de ChatGPT dépend fortement du niveau de confiance, des compétences préalables, de la capacité à formuler les demandes et de la possibilité de vérifier les réponses produites [33,34]. L'usage pédagogique de l'IA ne peut donc pas être dissocié d'une formation progressive à son utilisation critique.

Enfin, la revue de portée de Zintchem et al. offre un cadre particulièrement proche de la présente étude, en discutant l'IA générative dans l'apprentissage du raisonnement clinique infirmier. Les auteurs soulignent que les travaux disponibles restent encore peu nombreux et insistent sur la nécessité d'intégrer ces outils dans une approche pédagogique structurée, interactive et critique [30]. Cette perspective éclaire les résultats observés : l'agent conversationnel n'agit pas comme un simple dispositif technique, mais comme un médiateur d'interaction autour du raisonnement. Les résultats de cette étude pilote vont donc dans le sens d'un bénéfice pédagogique possible, mais conditionnel, dépendant du cadre d'utilisation, du temps disponible, des connaissances préalables et de la capacité des étudiants à maintenir une posture critique face à l'outil.

4. Un potentiel réel, mais une appropriation encore incomplète

Les résultats qualitatifs montrent que la plus-value potentielle de l'agent ne s'accompagne pas d'une appropriation simple ou immédiate. Les participants ont rapporté une charge cognitive importante, une consommation de temps élevée et une tension entre approfondissement du raisonnement et contraintes temporelles. Cette difficulté est cohérente avec les données de complétude : seuls 11 participants sur 21 ont rendu les deux cas attendus, 6 ayant fourni une réponse partielle et 4 n'ayant pas complété l'exercice. Ces éléments suggèrent que le dispositif nécessitait un cadrage plus important, notamment en termes de temps disponible, d'accompagnement et de familiarisation préalable avec l'outil.
Cette interprétation est également cohérente avec le score médian du System Usability Scale, qui situait l'outil dans une zone limite de l'acceptabilité, avec une variabilité interindividuelle importante. Les participants déjà familiarisés avec l'IA semblaient mieux en identifier les usages possibles, tandis que ceux qui l'utilisaient peu ou pas rapportaient davantage de déstabilisation, de charge cognitive ou de perte de maîtrise. L'appropriation de l'agent ne semble donc pas dépendre uniquement de ses caractéristiques techniques, mais aussi du niveau de littératie IA, des connaissances disponibles pour l'interroger et de la capacité à évaluer ses réponses. Cette interprétation rejoint les travaux récents soulignant le rôle de la littératie IA, de la technophilie ou de la technophobie dans l'appropriation des chatbots et des outils d'IA générative en formation infirmière [33-35].

Enfin, les données montrent que l'acceptabilité de l'outil dépend aussi de la place qui lui est attribuée dans le raisonnement. Les participants insistaient sur la nécessité de conserver le contrôle de la décision clinique et de ne pas laisser l'agent se substituer au jugement professionnel. L'agent conversationnel apparaît donc pédagogiquement prometteur, mais son intérêt dépend d'un cadre explicite, progressif et critique. Sa valeur ne réside pas seulement dans sa disponibilité, mais dans la manière dont son usage est enseigné, encadré et intégré dans une activité de raisonnement clinique.

5. Limites de l'étude

Plusieurs limites doivent être soulignées. D'abord, l'effectif restait réduit et seuls 11 participants disposaient de données appariées complètes pour le critère principal. Cette complétude partielle limite la puissance statistique de l'étude et impose d'interpréter les résultats comme exploratoires. L'étude reposait également sur une séance unique, réalisée en distanciel et sous contrainte temporelle, ce qui a probablement influencé l'expérience rapportée par les participants et leur capacité à exploiter pleinement l'agent conversationnel.
La mesure du critère principal présente aussi des limites. Les productions écrites ont été cotées à l'aide d'une grille adaptée, mais aucune calibration préalable formelle des évaluateurs ni mesure d'accord inter-évaluateurs n'a été réalisée. La variabilité observée entre les deux séries de correction ne permet donc pas d'estimer de manière robuste l'effet de l'intervention, même si les deux corrections vont dans le même sens.
Le plan croisé expose enfin l'étude à des effets de séquence, de période et de contamination entre les deux conditions. Bien qu'une randomisation en séquences AB/BA ait été mise en place, aucune analyse spécifique de ces effets n'a été réalisée et aucune période de washout n'était prévue. Une amélioration en seconde phase pourrait donc refléter un effet d'apprentissage, de familiarisation avec la tâche ou de mobilisation progressive du raisonnement clinique, indépendamment de l'utilisation de l'agent conversationnel.
Enfin, plusieurs variables secondaires comportaient des données manquantes, et seule l'échelle SUS reposait sur un instrument validé, contrairement au questionnaire de satisfaction construit pour l'étude. Ces éléments limitent la portée généralisable des résultats.

6. Perspectives pédagogiques

Malgré ces limites, cette étude pilote apporte un résultat cohérent : un agent conversationnel basé sur l'IA semble pouvoir soutenir l'apprentissage du raisonnement clinique chez des étudiants IPA, non comme autorité clinique autonome, mais comme support de relance et de structuration réflexive.
Les résultats invitent moins à conclure que les apprenants sont déjà prêts à un usage spontané de l'IA qu'à envisager son intégration dans un programme d'enseignement structuré. Dans cette perspective, l'enjeu n'est pas seulement d'introduire l'outil dans la formation, mais d'enseigner explicitement ses usages, ses limites et les modalités de vérification critique qu'il impose.

CONCLUSION

Cette étude pilote avait pour objectif d'explorer l'apport potentiel d'un agent conversationnel basé sur l'intelligence artificielle générative dans l'apprentissage du raisonnement clinique chez des étudiants infirmiers en pratique avancée. Les résultats obtenus mettent en évidence un signal favorable : les scores médians de raisonnement clinique étaient plus élevés en condition avec IA qu'en condition sans IA, dans les deux séries de correction. Ce résultat doit toutefois rester interprété avec prudence en raison du faible effectif, de la complétude partielle des données, de l'absence de mesure formelle de l'accord inter-évaluateurs et du caractère exploratoire du dispositif.

Les données qualitatives permettent de mieux comprendre ce signal. L'agent conversationnel n'a pas été perçu comme une source autonome de réponses, mais plutôt comme un support de relance, de structuration et d'explicitation du raisonnement. Son intérêt semble donc moins résider dans la production d'une réponse correcte que dans sa capacité à soutenir un dialogue réflexif autour de la situation clinique. Cette contribution reste néanmoins conditionnelle : l'usage de l'outil demande du temps, des connaissances préalables, une posture critique et un cadre pédagogique explicite.

Ainsi, l'intégration d'un agent conversationnel dans la formation IPA apparaît comme une piste pédagogique prometteuse, mais non suffisante en elle-même. Les résultats invitent à poursuivre l'évaluation de ces dispositifs dans des études plus larges, mieux contrôlées et intégrant une formation progressive à l'usage critique de l'IA en contexte clinique.

Références :

1. Belgique. Loi coordonnée du 10 mai 2015 relative à l'exercice des professions des soins de santé [Internet]. Moniteur belge. 18 juin 2015. Disponible à: https://www.ejustice.just.fgov.be/cgi_loi/article.pl?language=fr&lg_txt=f&type=&sort=&numac_search=&cn_search=2015051006&caller=SUM&&view_numac=2015051006
2. Gerkens S, Lefèvre M, Bouckaert N, Levy M, Maertens de Noordhout C, Obyn C, et al. Performance du système de santé belge: rapport 2024. Bruxelles: Centre fédéral d'expertise des soins de santé (KCE); 2024. KCE Reports 376B. Disponible à: https://kce.fgov.be/sites/default/files/2024-01/KCE376B_HSPA2024_Rapport.pdf
3. Roulin MJ, Borrero P, Iori E. La pratique infirmière avancée: de quoi parle-t-on? Krankenpflege Soins Infirmiers. 2014;(4):68-71.
4. Hamric AB, Tracy MF. A definition of advanced practice nursing. In: Tracy MF, O'Grady ET, editors. Advanced Practice Nursing: An Integrative Approach. 6th ed. St. Louis: Elsevier; 2019. p. 61-79.
5. World Health Organization. Ageing and health [Internet]. Geneva: WHO; 2021 [cited 2026 Apr 21]. Available from: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health
6. World Health Organization Regional Office for Europe. Ticking timebomb: without immediate action, health and care workforce gaps in the European Region could spell disaster [Internet]. Copenhagen: WHO Europe; 2022 Sep 14 [cited 2026 Apr 21]. Available from: https://www.who.int/europe/fr/news/item/14-09-2022-ticking-timebomb–without-immediate-action–health-and-care-workforce-gaps-in-the-european-region-could-spell-disaster
7. World Health Organization Regional Office for Europe. *Health and care workforce in Europe: time to act* [Internet]. Copenhagen: WHO Europe; 2022. Available from: https://www.who.int/europe/publications/i/item/9789289058339
8. Organisation for Economic Co-operation and Development. Health at a Glance: Europe 2022. Paris: OECD Publishing; 2022. Disponible à: https://www.oecd.org/en/publications/health-at-a-glance-europe-2022_507433b0-en.html

9. Kilpatrick K, Savard I, Audet LA, Costanzo G, Khan M, Atallah R, et al. A global perspective of advanced practice nursing research: a review of systematic reviews. PLoS One. 2024;19(7):e0305008. doi:10.1371/journal.pone.0305008
10. McMenamin A, Turi E, Schlak A, Poghosyan L. A systematic review of outcomes related to nurse practitioner-delivered primary care for multiple chronic conditions. Med Care Res Rev. 2023;80(6):563-581. doi:10.1177/10775587231186720
11. Lecocq D, Mengal Y, Pirson M. Comment développer la pratique infirmière avancée dans des systèmes de soins de santé complexes? Santé Publique. 2015;S1(HS):105-110.
12. Newhouse RP, Stanik-Hutt J, White KM, Johantgen M, Bass EB, Zangaro G, et al. Advanced practice nurse outcomes 1990-2008: a systematic review. Nurs Econ. 2011;29(5):230-250.
13. Donald F, Martin-Misener R, Carter N, Donald EE, Kaasalainen S, Wickson-Griffiths A, et al. A systematic review of the effectiveness of advanced practice nurses in long-term care. J Adv Nurs. 2013;69(10):2148-2161. doi:10.1111/jan.12153
14. Kilpatrick K, DiCenso A, Bryant-Lukosius D, Ritchie D, Martin-Misener R, Carter N. Practice patterns and perceived impact of clinical nurse specialist roles in Canada: results of a national survey. Int J Nurs Stud. 2013;50(11):1524-1536. doi:10.1016/j.ijnurstu.2013.03.004
15. Conseil fédéral de l'art infirmier. Profil de fonction de l'infirmier de pratique avancée. Bruxelles: SPF Santé publique; 13 juin 2017. Disponible à: https://consultativebodies.health.belgium.be/sites/default/files/documents/profil_infirmier_pratique_avancee-verpleegkundig_specialist.pdf
16. Belgique. Arrêté royal du 14 avril 2024 fixant les activités cliniques et les actes médicaux que l'infirmier de pratique avancée peut exercer et les conditions selon lesquelles il peut les exercer. Moniteur belge. 26 avr 2024. Disponible à: https://etaamb.openjustice.be/fr/arrete-royal-du-14-avril-2024_n2024003886
17. International Council of Nurses. Guidelines on Advanced Practice Nursing [Internet]. Geneva: ICN; 2020. Available from: https://www.icn.ch/system/files/documents/2020-04/ICN_APN%20Report_EN_WEB.pdf
18. Tanner CA. Thinking like a nurse: a research-based model of clinical judgment in nursing. J Nurs Educ. 2006;45(6):204-211. doi:10.3928/01484834-20060601-04

19. Bryant-Lukosius D, DiCenso A. A framework for the introduction and evaluation of advanced practice nursing roles. J Adv Nurs. 2004;48(5):530-540. doi:10.1111/j.1365-2648.2004.03235.x
20. Hunter S, Arthur C. Nursing students' experiences of learning clinical reasoning. Nurse Educ Today. 2016;37:73-80. doi:10.1016/j.nedt.2015.11.016
21. Pérez-Perdomo A, Zabalegui A. Teaching strategies for developing clinical reasoning skills in nursing students: a systematic review of randomised controlled trials. Healthcare (Basel). 2024;12(1):90. doi:10.3390/healthcare12010090
22. Sim JJM, Rusli KDB, Seah B, Levett-Jones T, Lau Y, Liaw SY. Virtual simulation to enhance clinical reasoning in nursing: a systematic review and meta-analysis. Clin Simul Nurs. 2022;69:26-39. doi:10.1016/j.ecns.2022.05.006
23. Lei YY, Zhu L, Sa YTR, Cui XS. Effects of high-fidelity simulation teaching on nursing students' knowledge, professional skills and clinical ability: a meta-analysis and systematic review. Nurse Educ Pract. 2022;60:103306. doi:10.1016/j.nepr.2022.103306
24. Zhang F, Liu X, Wu W, Zhu S. Evolution of chatbots in nursing education: narrative review. JMIR Med Educ. 2024;10:e54987. doi:10.2196/54987
25. Chan MMK, Wan AWH, Cheung DSK, Choi EPH, Chan EA, Yorke J, et al. Integration of artificial intelligence in nursing simulation education: a scoping review. Nurse Educator. 2025;50(4):195-200. doi:10.1097/NNE.0000000000001851
26. Zhou Y, Tian X, Jiang S, Chen Y, Peng M. Using ChatGPT in nursing: scoping review of current opinions. JMIR Med Educ. 2024;10:e54297. doi:10.2196/54297
27. Gunawan J, Aungsuroch Y, Montayre J. ChatGPT integration within nursing education and its implications for nursing students: a systematic review and text network analysis. Nurse Educ Today. 2024;141:106323. doi:10.1016/j.nedt.2024.106323
28. Pham TD, Karunaratne N, Exintaris B, Liu D, Lay T, Yuriev E, et al. The impact of generative AI on health professional education: a systematic review in the context of student learning. Med Educ. 2025 Dec;59(12):1280-1289. doi:10.1111/medu.15746
29. Feigerlova E, Hani H, Hothersall-Davies E. A systematic review of the impact of artificial intelligence on educational outcomes in health professions education. BMC Med Educ. 2025;25:129. doi:10.1186/s12909-025-06719-5
30. Zintchem R, Ngono Zintchem MA, Ngae D, Tembe-Fokunang EA, Ntungwen Fokunang C. Intelligence artificielle générative et apprentissage du raisonnement clinique infirmier:

revue de portée, proposition d'une approche « socioconnectriste ». Recherche et Avancées en Sciences Infirmières. 2025;(2):141-177. doi:10.25965/reasci.606

31. Abujaber AA, Abd-Alrazaq A, Al-Qudimat AR, Nashwan AJ. A strengths, weaknesses, opportunities, and threats (SWOT) analysis of ChatGPT integration in nursing education: a narrative review. Cureus. 2023;15(11):e48643. doi:10.7759/cureus.48643
32. Amankwaa I, Ekpor E, Cudjoe D, Kobiah E, Fuseini AKJ, Diebieri M, Gyamfi S, Brownie S. Patterns, advances, and gaps in using ChatGPT and similar technologies in nursing education: a PAGER scoping review. Nurse Educ Today. 2025;153:106822. doi:10.1016/j.nedt.2025.106822
33. Aslan F. A descriptive phenomenological study of postgraduate nursing students' experiences with ChatGPT. J Prof Nurs. 2025;59:148-154. doi:10.1016/j.profnurs.2025.05.004
34. Blomquist J, Alderden J, Llewellyn S, Haynes M, Dworak E. Student and faculty perceptions and use of generative artificial intelligence in nursing education. J Nurs Educ. 2025;64(6):337-344. doi:10.3928/01484834-20250312-06
35. Amer FGM, Alanazi MA, Shaban M, Abdelgawad ME, Mohammed HH, Hassan S, Atrous MH, Ibrahim SSM, Hashem Mohamed A, Sabek EM. Artificial intelligence-based chatbots' and nursing students' academic performance: the mediating role of technophobia and technophilia. Int Nurs Rev. 2025;72(3):e70056. doi:10.1111/inr.70056

Annexes

Annexe 1 - LCJR-VF

Beauchamp, Jacinthe; Lalonde, Michelle; Fournier, Viviane; Mehiz, Sabrina; Pedrotti, Marco; Michel, Isabelle; Godbout, Pierre; Simoneau, Ivan; and Lasater, Kathie (2021) "Traduction et adaptation en langue française du Lasater Clinical Judgment Rubric: Une étude multicentrique," Quality Advancement in Nursing Education - Avancées en formation infirmière: Vol. 7: Iss. 2, Article 5. DOI: https://doi.org/10.17483/2368-6669.1277

LCJR-VF – Fiche de cotation

Attribuer un score de 1 à 4 pour chaque indicateur sur base de la grille LCJR-VF complète. 1 = Débutante / 2 = En développement / 3 = Accomplie / 4 = Exemplaire.

Domaine	Indicateur évalué	1	2	3	4
REMARQUER	Observation ciblée du patient / surveillance clinique	■	■	■	■
REMARQUER	Détection des écarts par rapport au normal / signes d'alerte	■	■	■	■
REMARQUER	Recherche active d'informations pertinentes (compléments, antécédents, monitorings)	■	■	■	■
INTERPRÉTER	Priorisation / tri des données cliniques importantes	■	■	■	■
INTERPRÉTER	Analyse / interprétation clinique cohérente de la situation	■	■	■	■
RÉPONDRE	Calme, maîtrise émotionnelle, posture professionnelle	■	■	■	■
RÉPONDRE	Plan d'intervention adapté, ajustements si évolution du patient	■	■	■	■
RÉPONDRE	Communication claire, précise et ciblée (équipe / patient)	■	■	■	■
RÉFLÉCHIR	Évaluation de l'effet des actions posées (auto-monitoring en temps réel)	■	■	■	■
RÉFLÉCHIR	Capacité d'auto-analyse critique après l'action	■	■	■	■
RÉFLÉCHIR	Intention d'amélioration / projection vers les bonnes pratiques futures	■	■	■	■
	Score total (somme des scores 1-4) :				_____ / 44

Dans le cadre de cette étude, l'indicateur « calme, maîtrise émotionnelle, posture professionnelle » n'a pas été coté en raison du format distanciel. Le score total utilisé pour l'analyse était donc de 40 points.

Annexe 2 - SUS

Gronier, G. & Baudet, A. (2021). Psychometric evaluation of the F-SUS: Creation and validation of the French version of the System Usability Scale. *International Journal of Human-Computer Interaction.* https://doi.org/10.1080/10447318.2021.1898828.

https://fabrique-formations.univ-lille.fr/fileadmin/user_upload/learning-fab/Documents/Echelles_REFA/Ergonomie/SUS_GRONIER_BAUDET_2021.pdf

pas du tout d'accord	très peu d'accord	moyennement d'accord	assez d'accord	tout à fait d'accord
1	**2**	**3**	**4**	**5**

[strongly disagree] <----------->[totally agree]

01	Je voudrais utiliser ce système fréquemment.	1	2	3	4	5
02	Ce système est inutilement complexe.	1	2	3	4	5
03	Ce système est facile à utiliser.	1	2	3	4	5
04	J'aurais besoin du soutien d'un technicien pour être capable d'utiliser ce système.	1	2	3	4	5
05	Les différentes fonctionnalités de ce système sont bien intégrées.	1	2	3	4	5
06	Il y a trop d'incohérences dans ce système.	1	2	3	4	5
07	La plupart des gens apprendront à utiliser ce système très rapidement.	1	2	3	4	5
08	Ce système est très lourd à utiliser.	1	2	3	4	5
09	Je me suis senti·e très confiant·e en utilisant ce système.	1	2	3	4	5
10	J'ai eu besoin d'apprendre beaucoup de choses avant de pouvoir utiliser ce système	1	2	3	4	5

Codification :
Les items 2, 4, 6, 8, 10 sont inversés.

Traitement des données :

- Pour les items impairs, soustraire un point au score brut sélectionné par l'utilisateur
- Pour les items pairs, calculer cinq moins le score brut sélectionné par l'utilisateur
- Faire la somme des 10 nouveaux scores ainsi recalculés
- Multiplier la somme des scores recalculés par 2,5. Ce qui donne un score total sur 100

7 qualificatifs ont été retenus pour qualifier l'utilisabilité d'un système :
25/100 = La pire qu'on puisse imaginer
35,7/100 = Mauvaise
50,9/100 = Acceptable
71,4/100 = Bonne
85,5/100 = Excellente
90,9/100 = La meilleure qu'on puisse imaginer

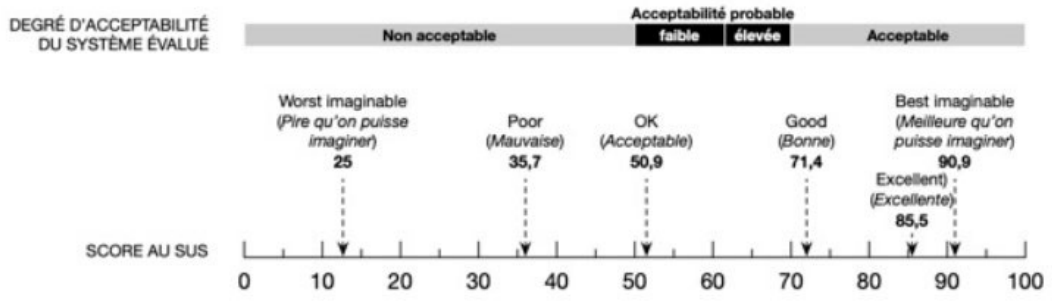


Figure 1 : Signification du score au SUS (adaptée de Bangor, Kortum et Miller, 2009).

Annexe 3 - Questionnaire démographique et de satisfaction

Une grille de recueil de données a été élaborée spécifiquement pour cette étude.
Elle comprend une section démographique permettant la description de l'échantillon et une section de satisfaction globale composée d'items à échelle de Likert.
Ce questionnaire, non validé psychométriquement mais conçu selon les principes de validité de contenu (pertinence des items) et de validité apparente (clarté et compréhension par les participants), visait à recueillir le ressenti des participants de manière simple, rapide et reproductible.
Les données démographiques permettront la description et la stratification de l'échantillon selon l'âge, le sexe, la formation, l'expérience professionnelle et le domaine d'activité.

Données sociodémographiques

Variable	Modalités de réponse proposées
Âge	☐ < 25 ans ☐ 25–34 ans ☐ 35–44 ans ☐ 45–54 ans ☐ ≥ 55 ans
Sexe	☐ Femme ☐ Homme ☐ Autre ☐ Préfère ne pas répondre
Années d'expérience clinique	☐ < 2 ans ☐ 2–5 ans ☐ 6–10 ans ☐ > 10 ans
Formation de base	☐ Bachelier ☐ Bachelier de spécialisation ☐ Master ☐ Autre (préciser) : _____
Statut / domaine actuel de pratique	☐ Infirmier employé ☐ Infirmier indépendant ☐ Étudiant ☐ Autre (préciser) : _____
Domaine clinique principal	Réponse libre : _____

Satisfaction globale

(Échelle de Likert à 5 niveaux : 1 = pas du tout d'accord / 5 = tout à fait d'accord)

Item	Affirmation	Échelle de réponse (1 à 5)
1	Je suis satisfait(e) de l'expérience proposée dans le cadre de cette étude.	☐1 ☐2 ☐3 ☐4 ☐5
2	Le dispositif m'a permis de réfléchir à ma pratique ou à mon raisonnement clinique.	☐1 ☐2 ☐3 ☐4 ☐5
3	L'activité correspondait à mes attentes.	☐1 ☐2 ☐3 ☐4 ☐5
4	Je participerais volontiers à une expérience similaire.	☐1 ☐2 ☐3 ☐4 ☐5
5	Satisfaction globale (note de 1 à 10)	☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8 ☐9 ☐10

Annexe 4 – Guide de questions pour le debriefing collectif

1. Perception de l'outil IA

Comment avez-vous perçu l'utilisation de l'intelligence artificielle dans l'exercice de raisonnement clinique ? (facilitation, limites, confiance, place potentielle dans la formation…)

2. Apport dans la réflexion clinique

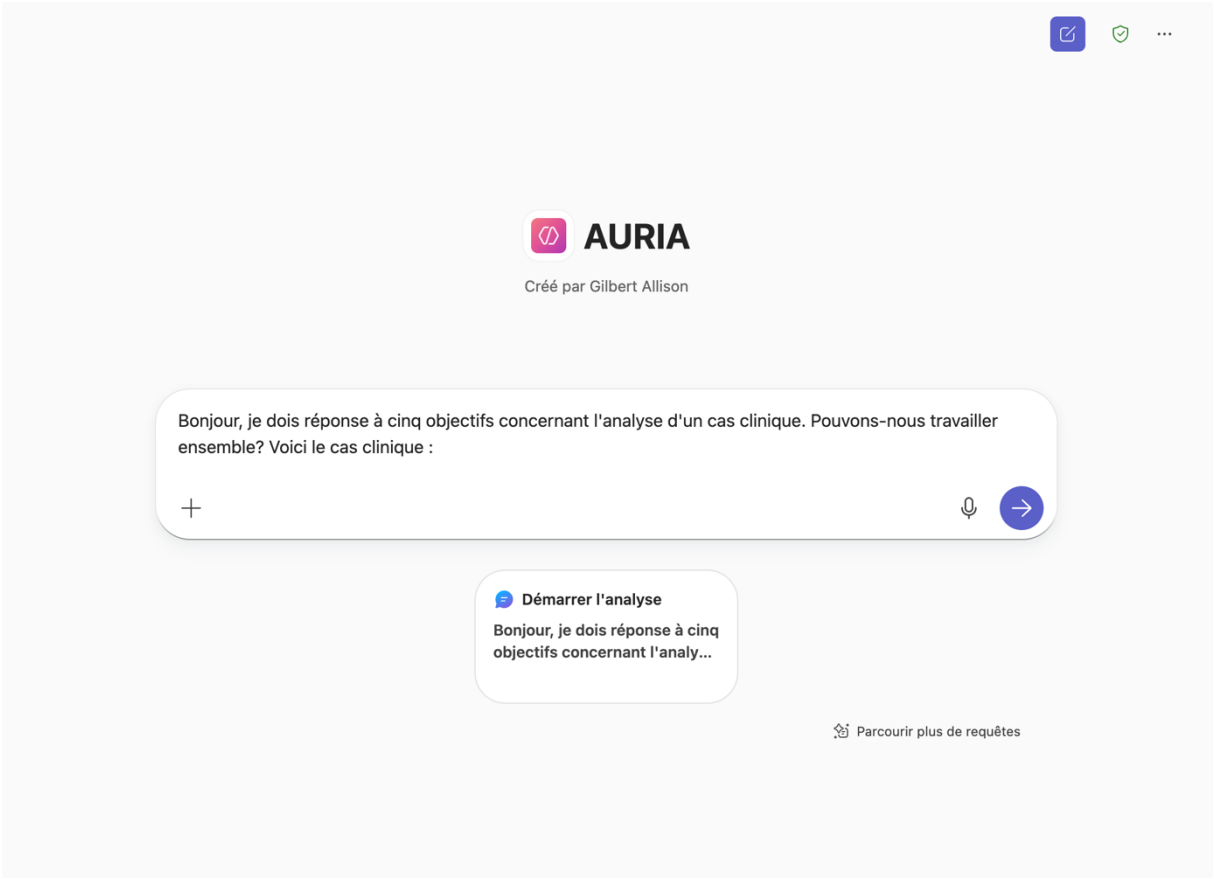
En quoi l'agent conversationnel a-t-il influencé votre manière d'analyser le cas ou de structurer votre raisonnement ?

3. Ressenti global vis-à-vis du format pédagogique

Comment avez-vous vécu ce dispositif pédagogique global ?

(motivation, engagement, équilibre entre accompagnement et autonomie…)

Annexe 5 - Illustration de l’interface et du cadre d’utilisation de l’agent AURIA



Annexe 6 - cas A

Cas A – Médecine aigue

Vous êtes infirmier en pratique avancé au sein d'un service d'urgence ayant implémenté une voie de soins rapides où votre fonction est d'évaluer les patients, établir un premier bilan, réaliser les premiers soins et examens sous supervision du médecin urgentiste, établir un premier diagnostic et entamer le traitement en collaboration avec le médecin urgentiste.

Il est dix-huit heures passées lorsque vous voyez entrer un homme de 64 ans, accompagné par son voisin. Il avance avec une boiterie marquée, le visage crispé, et dès qu'il vous aperçoit, il s'empresse de vous interpeller d'une voix inquiète :

« Regardez ça… Je suis tombé il y a quatre jours dans l'escalier de la cave. J'ai glissé sur la dernière marche, je me suis cogné fort, et ça m'a fait cette vilaine plaie. Je veux éviter une infection à tout prix. Et mon vaccin contre le tétanos… il doit être périmé, je n'y ai pas pensé depuis des années. »

Il insiste, vous tend la jambe pour que vous voyiez l'éraflure. La plaie que vous objectivez est superficielle, mais il semble très préoccupé. Il poursuit, en fronçant les sourcils :

« Depuis la chute, mon genou me fait mal, surtout quand je plie la jambe. Je me demande si je n'ai pas besoin de points de suture… ou d'une radio, parce que j'ai vraiment pris un choc.»

Il ajoute, presque suppliant, car il semble très inquiet :

« Je ne voudrais pas que ça s'infecte… Vous savez, avec mon âge, ça peut vite tourner mal. Surtout que je suis diabétique moi…»

En décrivant la chute, il marque une pause car semble essoufflé lors des explications :

« J'ai d'abord glissé… mais juste avant de tomber, j'ai senti comme une traction brusque dans mon mollet gauche. Ça m'a surpris, et puis j'ai perdu l'équilibre. »

Il grimace en se souvenant, puis reprend :

« J'ai eu mal sur le coup, mais je pensais que c'était juste à cause du choc. »

Le voisin, resté à ses côtés, intervient alors :

« Depuis une semaine, il se plaignait d'un tiraillement dans le mollet gauche, ça a peut-être favorisé sa chute. Il marchait plus lentement, devait s'arrêter quand on allait aux magasins. Il disait que c'était des crampes. »

Puis il ajoute :

« Ces derniers jours, il bricolait beaucoup, assis longtemps, souvent les jambes croisées. »

Le patient acquiesce, mais minimise :

« Oui, j'ai bricolé pas mal… mais bon, tout le monde a les jambes lourdes passé un certain âge. C'est surtout cette fichue chute qui m'a mis dedans ! »

Vous poursuivez l'entretien. Il vous confie qu'il souffre d'hypertension et d'hypercholestérolémie ainsi que de diabète, qu'il a un peu de surcharge pondérale et qu'il ne se souvient plus de la date de son dernier vaccin contre le tétanos. Il précise qu'il prend de l'amlodipine 10mg une fois par jour et de la simvastatine 20mg une fois par jour. Il prend également de la Metformine 500mg 3 fois par jour.

En approfondissant, il lâche presque comme une remarque secondaire :

« Regardez ! J'suis sûr que j'ai une fichue infection ! Depuis ce matin, ma jambe gauche est plus chaude… la douleur s'est aggravée… marcher est devenu difficile. »

Puis, en haussant les épaules :

« Et puis cette histoire ça me fatigue tellement je me tracasse. Regardez je suis tout essoufflé moi depuis que ça m'est arrivé !»

À l'examen, vous constatez que sa tension artérielle est à 155/88mmHg, sa fréquence cardiaque à 114 battements par minute, sa saturation en oxygène à 93% à l'air ambiant et sa température à 37,4°C. Le mollet gauche est douloureux en profondeur, tendu, avec un œdème discret et une chaleur locale. La jambe présente une rougeur modérée, mal délimitée. Quant au genou droit, il porte bien une éraflure superficielle sans écoulement suspect constaté.

Maintenant que vous avez pris connaissance du cas, veuillez répondre aux cinq objectifs :

1) **Résumez le cas complexe identifié.**
2) **Identifiez les éléments clés qui nécessitent une attention particulière et justifiez votre choix.**
3) **Identifiez les problèmes de santé du patient.**
4) **Établissez un plan d'action sur base des risques et attentes pluridisciplinaires requises.**
5) **Réalisez une synthèse des éléments précédents à soumettre à l'enseignant.**

Q1 = Résumez le cas complexe identifié

Insérer votre réponse ici

Q2 = Identifiez les éléments clés qui nécessitent une attention particulière et justifiez votre choix

Insérer votre réponse ici

Q3 = Identifiez les problèmes de santé du patient

Insérer votre réponse ici

Q4 = Établissez un plan d'action sur base des risques et attentes pluridisciplinaires requises

Insérer votre réponse ici

Q5 = Réalisez une synthèse des éléments précédents à soumettre à l'enseignant

Insérer votre réponse ici

Annexe 7 - cas B

Cas B – Médecine chronique

Vous êtes infirmier en pratique avancée au sein d'une maison médicale et vous exercez une fonction de coordination de soins pour les patients chroniques nécessitant un suivi plus rapproché.

Ce matin, un homme de 58 ans, connu pour un diabète de type 2 avec des contrôles glycémiques complexes, revient pour une consultation de coordination prévue par le médecin traitant de la maison médicale suite à la mise en évidence à la dernière consultation d'une plaie sur pied diabétique.

Il entre en traînant l'avant-pied droit, s'assied avec précaution et vous regarde, hésitant : **« Je viens pour le pied… la plaie sous le pied droit. Ça fait plusieurs semaines et votre collègue voulait que je sois revu donc me voilà ! »**

Vous lui rappelez que des consignes claires avaient été données et lui demandez si il les a suivies, notamment le traitement local pour le pied, les pansements, la reprise du traitement diabétique et le contrôle radiologique.

Il soupire, se frotte les mains, et cherche ses mots : **« J'ai laissé respirer, comme on fait… L'air, ça cicatrise plus vite. Et puis mon voisin m'a dit : "Mets de l'alcool, ça désinfecte tout." Alors j'ai mis de l'alcool à 90. Quand ça pique fort, c'est que ça travaille, non ? Il m'a dit aussi : "Arrête la metformine, ça fatigue le foie." J'ai arrêté pendant quelques jours. Le jus de citron, c'est naturel, ça nettoie le sucre. »**

Vous lui demandez s'il a pu se rendre au contrôle radiologique programmé par le médecin.

Il lève les yeux au plafond, agacé par avance : **« Franchement… c'est à vingt-cinq kilomètres. Pas de voiture. Le bus, c'est la matinée entière. Et si je m'arrête, je perds. Je suis intérimaire : pas de jours, pas de paye. Et la dernière fois, on m'a dit "il faut revenir" pour un autre cliché. J'ai pensé : ça allait pas si mal, ça pouvait attendre. »**

Vous revenez sur les pansements et la décharge de l'appui.

Il hausse les épaules et montre sa chaussette tachée : **« J'ai mis du sopalin avec de l'alcool, ça tient mieux dans la chaussure. Le pansement médical, ça coûte. Et l'attelle de décharge, je l'ai reçue, mais je dois marcher pour le boulot. Je ne peux pas me promener avec un truc qui m'empêche d'appuyer. Et puis… quand ça suinte, c'est bon signe, on m'a dit que "ça sort les saletés". »**

Vous lui demandez comment il ressent l'évolution au niveau de son pied : **« Depuis dix jours, je boîte un peu, ça fait mal à l'appui, mais c'est supportable. J'ai vu pire. Ma fille m'a dit que mes chaussures sentent… c'est sûrement une mycose. Et j'ai une petite toux sèche depuis quelques jours, c'est l'air froid du matin. Depuis deux jours, je suis fatigué, j'ai des frissons et moins d'appétit… Ça doit être un petit virus. Ça va passer. »**

Vous revenez sur la glycémie et l'observance thérapeutique à son traitement.

Il secoue la tête, presque amusé : **« Les bandelettes ? Trop cher. Je me fie à ce que je sens. Quand j'ai la bouche sèche, je bois du citron pressé, ça casse le sucre. Les comprimés… on verra plus tard, c'est lourd pour le foie, on m'a dit. Et puis, si je prends trop de médicaments, je vais devenir dépendant, hein ? »**

Vous relevez un front moite et un air fatigué. Sa tension est à 148/92, sa fréquence à 96bpm, sa saturation à 97 %, et sa température à 37,8 °C. Lorsqu'il retire sa chaussette, une odeur nette se dégage et la semelle intérieure est humide. Le pansement bricolé est imbibé. La plaie plantaire, au niveau de l'avant-pied droit, s'étale sur environ deux centimètres et demi ; les bords sont irréguliers, le fond jaunâtre, un écoulement léger apparaît à la pression, la chaleur locale est marquée, l'érythème dessine une nappe diffuse. La sensibilité à la piqûre est diminuée, mais la pression sur les têtes métatarsiennes déclenche une douleur profonde.

Le patient exprime : « Je sens moins sous les orteils depuis des années. Mais quand vous appuyez sur l'os, oui, ça tire. C'est normal, non ? Et pour l'odeur, j'ai mis du talc. Le talc, ça absorbe. »

Vous évaluez la recoloration capillaire : elle dépasse trois secondes. Le pouls pédieux droit est difficile à percevoir. La glycémie capillaire affiche 298 mg/dL. Les résultats labo confirment une CRP à 56 mg/L et une leucocytose à 11 200/mm³. Le dosage de HbA1c est à 9,1 %. La radiographie antérieure mentionnait le doute d'une érosion métatarsienne, sans contrôle effectué.

Maintenant que vous avez pris connaissance du cas, veuillez répondre aux cinq objectifs :

1) **Résumez le cas complexe identifié.**
2) **Identifiez les éléments clés qui nécessitent une attention particulière et justifiez votre choix.**
3) **Identifiez les problèmes de santé du patient.**
4) **Établissez un plan d'action sur base des risques et attentes pluridisciplinaires requises.**
5) **Réalisez une synthèse des éléments précédents à soumettre à l'enseignant.**

Q1 = Résumez le cas complexe identifié

Insérer votre réponse ici

Q2 = Identifiez les éléments clés qui nécessitent une attention particulière et justifiez votre choix

Insérer votre réponse ici

Q3 = Identifiez les problèmes de santé du patient

Insérer votre réponse ici

Q4 = Établissez un plan d'action sur base des risques et attentes pluridisciplinaires requises

Insérer votre réponse ici

Q5 = Réalisez une synthèse des éléments précédents à soumettre à l'enseignant

Insérer votre réponse ici

Annexe 8

L'annexe 8 présente les extraits mobilisés, les codes initiaux, les regroupements successifs et la construction des thèmes.

Axe protocole	Thème candidat	Code (action/constat)	Extrait verbatim (texte exact)	Intervenant
Ressenti global (format pédagogique)	Ressenti à chaud	Qualifie l'expérience d'intéressante	« Intéressant. »	ID9
Perception de l'outil IA	Innovant	Découvre l'IA / Qualifie l'expérience d'intéressante	« perso je n'avais jamais utilisé l'IA donc du coup j'ai trouvé ça assez intéressant comme système»	ID21
Ressenti global (format pédagogique)	Friction technique	Frustration liée à l'outil informatique	« un peu frustré de l'outil informatique qui est un peu limitant »	ID7
Perception de l'outil IA	Attentes vs réalité de l'agent	Déstabilisation puis préférence	« je m'attendais pas du tout à ça donc ça m'a un petit peu déstabilisée. Mais c'était mieux, du coup, c'était mieux comme outil. »	ID4
Perception de l'outil IA	Modèle "guide"	Stimulation du raisonnement approfondi Maintien de l'autonomie réflexive	« j'ai compris que c'est me poussait beaucoup plus dans la réflexion (…) donc d'avoir toujours le volant quoi »	ID17
Perception de l'outil IA	Contraintes temporelles	Manque de temps	« j'ai fini par tout bâcler parce que c'était trop long pour le quart d'heure. »	ID8
Ressenti global (format pédagogique)	Contraintes temporelles	Juge l'expérience "pas bonne" dans ce timing	« Mais là c'est pas une bonne expérience. »	ID8
Ressenti global (format pédagogique)	Rejet à chaud	Exprime envie d'arrêter	« Moi j'ai juste envie de le « baquer » «	ID8
Ressenti global (format pédagogique)	Contraintes temporelles	Tension entre exigence qualitative et contraintes temporelles	« il me donnait envie (…) de bien faire un truc. Beau mais là je l'ai pas fait parce que c'est impossible en 1/4 d'heure impossible »	ID8
Perception de l'outil IA	Contraintes temporelles	Usage conditionnel du temps disponible	« Si j'avais eu le temps. »	ID8
Apport dans la réflexion clinique	Temps vs performance	Préfère sans IA car plus rapide	« j'aurais préféré le faire toute seule. J'aurais été plus vite. »	ID8

Axe protocole	Thème candidat	Code (action/constat)	Extrait verbatim (texte exact)	Intervenant
Apport dans la réflexion clinique	Complétude	Voit l'agent utile pour être complet	« il est utile pour être sûr d'être bien complet. Il repose des questions, as-tu bien pensé à ça ? »	ID8
Ressenti global (format pédagogique)	Analogie pédagogique	Compare à supervision/stage ("pourquoi ?")	« j'avais l'impression d'être en stage quand la MFP nous pose des questions. Et pourquoi et pourquoi ? Pourquoi ? »	ID13
Apport dans la réflexion clinique	Profondeur vs temps disponible	Pousse à réflexion plus lointaine mais chronophage	« ça pousse à la réflexion (…) mais effectivement, il faut avoir du temps et on n'a pas toujours le temps. »	ID13
Apport dans la réflexion clinique	Non-omission	Dit que l'outil aide à ne rien négliger	« il pousse vraiment à la réflexion, à aller plus loin, à vraiment ne rien négliger, rien oublier »	ID2
Perception de l'outil IA	Contextualisation d'usage	Distingue "aigu" vs "chronique"	« si on demande (…) d'être dans l'aigu (…) il va poser problème (…) Mais pour des trucs chroniques qui suivent, c'est super intéressant »	ID2
Perception de l'outil IA	Choix d'usage situationnel	Décide selon disponibilité temporelle	« il faut pouvoir se dire là j'ai le temps-là j'ai pas le temps »	ID2
Apport dans la réflexion clinique	Stratégie personnelle	Raisonne d'abord seul puis compare à l'IA	« Je me suis d'abord posé la réflexion sur moi (…) et puis après en fait il a juste confirmé ce que je pensais »	ID1
Apport dans la réflexion clinique	Confirmation	IA confirme sans enrichir	« il a juste confirmé ce que je pensais »	ID1
Perception de l'outil IA	Utilité selon complexité	Utile si complexe, peu utile si routinier	« si on arrive sur des situations complexes (…) ça peut pousser à la réflexion. Mais dans des cas (…) assez routinier (…) je ne vois pas l'intérêt »	ID1

Axe protocole	Thème candidat	Code (action/constat)	Extrait verbatim (texte exact)	Intervenant
Perception de l'outil IA	Utilité selon contexte clinique	Peu utile en aigu (pour elle)	« si c'est une prise en charge urgente aiguë donc je voyais pas l'utilité parce que je sais ce que je dois faire »	ID1
Perception de l'outil IA	Usage conditionné par l'incertitude professionnelle	Utiliserait si obstacles en chronique	« si c'est une situation chronique (…) où moi-même j'ai pas de réflexion (…) alors là je l'utiliserais »	ID1
Apport dans la réflexion clinique	Agent limité (fonction)	Décrit l'agent comme reformulation/recentrage	« il faisait principalement que de la reformulation, du recentrage »	ID7
Apport dans la réflexion clinique	Peu de valeur ajoutée	N'aide pas au-delà de ce qu'il avait	« il m'a pas aidé à voir plus que ce que j'avais en tête »	ID7
Apport dans la réflexion clinique	Influence par formulation	Reconnaît l'effet de son prompt	« moi je l'ai influencé effectivement de par ma formulation »	ID7
Apport dans la réflexion clinique	Manque de propositions nouvelles	IA ne propose "rien en dehors"	« il m'a pas proposé grand-chose en dehors de ce que j'avais en tête »	ID7
Apport dans la réflexion clinique	Complétude	Ajoute des éléments oubliés	« m'a permis de mettre en place des choses que j'avais oubliées ou que j'avais pas inclus dans mon raisonnement »	ID2
Apport dans la réflexion clinique	Confirmation	Confirme idée de base	« l'IA m'a juste confirmé que mon idée de base était bonne »	ID2
Apport dans la réflexion clinique	Raisonnement guidé par l'IA	Part directement avec IA (sans avis préalable)	« j'ai pas posé mon propre avis à l'avance. J'ai attaqué avec IA directement »	ID8
Charge cognitive / temporalité (se retrouve dans axes 1-3)	Contraintes temporelles	Perte de temps	« j'ai perdu un temps de dingue »	ID8
Apport dans la réflexion clinique	Absorption par l'outil	Absorption cognitive par l'outil	« je me suis laissé manger par l'IA (…) je me suis presque complètement oublié »	ID12
Charge cognitive / temporalité	Énergivore	Décrit coût énergétique	« c'est extrêmement énergivore »	ID12
Perception de l'outil IA	Besoin de formation	Affirme nécessité d'être formé	« il faut être formé à ça »	ID12
Apport dans la réflexion clinique	Usage partiel	Préconise usage partiel	« j'aurais probablement dû l'utiliser par petites touches (…) et pas me focaliser autant dessus »	ID12

Axe protocole	Thème candidat	Code (action/constat)	Extrait verbatim (texte exact)	Intervenant
Perception de l'outil IA	Besoin de bases	Conditionne l'usage à des connaissances	« il faut déjà nous-même savoir plus ou moins de quoi on parle »	ID3
Perception de l'outil IA	Compétence d'usage	Nécessité de prérequis professionnels	« si on a, on n'a pas des compétences (…) on ne saurait pas l'utiliser à bonne escient »	ID3
Ressenti global (format pédagogique)	Expérience pédagogique innovante	Apprécie la dimension technologique	« Ça fait du bien de voir un petit peu de technologie et de futur »	ID7
Ressenti global (format pédagogique)	Conscience des risques	Mentionne dangers/risques de l'utilisation de l'IA	« même si il y a les dangers et tout, le risque qu'on doit avec lequel on doit travailler »	ID7
Ressenti global (format pédagogique)	Satisfaction globale	Expérience positive	« c'était un chouette cours, un chouette module »	ID7
Ressenti global (format pédagogique)	Utilité conditionnelle	"énorme plus-value" sous conditions	« je trouve que c'est une énorme plus-value (…) même si il faut y faire attention »	ID2
Ressenti global (format pédagogique)	Engagement dans le dispositif pédagogique	Temps perçu rapide	« les 4 h sont passées super vite »	ID4
Ressenti global (format pédagogique)	Améliorations logistiques	Propose préparation / dépôt Teams	« nous faire faire à l'avance (…) déposer plutôt les fichiers sur le teams plutôt que d'aller par mail »	ID4
Perception de l'outil IA	Élargissement des outils de l'IA	Prise de conscience que l'outil peut servir autrement que la reformulation	« d'habitude j'utilise l'IA surtout pour reformuler mes textes… et j'ai envie de refaire l'exercice avec ma manière habituelle »	ID8

Axe protocole	Thème candidat	Code (action/constat)	Extrait verbatim (texte exact)	Intervenant
Ressenti global (format pédagogique)	Enthousiasme + curiosité	Curiosité d'un outil non conditionné	« j'ai envie de faire l'exercice là avec mon utilisation habituelle (…) je serais quand même curieuse de le réessayer »	ID8
Ressenti global (format pédagogique)	Satisfaction	Satisfaction	« c'était top »	ID8
Charge cognitive / temporalité	Charge émotionnelle ambivalente de l'expérience	Énergivore + soulagement	« c'est quelque chose qui est très énergivore (…) je suis très contente que ça se termine »	ID13
Ressenti global (format pédagogique)	Apprentissage par erreurs	Retient "essais erreurs"	« On apprend d'essais erreurs »	ID13
Ressenti global (format pédagogique)	Regret / possibilité de correction	Aurait voulu corriger	« j'ai fait énormément d'erreurs que j'aurais voulu corriger »	ID13
Ressenti global (format pédagogique)	Valeur du groupe	Apprend via comparaison aux autres	« ça me permet de savoir ce que les autres auraient fait (…) et de peaufiner mes actions »	ID13
Ressenti global (format pédagogique)	Frustration (participation)	N'a pas su participer à la 2e partie	« je suis super frustrée parce que j'ai pas su participer à la 2e partie »	ID6