

AIX-MARSEILLE UNIVERSITÉ

FACULTÉ DE MÉDECINE DE MARSEILLE
ÉCOLE DOCTORALE : SCIENCES DE LA VIE ET DE LA SANTÉ

THÈSE

Présentée et publiquement soutenue devant
LA FACULTÉ DE MÉDECINE DE MARSEILLE

Développement des outils pour la formation à l'hystérectomie totale
par voie coelioscopique sur un simulateur de réalité virtuelle

Le 9 Mai 2017

par Patrice CROCHET

Né le 4 janvier 1978 à Romilly-sur-Seine (10)

Pour obtenir le grade de DOCTEUR de l'UNIVERSITÉ AIX-MARSEILLE

Spécialité : Pathologie humaine - Recherche clinique et santé publique

Membres du jury de la Thèse :

Monsieur le Professeur Léon BOUBLI

Président du jury

Monsieur le Professeur Aubert AGOSTINI

Directeur de thèse

Monsieur le Professeur Olivier GRAESSLIN

Rapporteur

Monsieur le Professeur Benoit RABISCHONG

Rapporteur

REMERCIEMENTS :

Aux membres du Jury :

Au Pr Léon BOUBLI,

merci d'avoir accepté de présider ce jury. Le semestre passé dans votre service reste un excellent souvenir. Votre actif soutien depuis le début de ce projet a été une grande aide.

Au Pr Aubert AGOSTINI,

merci d'avoir accepté de diriger cette thèse et de m'avoir fait confiance tout au long de ma formation à la Conception.

Au Pr Olivier GRAESSLIN,

merci d'avoir accepté d'être rapporteur de ma thèse, et de siéger dans ce jury. Votre service est celui dans lequel j'ai décidé de devenir Gynécologue Obstétricien.

Au Pr Benoit RABISCHONG,

merci d'avoir accepté d'être rapporteur de ma thèse, et de siéger dans ce jury. Je suis admiratif de ton talent à promouvoir sans relâche la simulation pour l'enseignement de la chirurgie gynécologique et espère que nous pourrons mener des projets communs à l'avenir.

Au Dr Rajesh AGGARWAL,

thank you for having co-directed this thesis. Thank you for having welcomed me at St. Mary's a few years ago, and taught me rigorous methodology in the field of pedagogy. Your constant support is of great value to me. I wish you and Shilpee and children all the best in Montreal.

Au Pr Stéphane BERDAH,

merci de m'avoir soutenu et accueilli dans ton laboratoire.

À mes parents, Claudine et Gérard

À ma sœur Diane, à Gergely, à Livia et David

À Audrey

À mes amis : Julien, Guillaume, Frédérique et Thomas, Emilie et Remi, Alan

Au noyau dur du 6^{ème} : Claire, Maxime, Emmanuelle, Laura, Sabine, Julie

À la Sim' Team de St.Mary's, à Pram, à Amit, à Sophie et Antoine

*À mes chers Maîtres et collègues de la Conception : Pr Blanc, Pr Gamberre, Pr Cravello,
Pr Courbière, Hélène, Béatrice, et à Françoise*

À tous les participants des études, merci pour votre temps et votre disponibilité

RÉSUMÉ

Rationnel : L'intérêt de la simulation pour apprendre la chirurgie mini-invasive est démontré principalement pour les gestes de base. Aucune formation dédiée à l'hystérectomie totale coelioscopique (HC) n'a été développée sur simulateur. Une telle formation pourrait améliorer la sécurité des patientes et réduire la courbe d'apprentissage au bloc opératoire (BO) pour les internes. Ceux-ci ont en effet un accès limité à cette intervention complexe en tant que premier opérateur.

Objectifs : Évaluer la validité d'un programme reproduisant une HC sur un simulateur de réalité virtuelle (RV), et développer un parcours d'entraînement basé sur des preuves. Développer une échelle d'évaluation spécifique des compétences techniques pour l'HC (H-OSATS), puis tester sa validité lorsque l'intervention est réalisée sur un simulateur de RV.

Méthodes : Au cours d'un travail préliminaire, un curriculum a été développé pour la cholécystectomie coelioscopique. Une deuxième étude a montré que l'application de la pratique délibérée sur simulateur de RV permettait d'améliorer la qualité opératoire lors d'une cholécystectomie. **Étude 1.** Quarante participants ont été évalués lors de la réalisation d'une HC sur un simulateur de RV. La preuve de validité retenue était la capacité du programme de RV à différencier 3 niveaux d'expérience (expérimenté n=8, intermédiaire n=8 et inexpérimenté n=14). Les opérateurs inexpérimentés ont réalisé 10 interventions pour analyse de la courbe d'apprentissage. Les performances étaient évaluées par les paramètres du simulateur et par une échelle d'évaluation globale de la qualité opératoire (OSATS).

Étude 2. L'échelle H-OSATS a été développée par une analyse séquentielle, puis une méthode Delphi pour obtenir un consensus auprès d'experts internationaux. La validité d'H-OSATS a été testée en évaluant les procédures réalisées par les 3 groupes d'opérateurs d'expérience différente sur un simulateur de RV.

Résultats : **Étude 1.** L'analyse quantitative montrait une différence entre les groupes expérimentés, intermédiaires et inexpérimentés pour le temps opératoire (1369, 2385 et 3370s; $p < 0.001$), le nombre de mouvements (2033, 3195 et 4056; $p = 0.001$), la distance parcourue (3390, 4526 et 5749 cm; $p = 0.002$), le respect pour les tissus (24, 40 et 84; $p = 0.01$) et le nombre de plaies vésicales (0.13, 0 et 4.27; $p < 0.001$). Une courbe d'apprentissage existait pour tous ces critères avec plateau entre le 2^{ème} et le 6^{ème} essai. Les scores de qualité OSATS

étaient différents entre les groupes lors du premier essai (22, 15 et 8 respectivement ; $p < 0.001$) et du deuxième essai (25.5, 19.5 et 14 ; $p < 0.001$). **Étude 2.** L'analyse séquentielle identifia 76 étapes consécutives. 14 experts ont répondu au Delphi, et un consensus fut obtenu pour l'intégration de 64 étapes. Il y avait une différence entre les groupes expérimentés, intermédiaires et inexpérimentés pour le score H-OSATS (178.25, 155.9 et 133; $p = 0.002$). Il y avait une bonne corrélation entre les scores H-OSATS et OSATS ($\rho = 0.928$; $p < 0.001$).

Conclusion : Le programme de RV pour l'HC a montré des preuves de validité et permis la construction d'un curriculum d'entraînement fondé sur une démarche scientifique. L'étape suivante sera de déterminer si ce programme améliore les compétences techniques lors du transfert au BO. L'échelle H-OSATS est valide pour évaluer les performances techniques sur le simulateur de RV. Son utilisation permettra de faciliter la pratique délibérée lors de la formation à l'HC, du simulateur de RV au BO.

Mots-clefs : éducation, gynécologie, réalité virtuelle, formation, évaluation.

Nombre de mots : 532

ABSTRACT

Background: Substantial evidence in scientific literature supports the use of simulation for training on laparoscopic basic skills. So far, there is no simulation-based training program specifically dedicated to total laparoscopic hysterectomy (LH). Such training could reduce learning curve and provide safe implementation in the operating room (OR) for junior surgeons who have limited access to this complex procedure as a primary operator.

Objectives: To evaluate the validity of a LH program on a virtual reality (VR) simulator, and to develop an evidence-based and stepwise training curriculum. To develop an objective scale specific for the assessment of technical skills for LH (H-OSATS) and to demonstrate feasibility of use and validity in a VR setting.

Methods: In a preliminary study, an evidence-based VR training curriculum was developed for laparoscopic cholecystectomy. A second study applied deliberate practice to a laparoscopic cholecystectomy VR training, leading to higher quality of surgical skills. **Study 1.** Forty participants were tested for the LH procedure on a high-fidelity VR simulator. Validity evidence was proposed as the ability to differentiate between 3 levels of experience (experienced n=8, intermediate n=8 and inexperienced n=14). Inexperienced subjects performed ten repetitions for learning curve analysis. Outcome measures were simulator-derived metrics and Objective Structured Assessment of Technical Skills (OSATS) scores.

Study 2. The H-OSATS scale was developed using a hierarchical task analysis (HTA) and a Delphi panel of international experts. H-OSATS was evaluated by reviewing video recordings of LH performed on a VR simulator by 3 groups of operators of different experience level.

Results: Study 1. Quantitative analysis found intergroup differences between experienced intermediate and inexperienced groups for time (1369, 2385 and 3370 s; $p < 0.001$), number of movements (2033, 3195 and 4056; $p = 0.001$), path length (3390, 4526 and 5749 cm; $p = 0.002$), respect for tissue (24, 40 and 84; $p = 0.01$) and number of bladder injuries (0.13, 0 and 4.27; $p < 0.001$). Learning curves plateaued at the 2nd to 6th repetition. Further qualitative analysis found intergroup OSATS score differences at first repetition (22, 15 and 8 respectively; $p < 0.001$) and second repetition (25.5, 19.5 and 14; $p < 0.001$). **Study 2.** A total of 76 discrete steps were identified by the HTA. 14 experts completed the Delphi questionnaire. 64 steps

reached consensus and were integrated in the scale. There was a significant difference between the experienced, intermediate and inexperienced group for total H-OSATS scores (178.25, 155.9 and 133 respectively; $p=0.002$). High correlations were found between total H-OSATS scores and total OSATS scores ($\rho=0.928$; $p<0.001$).

Conclusion: The LH program accrued validity evidence and allowed the development of a VR training curriculum using a structured scientific methodology. A next step will be to evaluate whether this training program leads to improved operative skills when transferred to the OR. The H-OSATS scale is a valid instrument for assessment of technical performances on a VR simulator. The implementation of this scale is expected to facilitate deliberate practice during LH training from a VR setting to the OR.

Key words : education, gynecology, virtual reality, training, assessment.

Word count : 484

LISTE DES ABRÉVIATIONS

BO : bloc opératoire

RV : réalité virtuelle

OSATS : objective scale for assessment of technical skills

H-OSATS : objective scale for assessment of technical skills for total laparoscopic
hysterectomy

HC : hystérectomie totale par voie coelioscopique

HTC : habileté technique chirurgicale

TABLE DES MATIÈRES

| | |
|--|-----------|
| 1. PRÉAMBULE | 11 |
| 2. INTRODUCTION | 12 |
| 3. ANALYSE BIBLIOGRAPHIQUE | 14 |
| 3.1 Stratégies d'apprentissage en chirurgie | 14 |
| 3.1.1 Le compagnonnage | 14 |
| 3.1.2 La pratique délibérée | 15 |
| 3.1.3 Le curriculum | 17 |
| 3.1.3.1 Définition | 17 |
| 3.1.3.2 Curricula en chirurgie mini-invasive | 18 |
| 3.2 Simulation : outil de formation chirurgicale | 21 |
| 3.2.1 Les niveaux de Kirkpatrick | 21 |
| 3.2.2 Impact dans le domaine de la santé | 22 |
| 3.2.3 Impact en chirurgie mini-invasive | 23 |
| 3.3 Évaluation des compétences techniques chirurgicales | 23 |
| 3.3.1 Les modes d'évaluation | 24 |
| 3.3.1.1 L'évaluation quantitative | 24 |
| 3.3.1.2 L'évaluation qualitative | 25 |
| 3.3.1.3 La détection des erreurs | 28 |
| 3.3.1.4 Les conséquences opératoires | 28 |
| 3.3.2 La compétence technique | 29 |
| 3.4 Validité en sciences de l'éducation | 29 |
| 3.5 Simulation en cœlioscopie gynécologique | 33 |
| 3.5.1 Les modèles décrits | 33 |
| 3.5.1.1 Exercices non spécifiques | 33 |
| 3.5.1.2 Gestes spécifiques de la chirurgie gynécologique | 34 |
| 3.5.2 Les curricula | 36 |
| 3.5.2.1 Utilisation du FLS en gynécologie | 36 |
| 3.5.2.2 Le curriculum de Shore | 36 |
| 3.5.2.3 Le curriculum pour la salpingectomie | 36 |

| | |
|--|------------|
| 4. EXPOSÉ DES TRAVAUX PRÉLIMINAIRES | 38 |
| 4.1 Curriculum sur simulateur de réalité virtuelle (RV) pour la cholécystectomie | 38 |
| 4.1.1 Résumé | 38 |
| 4.1.2 Publication | 40 |
| 4.2 La pratique délibérée sur simulateur de RV améliore la qualité opératoire | 48 |
| 4.2.1 Résumé | 48 |
| 4.2.2 Publication | 49 |
| 5 DÉVELOPPEMENT D'UN CURRICULUM SUR SIMULATEUR DE RV POUR L'HYSTÉRECTOMIE TOTALE PAR VOIE CŒLIOSCOPIQUE (HC) | 57 |
| 5.1 Résumé | 57 |
| 5.2 Publication | 60 |
| 6 DÉVELOPPEMENT D'UNE ÉCHELLE D'ÉVALUATION SPÉCIFIQUE DE L'HC: | |
| H-OSATS | 69 |
| 6.1 Résumé | 69 |
| 6.2 Article soumis | 71 |
| 7. DISCUSSION ET PERSPECTIVES | 110 |
| 7.1 Intérêt d'un programme de simulation pour l'HC | 110 |
| 7.2 Intégration du curriculum de RV dans la formation globale des internes | 115 |
| 7.3 Intérêt de l'échelle spécifique H-OSATS | 117 |
| 7.4 Perspectives | 118 |
| 7.4.1 Étude de la validité de l'échelle H-OSATS pour évaluer les compétences au BO | 118 |
| 7.4.2 Validation de l'impact du curriculum de RV sur l'habileté technique pour l'HC | 119 |
| 8. RÉFÉRENCES BIBLIOGRAPHIQUES | 123 |
| 9. ANNEXES | 132 |
| 9.1 Structure pour développer un curriculum basé sur la simulation, d'après Zevin et al. | |
| 9.2 Utilisation des simulateurs pour former les internes de chirurgie gynécologique en France : un état des lieux en 2013 | |

1 PRÉAMBULE

J'ai effectué le premier cycle des études médicales et l'externat à la Faculté de Médecine de Reims. De ce parcours initiatique commun à tous les futurs médecins, je garde d'excellents souvenirs et surtout une base indispensable à tout exercice médical. Lors de ces années, le mot de simulation n'était jamais employé par nos enseignants. Pourtant, les séances de dissections au laboratoire d'anatomie étaient déjà un apprentissage lié à ce concept.

Le potentiel de cette méthode de formation ne m'apparut que plus tard, lors d'un stage à Stanford dans l'équipe du Dr Camran NEZHAT. Les interventions réalisées étaient presque exclusivement coelioscopiques et cette expérience riche d'enseignements. Les étudiants étrangers ne sont pas admis à participer aux soins dans le système américain, et les seuls outils disponibles pour s'exercer étaient les pelvi-trainers de l'hôpital. C'est au retour à Marseille que je pris conscience que ces heures d'entraînement hors du bloc m'avaient donné une aisance gestuelle qu'il aurait été plus difficile d'acquérir dans le cadre du compagnonnage.

Cette expérience m'a incité à poursuivre dans ce domaine, et à réaliser une année recherche à Londres en 2008 dans le laboratoire du Pr DARZI. Au sein de cette équipe pionnière, j'ai pris conscience de l'intérêt des nouvelles technologies, et en particulier de la simulation, pour faire évoluer le système d'enseignement de la chirurgie. Ceci pour répondre aux attentes de la société, mais également pour améliorer les compétences des chirurgiens et leur capacité à maîtriser les nouvelles techniques issues de la révolution endoscopique.

2 INTRODUCTION

L'enseignement hors du bloc opératoire (BO) s'est largement développé ces 10 dernières années.¹ Les changements d'organisation du temps de travail et les contraintes limitant l'accès au BO pour la formation favorisent l'utilisation de la simulation. Celle-ci est également encouragée par les pouvoirs publics et les usagers pour des raisons éthiques, ce que résume bien la formule du rapport de l'HAS en 2012 : « ne jamais faire un geste chirurgical pour la première fois sur un patient ». Délocaliser une partie de l'enseignement est aussi une opportunité pour améliorer la formation et répondre aux défis posés par l'évolution rapide des conditions d'exercice de la chirurgie. La coelioscopie a beaucoup contribué au développement de la simulation. Les axes de recherche ont jusqu'ici principalement porté sur l'enseignement des compétences techniques, montrant un impact positif sur les performances au BO.² En coelioscopie gynécologique, les applications à une large échelle sont actuellement limitées aux compétences de base. Les travaux de cette thèse sont issus de l'hypothèse qu'un enseignement sur simulateur de réalité virtuelle (RV) peut trouver sa place pour une intervention gynécologique complexe, l'hystérectomie totale par voie coelioscopique (HC). En effet, cette intervention est peu accessible aux internes en tant qu'opérateur, et associée à une longue courbe d'apprentissage sur le plan des compétences techniques et de la prise de décision opératoire.

Deux travaux préliminaires sont d'abord présentés, réalisés lors d'une année recherche à l'Imperial College de Londres, sous la direction du Dr AGGARWAL. Le premier travail est la création d'un curriculum d'entraînement sur un simulateur de RV pour une intervention coelioscopique complète : la cholécystectomie. Le 2^{ème} travail illustre la mise en œuvre d'une stratégie d'amélioration de la performance, la pratique délibérée, lors d'une

formation sur simulateur de RV pour la cholécystectomie.

Les travaux de cette thèse ont été co-dirigés par le Pr Aubert AGOSTINI (Aix-Marseille Université) et le Dr Rajesh AGGARWAL (McGill University). Ils ont été réalisés à l'Université d'Aix-Marseille entre 2013 et 2016, au Centre d'Enseignement et de Recherche Chirurgicale (CERC) dirigé par le Pr Stéphane BERDAH.

Le premier objectif de cette thèse était de développer un curriculum d'entraînement pour l'HC sur un simulateur de RV selon une méthode scientifique. Le deuxième objectif était de construire une échelle d'évaluation spécifique de cette intervention, qui puisse être un outil utile à l'entraînement sur simulateur de RV et favorisant la pratique délibérée.

Les études publiées sont présentées sous la forme article. Pour une meilleure lisibilité, la police de l'article soumis est différente de celle utilisée pour le reste de la thèse.

3 ANALYSE BIBLIOGRAPHIQUE

3.1 Stratégies d'apprentissage en chirurgie

3.1.1 *Le compagnonnage*

Le compagnonnage a toujours eu cours dans le domaine chirurgical. Cet apprentissage « au bloc » et « au lit du malade » a été théorisé par William Halsted au début du 20^{ème} siècle, et reste aujourd'hui un pilier incontournable de la formation des jeunes chirurgiens.³ Ce système attribue au chirurgien sénior un double rôle : le soin apporté à son patient et simultanément la formation de son élève. Le sénior qui assure donc à la fois la fonction de modèle et d'enseignant, accompagne les progrès de son élève en lui faisant faire des gestes de plus en plus complexes au BO.¹

Cette méthode a fait ses preuves et permis de former de nombreux chirurgiens compétents. Son succès repose sur l'implication des protagonistes et un enseignement en trois temps : avant, pendant et après le geste opératoire réalisé.⁴ Cependant l'évolution du système de soins dans les pays industrialisés contribue à limiter l'application du compagnonnage dans cette forme classique. Il faut du temps pour que le senior connaisse le niveau de l'interne et adapte son enseignement au BO. Or le temps passé par les internes à l'hôpital a diminué et est actuellement fixé par la loi à un maximum de 48 heures hebdomadaires au niveau européen.⁵ De plus, les avancées technologiques augmentent considérablement les coûts, alors que les contraintes financières sont devenues un enjeu majeur de la pratique chirurgicale. Les opportunités de formation sont donc moins importantes que par le passé.

Le compagnonnage a plusieurs limites intrinsèques. La qualité de la relation sénior - interne influe beaucoup sur son efficacité, qui est remise en cause lorsqu'il n'y a pas une

bonne entente entre les deux parties. De plus, le sénior dispense l'enseignement et juge les résultats de celui-ci, ce qui est préjudiciable à l'objectivité de la formation. Selon une enquête auprès des intéressés, le retour d'expérience prodigué au BO est jugé bien plus sévèrement par les internes que par les séniors.⁶ Les opportunités de gestes techniques sont dépendantes du programme opératoire et il n'y a pas d'articulation logique entre le niveau de l'interne et le niveau de difficulté des gestes susceptibles de lui être délégués. Par ailleurs, il n'y a pas de répétition possible des gestes pour l'interne qui devra attendre une nouvelle opportunité. Le BO est un lieu peu propice à l'apprentissage, car bruyant et stressant. Le sénior responsable y est dans une situation ambiguë vis à vis de son patient, car les priorités des soins et celles de la formation de l'interne ne sont pas convergentes. Les légitimes attentes de la société et les principes éthiques énoncés dans le rapport de l'HAS de ne jamais faire un geste pour la première fois sur un patient sont en contradiction avec tous ces aspects du compagnonnage.⁷ Il est donc nécessaire de développer une méthode complémentaire de formation qui puisse permettre aux élèves de débiter leur courbe d'apprentissage en dehors du BO (*pre-trained novices*).⁸

3.1.2 La pratique délibérée

Anders Ericsson est un psychologue suédois qui a travaillé sur la manière dont s'acquiert l'expertise. Selon lui, les différences de niveau entre professionnels dépendent de la mise en œuvre d'un type spécifique d'entraînement, appelé « pratique délibérée ».⁹ Cette théorie est basée sur l'hypothèse que la performance d'individus ayant acquis le statut d'expert est le fruit de la mise en œuvre d'activités « délibérément » choisies pour leur capacité à améliorer les performances et à les maintenir au plus haut niveau. Cette théorie repose sur cinq éléments clefs : des exercices bien définis focalisés sur les gestes importants,

la notion d'effort, la motivation, un retour d'expérience immédiat (feedback) et la répétition d'exercices de difficulté croissante. Il a ainsi été prouvé dans d'autres domaines nécessitant un haut niveau de performance technique - tels que la musique classique, les échecs ou le sport - que le niveau d'expertise dépend directement du temps consacré aux activités dites de pratique délibérée.

Dans sa publication initiale, Ericsson présente une étude retraçant le parcours de trois catégories différentes de violonistes. Il souligne que les meilleurs interprètes (tous solistes ayant un potentiel au niveau international) consacraient plus de temps à répéter seuls (ce qui est considéré par tous les violonistes comme fondamental pour s'améliorer), et ce quotidiennement pendant une période de plus de dix ans. De plus, ils réalisaient cet entraînement avec le but précis d'améliorer leur niveau et reconnaissaient que ces temps d'entraînement leur demandaient des efforts et étaient même fastidieux (en contraste avec le témoignage des violonistes de niveau inférieur). Ils consacraient également moins de temps aux loisirs dans leur emploi du temps. Un autre exemple typique de pratique délibérée est la prescription par un coach d'activités spécifiquement construites dans le but d'améliorer un aspect particulier de la performance d'un athlète. Dans le cas des joueurs de tennis, l'entraînement à la volée de revers est particulièrement démonstratif : les occasions de réaliser ce geste sont imprévisibles et peu fréquentes au cours d'un match. L'entraîneur peut donc demander au joueur de répéter spécifiquement des volées de revers, d'abord d'exécution facile puis plus difficile, avec pour but ultime d'être capable de l'exécuter de manière parfaite lorsque l'occasion se présentera au cours du jeu. Ces stratégies consistant à travailler isolément un geste technique permettent aux athlètes de dépasser leur niveau de performance globale. De plus, le fait de recevoir une évaluation immédiate ainsi que des

conseils au cours de ces séances de répétition d'un même mouvement permet aux athlètes d'accélérer encore leur progression.

Cette théorie est adaptable dans le domaine de la médecine et de la chirurgie.⁹ L'apprentissage de la chirurgie est fondé sur la répétition des gestes, le retour d'expérience d'un formateur, la définition d'objectifs de performance et la mise en place de programmes d'entraînement structurés. Ces conditions sont difficiles à réunir en salle opératoire, mais l'émergence des simulateurs de coelioscopie donne l'occasion d'appliquer plus facilement les principes de la pratique délibérée que dans le cadre du compagnonnage. La pratique délibérée appliquée lors d'un entraînement sur simulateur de RV a montré son efficacité pour la cholécystectomie par coelioscopie lors du transfert sur un modèle porcin (**travail préliminaire n°2**)^{10,11} et au BO.¹²

3.1.3 *Le curriculum*

3.1.3.1 Définition

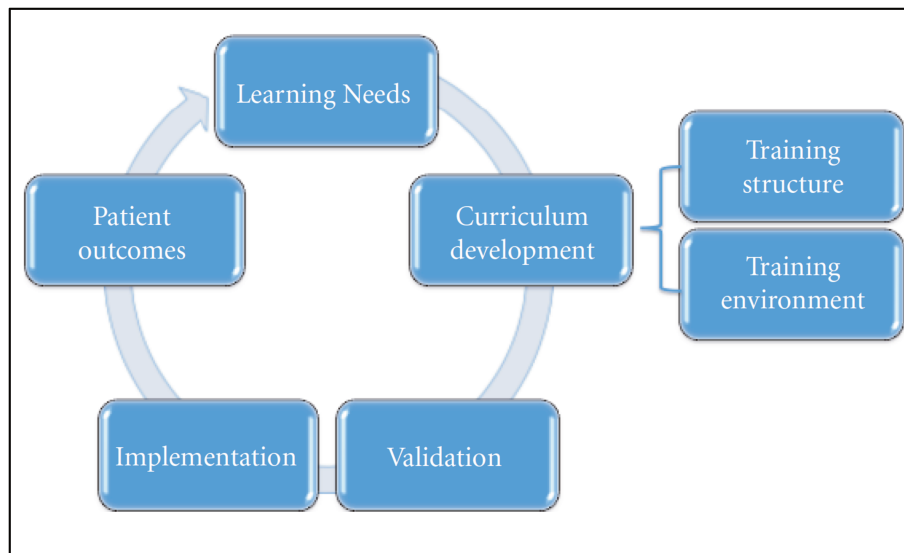
Le curriculum est un terme qui désigne, dans la littérature anglo-saxonne, un parcours éducatif constitué par des activités d'apprentissage structurées et une évaluation des acquis. Le General Medical Council est l'équivalent pour le Royaume-Uni de l'Ordre des Médecins et assure également une mission de régulation des lieux d'enseignement. Il donne la définition suivante d'un curriculum :

« A statement of the intended aims and objectives, content, experiences, outcomes and processes of an educational programme including:

- a description of the training structure [entry requirements, length and organisation of the programme including its flexibilities, and assessment system]*
- a description of expected methods of learning, teaching, feedback and supervision ».*

Le curriculum ainsi défini n'est pas établi une fois pour toutes, mais il doit être évalué périodiquement et éventuellement modifié, pour plus d'efficacité ou bien pour s'adapter à l'évolution des outils d'apprentissage (Fig A).¹³

Figure A : Processus évolutif d'un curriculum, d'après Ahmed et al.¹³



3.1.3.2 Curricula en chirurgie mini-invasive

La chirurgie mini-invasive nécessite un entraînement technique très spécifique et a beaucoup contribué au développement des modèles de simulation chirurgicale dans les centres de formation. Cet environnement sécurisé permet de limiter les contraintes liées à l'apprentissage au BO. Cependant le simulateur est uniquement un outil dont l'utilisation influe beaucoup sur le rendement de cet apprentissage. Le curriculum doit remplir un certain nombre de critères pour être efficace¹⁴ :

- Les conditions d'utilisation du simulateur ont prouvé leur efficacité.
- Les exercices sont de complexité croissante, suivant un enchaînement logique par rapport au but recherché.

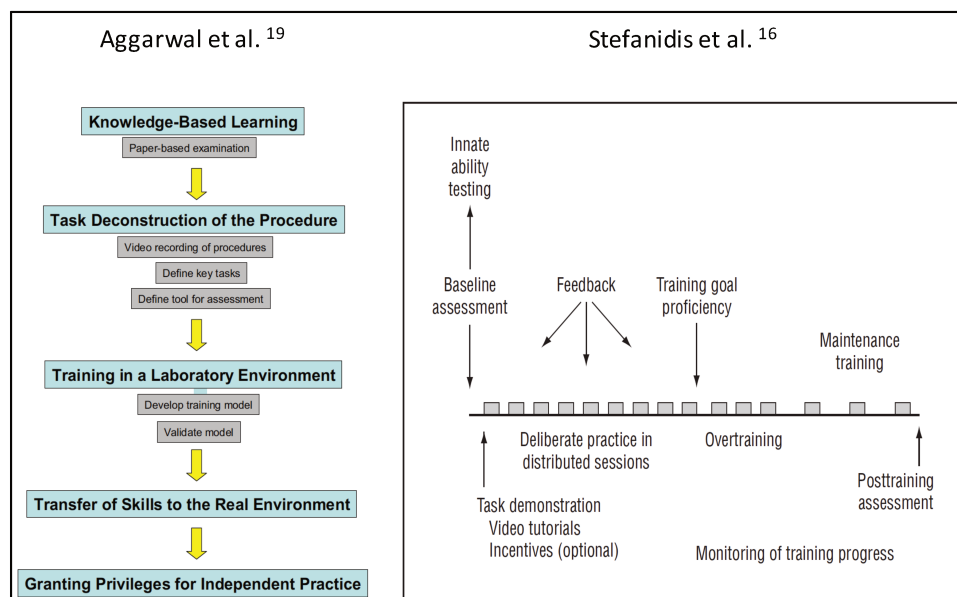
- La répétition des séances d'entraînement est codifiée. Une pratique distribuée dans le temps est plus efficace que le même nombre d'heures effectuées dans une même journée.¹⁵ Des séances d'une heure par jour maximum sont recommandées selon des avis d'experts.¹⁶
- Les critères de performance attendue ont été définis par celle d'opérateurs qualifiés. Le niveau atteint par les élèves est ainsi plus homogène que si le critère était le nombre de séances passées sur le simulateur et permet à chacun de suivre sa propre courbe d'apprentissage jusqu'à atteindre ce niveau prédéfini.
- Des interventions peuvent être proposées pour faciliter l'apprentissage au cours du curriculum. Le feedback entre les séances par l'intermédiaire d'un formateur qualifié est très efficace¹⁷, de même que l'analyse rétrospective des performances enregistrées sur supports vidéos.¹⁸

La structure d'un curriculum de formation chirurgicale par la simulation a été théorisée en 5 points par Aggarwal et al. (Fig B) :¹⁹ Après une première étape d'acquisition de connaissances théoriques, les éléments clefs de l'intervention et les erreurs à éviter sont enseignées ; l'élève est ensuite invité à répéter en laboratoire des exercices de difficulté croissante sur un simulateur. Lorsque les critères de performance attendus ont été atteints sur le simulateur, l'étape suivante est de confirmer que les compétences acquises en laboratoire sont valables dans des conditions réelles, ou bien permettent de diminuer significativement la courbe d'apprentissage au BO. Cette dernière étape du transfert dans des conditions réelles est un élément essentiel pour juger de l'efficacité d'un curriculum.

Une autre proposition de structure de curriculum a été proposée par Stefanidis et Heniford:¹⁶ dans ce schéma, les compétences techniques de base de l'élève sont évaluées,

puis la procédure chirurgicale est enseignée à l'aide de vidéos didactiques, avant de débiter des séances d'entraînement sur un simulateur en appliquant les principes de la pratique délibérée ; l'entraînement est poursuivi au-delà du moment où le seuil de performance est atteint dans le but de consolider les acquis et de les maintenir ; enfin une évaluation après le programme d'entraînement est réalisée.

Figure B : 2 propositions de structure de curriculum pour la formation chirurgicale



Un consensus entre experts internationaux de différentes spécialités a été publié par Zevin et al. pour le développement d'un curriculum de formation chirurgicale.²⁰ Cette proposition reprend les principes énoncés par Aggarwal et Stefanidis et distingue plusieurs étapes :

- Une analyse préalable pour identifier les besoins
- La construction du curriculum en utilisant un modèle de simulation et en associant :
 - Une composante cognitive
 - Une composante technique

- Une phase de validation du curriculum pour démontrer son efficacité au bloc opératoire, avant de l'appliquer à une large échelle et de proposer des améliorations
- Une phase de maintenance des compétences dans le temps.

Deux éléments sont spécifiés par les auteurs : l'intégration d'un module d'entraînement aux compétences non-techniques (capacité à travailler en équipe, gestion du stress...), et la création d'un outil d'évaluation spécifique de l'intervention concernée. Le détail de ce consensus est présenté en **annexe 1**.

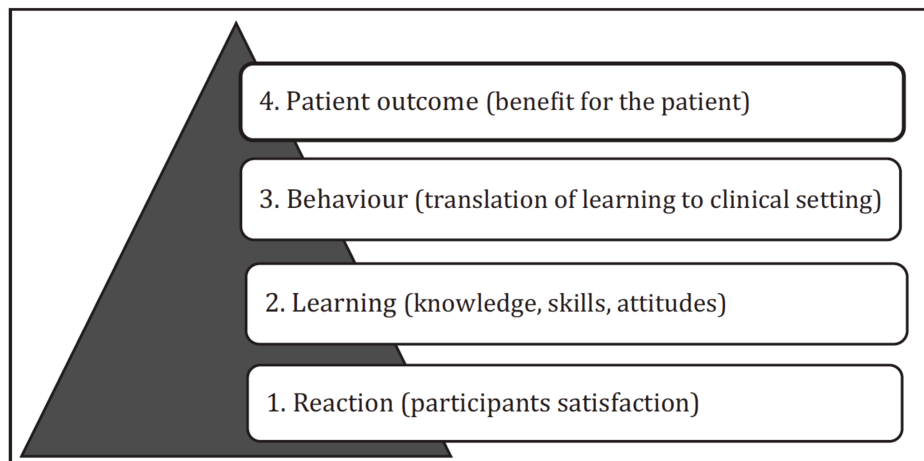
À ce jour, un seul curriculum (sur un pelvi-trainer) est utilisé à une large échelle pour l'apprentissage des gestes de base de la cœlioscopie : le programme Fundamental of Laparoscopic Surgery.^{21,22} Il est requis dans le cadre de la certification par l'American Board of Surgery depuis 2009.

3.2 La simulation : outil de formation chirurgicale

3.2.1 les niveaux de Kirkpatrick

La mesure de l'efficacité d'un exercice de simulation pour les professionnels de santé est mesurée selon la classification de Kirkpatrick qui comporte 4 niveaux.²³ Le premier niveau recueille la satisfaction des participants vis-à-vis du programme réalisé, sans que l'efficacité du modèle soit directement évaluée. Le deuxième niveau mesure l'apprentissage réalisé sur le simulateur, en termes d'attitude dans une situation donnée ou de compétences acquises pour réaliser un geste chirurgical. Le troisième niveau correspond à l'impact d'une formation de simulation sur le comportement lors du retour à une situation réelle clinique, ce qui est communément appelé le transfert. Enfin, le 4^{ème} niveau mesure l'impact du programme de simulation sur l'amélioration éventuelle de la santé des patient(e)s (Fig C).

Figure C : les 4 niveaux de Kirkpatrick ²³



3.2.2 Impact dans le domaine de la santé

Cook et al. ont publié en 2011 une méta-analyse au sujet de l'impact de la simulation pour former des professionnels de santé, comparativement à l'absence d'intervention ou en complément des pratiques habituelles.²⁴ Les critères d'inclusion choisis étaient très larges : les modèles étaient des mannequins synthétiques de base ou haute fidélité, des tissus réels humains ou d'animaux et des simulateurs de RV ; les professionnels étaient de tous niveaux de formation et incluaient chirurgiens, médecins, dentistes, vétérinaires et infirmières ; les méthodologies étaient diverses : études avant-après, études comparatives randomisées et non randomisées. Les résultats étaient analysés selon les niveaux de Kirkpatrick en termes d'apprentissage et d'éventuel effet mesurable sur les résultats des patients. 609 études ont été incluses. Les compétences acquises étaient le plus souvent évaluées sur le modèle de simulation ayant servi à l'entraînement et 32 études seulement permettaient de mesurer un effet sur des patients (niveau 4 de Kirkpatrick). Le nombre médian de participants était de 24 (interquartile range : 15-47). Les résultats montraient que la simulation est associée à de meilleurs résultats, avec un impact positif en terme d'apprentissage et un impact modéré mais

positif concernant l'effet sur les patients (pooled effect size : 0.50 (95% CI, 0.34-0.66; P<0.001). L'association entre simulation et résultats était hétérogène, étant donné le large spectre des critères d'inclusion. Cependant la simulation était presque toujours associée à un bénéfice (excepté dans 4% des études), ce qui conduit les auteurs à conclure que ce travail démontre le bénéfice de la simulation comme outil de formation dans le domaine de la santé, comparé à l'absence d'intervention. La question qui prévaut désormais est donc d'optimiser la façon d'utiliser « l'outil simulation » pour la formation des professionnels de santé, et ceci au meilleur coût.

3.2.3 *Impact en chirurgie mini-invasive*

Une méta-analyse publiée par la même équipe rassemblant 151 études a confirmé l'impact positif de la simulation pour l'entraînement à la coelioscopie, comparée à l'absence d'intervention.²⁵ Il est important de souligner que les exercices évalués étaient principalement des tâches simples (n=173) ou des exercices de suture (n=86) et plus rarement des procédures entières en chirurgie digestive (n=35), gynécologie (n=12) ou urologie (n=12). Par ailleurs, la mesure de l'efficacité de l'intervention était presque exclusivement selon les niveau 2 et 3 de Kirkpatrick, et donc l'impact sur la santé des patients n'était pas évalué.

3.3 Évaluation des compétences techniques chirurgicales

La chirurgie coelioscopique nécessite un apprentissage psychomoteur spécifique par rapport à la chirurgie ouverte, puisque les instruments sont mobilisés à partir du point fixe des trocars pariétaux. L'évaluation du geste opératoire est bien sûr essentielle, mais la compétence technique dépend aussi d'autres paramètres. La technique opératoire peut en effet comporter des variantes selon les écoles (ex : la réalisation ou non d'une fenestration du

ligament large lors d'une HC). D'autre part la capacité de l'opérateur de s'adapter à l'anatomie et à obtenir une bonne exposition est importante ; enfin la chronologie des temps opératoires doit être respectée. Cependant une certaine flexibilité dans l'ordre des gestes est souvent sans conséquences (elle est même nécessaire en cas de présentation anatomique inhabituelle), et elle ne saurait être reprochée comme défaut technique. Les outils d'évaluation utilisés sont surtout descriptifs et parfois prennent en compte les conséquences du geste pour juger de sa technicité.

3.3.1 *Les modes d'évaluation*

3.3.1.1 L'évaluation quantitative

- Le temps opératoire

Le temps opératoire est souvent utilisé pour évaluer les compétences techniques.²¹ C'est le critère principal dans la littérature pour l'évaluation des courbes d'apprentissage en coelioscopie.²⁶ Si la notion d'un temps opératoire diminué est intuitivement associée à la compétence des opérateurs expérimentés, ce critère ne permet pas d'évaluation directe de la qualité.

- Les paramètres métriques

Le nombre de mouvements réalisés par les mains du chirurgien, ou bien la distance parcourue par les instruments sont obtenus en laboratoire par un dispositif de capteurs placé sur les mains de l'opérateur, comme par exemple le procédé ICSAD dont l'utilisation est bien documentée.²⁷ Les paramètres métriques sont automatiquement enregistrés sur les simulateurs de RV, ce qui facilite le recueil de ces informations.²⁸ Ces paramètres sont beaucoup plus complexes à recueillir au BO.

3.3.1.2 Évaluation qualitative

- L'évaluation subjective

C'est la méthode d'évaluation employée dans le cadre du compagnonnage. Elle est la plus utilisée et elle reste incontournable dans la pratique courante.²⁹ Par nature subjective, cette évaluation est peu reproductible.³⁰

- Les outils binaires :

Un score binaire global « satisfaisant / non satisfaisant » est équivalent au mode d'évaluation subjectif du compagnonnage, avec les mêmes avantages et inconvénients.³¹ Les checklists sont des outils binaires appliqués à un geste chirurgical décomposé en séquences selon un ordre prédéfini. Chaque séquence est évaluée par un observateur qui juge si elle a été correctement réalisée ou non. Des checklists ont été développées pour des gestes fondamentaux comme la suture intra-corporéale³² et pour des interventions complètes.³³ Weizman et al. ont récemment publié une checklist pour évaluer la suture du dôme vaginal sur un pelvi-trainer.³⁴

- Les échelles d'évaluation

Les échelles d'évaluation regroupent plusieurs critères jugeant de la qualité technique d'une opération. Un observateur qualifié évalue chaque critère sur une échelle de Likert (le plus souvent de 1 à 5). La note attribuée pour chaque critère est guidée par une description de la manière dont est réalisé le geste. S'il reste subjectif, ce procédé de notation est plus structuré afin d'être précis et reproductible.

- Échelles globales

Les échelles globales évaluent les compétences techniques sans tenir compte des spécificités d'un geste opératoire donné. Elles sont donc utilisables quelle que soit

l'intervention. Le modèle de référence est l'échelle OSATS développée par l'équipe de Reznick à Toronto, qui comporte 7 critères (Fig D).³⁵ Initialement testée lors de séances opératoires réalisées par des internes en chirurgie ouverte sur modèle porcin vivant, son utilisation a depuis été validée dans d'autres circonstances opératoires, y compris au BO pour la chirurgie mini-invasive.³⁶ L'utilisation possible de l'échelle OSATS via une analyse de vidéos opératoires a également été démontrée.³⁷

Figure D : Objective structured assessment of technical skill (OSATS) Global Rating Scale, d'après Martin et al.³⁵

| | 1 | 2 | 3 | 4 | 5 |
|--|--|---|--|---|--|
| Respect for tissue | Frequently used unnecessary force on tissue or caused damage by inappropriate use of instruments | | Careful handling of tissue but occasionally caused inadvertent damage | | Consistently handled tissues appropriately with minimal damage |
| Time and motion | Many unnecessary moves | | Efficient time/motion but some unnecessary moves | | Economy of movement and maximum efficiency |
| Instrument handling | Repeatedly makes tentative or awkward moves with instruments | | Competent use of instruments although occasionally appeared stiff or awkward | | Fluid moves with instruments and no awkwardness |
| Knowledge of instruments | Frequently asked for the wrong instrument or used an inappropriate instrument | | Knew the names of most instruments and used appropriate instrument or the task | | Obviously familiar with the instruments required and their names |
| Use of assistants | Consistently placed assistants poorly or failed to use assistants | | Good use of assistants most of the time | | Strategically used assistant to the best advantage at all times |
| Flow of operation and forward planning | Frequently stopped operating or needed to discuss next move | | Demonstrated ability for forward planning with steady progression of operative procedure | | Obviously planned course of operation with effortless flow from one move to the next |
| Knowledge of specific procedure | Deficient knowledge. Needed specific instruction at most operative steps | | Knew all important aspects of the operation | | Demonstrated familiarity with all aspects of the operation |

L'échelle GOALS est un autre modèle d'échelle globale, développée par l'équipe de Fried pour la chirurgie coelioscopique au BO.³⁸ Les checklists sont plus rigides que les échelles d'évaluation, lesquelles permettent de mieux discriminer le niveau technique des internes.³⁹ Cependant ces échelles d'évaluation ont surtout montré leur efficacité pour comparer des internes de niveaux différents ou bien des internes à des chirurgiens séniors qualifiés.⁴⁰ À de rares exceptions près^{41,42}, ces outils n'ont pas prouvé leur efficacité pour évaluer

comparativement des chirurgiens qualifiés et doivent être utilisés avec prudence dans le cadre de la certification et de la post-certification.

- Échelles spécifiques d'une intervention chirurgicale

Les échelles spécifiques suivent le même principe que les échelles globales, mais les critères d'évaluation sont adaptés spécifiquement à l'intervention, et produits par une analyse séquentielle des différents temps opératoires. Les étapes clés à intégrer dans l'échelle peuvent être sélectionnées arbitrairement⁴³ ou être légitimées par un groupe d'experts grâce à la méthode Delphi.^{20,44}

Plusieurs échelles spécifiques d'interventions gynécologiques coelioscopiques ont été publiées. La première échelle est dédiée à la salpingectomie.⁴³ Deux échelles spécifiques de l'HC ont été publiées : la première incluait une liste de gestes opératoires définis par un consensus d'experts internationaux mais sa validité n'a pas été testée.⁴⁵ La deuxième échelle concernait l'HC robot-assistée. Ses étapes étaient définies d'après l'avis de 5 experts et elle incluait un processus de validation au BO. Les différences de scores étaient significatives entre trois groupes d'opérateurs de niveaux différents (experts, internes expérimentés et internes inexpérimentés : 75.6 vs 71.3 vs 69 sur un score maximum possible de 80), Les faibles écarts étaient expliqués par les auteurs en raison de la supervision des internes.⁴⁶

- Échelles combinées :

Les échelles combinées associent les critères d'une échelle globale de type OSATS à ceux d'une checklist ou d'une échelle spécifique.⁴⁰

3.3.1.3 Détection des erreurs

La compétence technique peut être évaluée par la mesure des erreurs opératoires. Un score général d'erreurs pour la coelioscopie a été développé par Bonrath et al. Celui-ci comporte 9 compétences (dont l'utilisation des énergies, la dissection et la suture) évaluées sous la forme du type et du nombre d'erreurs commises (4 catégories d'erreurs sont répertoriées : trop de force ou de distance / pas assez de force ou de distance / mauvaise orientation / mauvaise visualisation).⁴⁷ Cet outil a permis de discriminer des groupes d'opérateurs de niveau différent (score OSATS > ou < 28) réalisant une hystérectomie coelioscopique.⁴⁸ Un score d'erreurs spécifique d'une intervention, incluant une description des erreurs possibles pour chaque étape a été publié pour la cholécystectomie.⁴⁹

3.3.1.4 Les conséquences opératoires

La compétence technique peut être évaluée indirectement par la mesure des conséquences du geste : complications per et post opératoires ; suites opératoires ; résultats attendus du geste (marge de résection et taux de survie sans récurrence en oncologie, taux de récurrence pour la chirurgie du prolapsus).⁵⁰ L'étude de référence, réalisée dans le domaine de la coelioscopie bariatrique, établit un lien entre la mesure objective de la qualité technique opératoire et les résultats des patients.⁴² Cette méthode d'évaluation est cependant limitée par de nombreux facteurs confondants : indications, comorbidités des patientes, qualité de la prise en charge des autres intervenants au bloc et dans les suites. En outre, elle ne permet pas de donner un feedback et son intérêt est limité dans le cadre de la formation.

3.3.2 La compétence technique

Le terme de compétence technique est ambigu dans la littérature, car souvent désigné par le terme *proficiency*, qui se réfère à un niveau technique défini par des performances d'experts et donc théoriquement très élevé. Or la définition de la compétence technique correspond plus à la capacité de réaliser un geste ou une intervention complète avec un niveau minimum, de manière indépendante, en respectant des règles de sécurité.³⁶ La définition de ces pré-requis minimum de compétences est mieux adaptée à l'évaluation sommative.^{51,52}

Selon une revue de Szasz et al, la compétence technique est évaluée dans la littérature en utilisant 5 types de stratégie: les scores issus d'échelles d'évaluation, les critères de référence préalablement établis (*benchmark*) par des performances d'experts ou d'opérateurs compétents de niveau moins élevé, les outils binaires, les conséquences opératoires ou bien les nouveaux outils.³⁶ Dans ces nouveaux outils, les auteurs classent les scores d'erreurs, mais également d'autres échelles évaluant le niveau d'indépendance de l'interne lorsqu'il réalise un geste dans le cadre du compagnonnage.^{53,54}

3.4 La validité en sciences de l'éducation

L'évaluation de la compétence technique des élèves chirurgiens est essentielle pour juger de leurs progrès au cours de la formation et in fine garantir leur qualification. Cette évaluation intervient à plusieurs niveaux : au BO bien entendu, mais aussi lors d'exercices de simulation. La confiance que l'on peut accorder aux méthodes utilisées pour évaluer les opérateurs dépend de leur validité, qui est donc une question fondamentale en sciences de l'éducation.

Selon la définition traditionnelle, un test psychomoteur est valide s'il mesure effectivement ce qu'il prétend mesurer. Selon cette approche, mentionnée jusqu'en 1974 par l'American Educational Research Association, c'est le test qui est considéré comme plus ou moins valide, et cette notion de **validité** est scindée en différents types distincts :

- Apparente (*face validity*) : détermine si le test ressemble à une situation réelle, c'est-à-dire s'il apparaît subjectivement comme approprié.
- Contenu (*content validity*) : détermine si le domaine de compétence concerné est effectivement mesuré par le test. Il s'agit d'un jugement subjectif, réalisé par des experts du domaine étudié.
- Construction (*construct validity*) : est-ce que le résultat obtenu par le test est fidèle à la performance réelle. La validité de construction s'appuie souvent sur des hypothèses concernant les résultats attendus dans différents cas de figure : par exemple la capacité à discriminer des opérateurs de niveau différent.
- Critère (*criterion validity*) : deux formes sont reconnues :
 - Convergence (*concurrent validity*) : liens de corrélation entre les résultats obtenus avec un test et ceux obtenus avec un autre test reconnu comme celui de référence mesurant la même chose.
 - Predictive (*predictive validity*) : valeur pronostique de la mesure présente sur une performance future (ex : l'évaluation d'un interne peut-elle prédire son niveau lorsqu'il sera sénior ?). Cette composante est très difficile à obtenir en chirurgie et soumise à de nombreux biais.

Selon cette approche de la validité d'un test, une deuxième notion y est toujours associée : la **fiabilité** (*reliability*), exprimée par un coefficient de corrélation et scindée en trois types :

- Test-retest (*test-retest reliability*) : la répétition d'une même mesure donne-t-elle les mêmes résultats pour une performance donnée ? C'est notamment important lorsque le test implique une évaluation subjective par un examinateur, comme dans les échelles du type OSATS.
- Inter-correcteurs (*inter rater reliability*) : détermine s'il y a un accord entre plusieurs observateurs qui notent de manière indépendante les mêmes phénomènes ou les mêmes sujets à l'aide du même instrument.
- Homogénéité (*internal consistency*) : Il s'agit de savoir si un test mesure bien la même dimension au travers de ses différentes parties ou items.

Le concept de validité a évolué en sciences de l'éducation et une définition différente a été adoptée par les sociétés américaines concernées depuis 1999⁵⁵, laquelle reprend les principes énoncés par Messick:⁵⁶ la validité consiste en un jugement basé sur des preuves empiriques et sur une argumentation de nature théorique qui vise à justifier l'interprétation des scores obtenus suite à l'administration d'un test dans un contexte donné (« *Validity refers to the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests* »). D'après cette définition, la validité est vue comme une entité unique correspondant à la validité de construction. Elle n'est pas une valeur fixe pour un test donné, mais correspond à l'ensemble des preuves qui confirment (ou infirment) l'interprétation du résultat d'un test, dans un contexte donné et pour une population donnée.

Si la validité est une entité unique, les sources prouvant cette validité peuvent avoir plusieurs facettes et sont classées en cinq catégories : le contenu (*content*), la manière de

répondre à la question (*response process*), la structure interne, les relations avec les autres variables et les conséquences.⁵⁷ Il y a une confusion dans la littérature concernant l'emploi du mot « validité », puisque beaucoup de termes utilisés se réfèrent à l'ancienne dénomination et en 2010, selon une revue de Korndorfer et al., un quart seulement des publications utilisant des simulateurs de chirurgie faisait référence à la validité dans les termes recommandés par les standards actuels.⁵⁸

Il n'est pas obligatoire et souvent difficile de réunir toutes les facettes de la validité au cours d'une même étude. Les facettes de validité le plus souvent explorées dans le domaine de l'évaluation de la technique chirurgicale sont le contenu (idéalement par des avis d'experts selon la méthode Delphi), la structure interne (par la mesure de la fiabilité des mesures réalisées) et les relations avec les autres variables.⁵⁹ Cette dernière facette est habituellement explorée en comparant des groupes d'opérateurs de niveaux différents et en faisant l'hypothèse que les opérateurs les plus qualifiés (« les experts ») auront de meilleurs scores que les moins qualifiés. Lorsqu'un test ne permet pas de discriminer deux groupes de niveaux différents prédéterminés, sa validité est sérieusement remise en question. Cependant la mesure d'une différence entre les deux groupes n'est pas suffisante à elle seule pour s'assurer de la validité d'un test. Les différences observées peuvent être expliquées par des facteurs confondants, comme l'habitude de tenir les pinces de coelioscopie lorsqu'on évalue la réalisation d'une intervention donnée. Une autre limite est liée aux mesures qui sont réalisées à l'échelle de la moyenne d'un groupe, sans tenir compte des différences à l'intérieur de ce groupe. En outre, le choix de groupes de niveaux très différents (ex : experts versus débutants) augmente les chances de montrer une différence, et n'est pas toujours en lien avec l'utilisation du test qui sera proposé en pratique à des internes tout au long de leur courbe

d'apprentissage.⁶⁰ L'inclusion d'un groupe de niveau intermédiaire lors de l'évaluation d'un test est un moyen de limiter ce biais.

3.5 La simulation en cœlioscopie gynécologique

La cœlioscopie a été expérimentée en chirurgie gynécologique dès les années 80.⁶¹ Les spécificités de cette nouvelle voie d'abord ont contribué au développement de simulateurs pour faciliter l'apprentissage. Les gynécologues ont été les premiers à opérer en cœlioscopie mais les programmes de simulations ont été développés par les chirurgiens digestifs, notamment suite à la constatation de l'augmentation des complications survenant lorsque des opérateurs qualifiés en chirurgie ouverte réalisaient des cholécystectomies sans avoir été formés à la cœlioscopie.⁶² L'initiative ayant eu le plus d'échos est le développement du pelvi-trainer MISTEL par l'équipe de Fried ²¹, à la base du programme d'entraînement Fundamental of Laparoscopic Surgery (FLS). Les exercices de simulation décrits en gynécologie utilisent un éventail de modèles similaires aux autres spécialités chirurgicales.⁶³

3.5.1 Les modèles décrits

3.5.1.1 Les exercices de base non spécifiques

- Les pelvi-trainers :

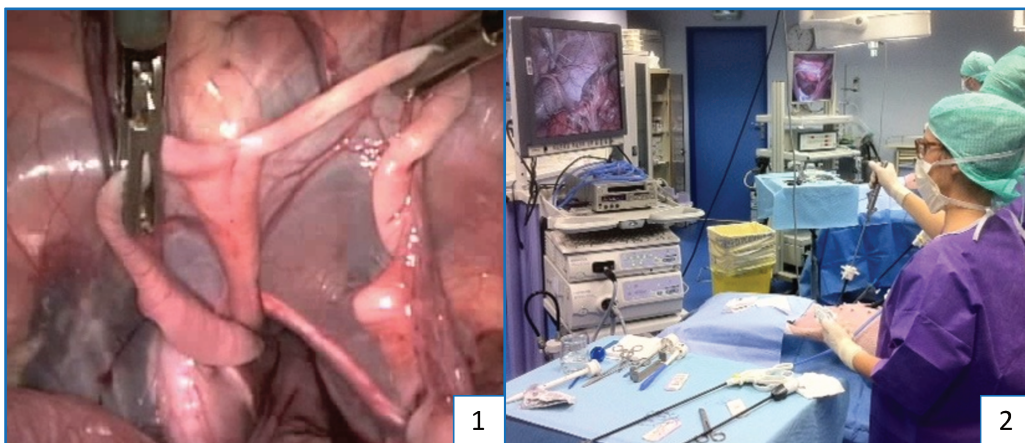
Le pelvi-trainer FLS est largement utilisé pour la formation des gynécologues.⁶⁴

D'autres modèles ont été développés, utilisant un pelvi-trainer de type Szabo, permettant de s'exercer à la coordination psycho-motrice (LASTT) et aux exercices de suture (SUTT).⁶⁵

- Les tissus réels :

Le modèle porcin vivant permet de réaliser des gestes non spécifiques de la gynécologie, mais très intéressants pour l'entraînement : suture, dissections, ligature d'un pédicule vasculaire lors de la néphrectomie (Fig E).⁶⁶

Figure E : Séance d'entraînement en laboratoire sur modèle porcin vivant : 1 vue opératoire de l'appareil génital ; 2 Laboratoire du centre d'enseignement et de recherche chirurgicale (Aix-Marseille Université).



- Les simulateurs de réalité virtuelle

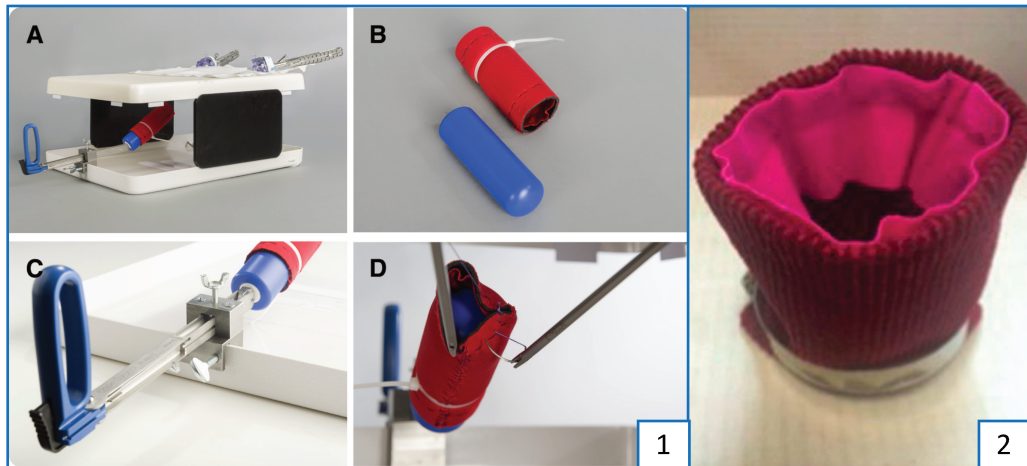
Plusieurs simulateurs commercialisés, comme le Lapsim*, incluent des exercices de bases qui ont été testés en gynécologie.^{67,68}

3.5.1.2 Les gestes spécifiques de la chirurgie gynécologique

- Les pelvi-trainers

Certains pelvi-trainers ont été adaptés pour l'entraînement à des interventions spécifiques, en y plaçant des tissus synthétiques : c'est le cas pour le temps de fermeture vaginale de l'HC (fig F)^{69,70}, et pour la sacrospinofixation.⁷¹

Figure F : Modèles de suture vaginale sur pelvi-trainer d'après : Tunitsky-Bitton et al. (1), King et al. (2)



- Les tissus réels :

Les interventions concernées sont surtout la chirurgie annexielle (salpingectomie) lors de l'entraînement sur cadavre ⁷², ou bien sur des modèles animaux vivants ovine ⁷³ ou porcine.⁷⁴ Une publication décrit la réalisation d'une hystérectomie par voie coelioscopique robot-assistée sur un modèle porcine vivant.⁷⁵ Les lymphadénectomies ont également été rapportées sur modèle porcine.⁷⁶ Le modèle animal, en particulier porcine, est peu utilisé pour des gestes spécifiques gynécologiques en raison de sa faible ressemblance avec l'appareil génital féminin (fig F1).

- Les simulateurs de RV :

Les deux simulateurs actuellement dominant sur le marché, le Lapsim* ou le Lapmentor*⁷⁷, proposent des programmes reproduisant des interventions gynécologiques complètes pour la chirurgie annexielle (ligature de trompe, salpingectomie, grossesse extra-utérine) ^{67,78-80} et l'hystérectomie. Les programmes concernant l'HC n'ont jamais été évalués.

3.5.2 *Les curricula*

3.5.2.1 Utilisation du FLS en gynécologie

Le curriculum FLS a été utilisé en gynécologie pour des internes dans un essai randomisé avec mesure de la qualité opératoire au bloc. L'intervention testée était simple (ligature de trompe) et la différence de qualité opératoire modérée mais significative par rapport au groupe contrôle.⁸¹

3.5.2.2 Le curriculum de Shore

Un curriculum a été construit pour les internes de gynécologie, en utilisant la méthode Delphi pour déterminer les étapes pédagogiques à intégrer et les outils de simulation à utiliser.⁸² Ce curriculum proposait plusieurs étapes dont une étape d'entraînement cognitif puis un entraînement réparti sur 7 semaines avec des plages horaires dédiées. Les simulateurs utilisés étaient le pelvi-trainer FLS ainsi qu'un simulateur de RV. Les exercices choisis étaient des exercices de base et de suture. Aucun exercice spécifique de la gynécologie n'était retenu. Le curriculum comprenait également un entraînement aux compétences non techniques. Ce curriculum a été évalué et a montré son efficacité par rapport à un groupe contrôle, lors du transfert au BO pour réaliser une suture intra-corporéale et une salpingectomie.⁸³

3.5.2.3 Le curriculum pour la salpingectomie :

Un curriculum d'apprentissage spécifique de la salpingectomie a été développé sur un simulateur de RV selon une méthodologie similaire au consensus de Zevin.²⁰ Un programme d'entraînement a d'abord été construit en utilisant les programmes de RV du Lapsim (exercices de base et salpingectomie pour la grossesse extra-utérine) ; les valeurs de référence étaient les performances obtenues par des experts selon des paramètres métriques

enregistrés automatiquement par le simulateur. Une échelle d'évaluation spécifique de la salpingectomie par voie coelioscopique a été construite et sa validité confirmée par évaluation anonyme de vidéos réalisées par des opérateurs de niveau différent.⁴³ La validité de ce curriculum a ensuite été prouvée par un essai randomisé multicentrique montrant que le groupe curriculum obtenait de meilleurs scores de qualité que le groupe compagnonnage, selon l'échelle spécifique.⁸⁴

4 TRAVAUX PRÉLIMINAIRES

4.1 Curriculum sur simulateur de RV pour la cholécystectomie

4.1.1 Présentation

Development of a virtual reality training curriculum for laparoscopic cholecystectomy.

Aggarwal R, Crochet P, Dias A, Misra A, Ziprin P, Darzi A. Publié dans le British Journal of Surgery. 2009 ;96(9) :1086-93.

L'objectif principal de ce travail était de construire, selon une méthode scientifique, un curriculum d'entraînement sur un simulateur de RV (LapMentor*) pour une intervention coelioscopique complète : la cholécystectomie.

- Les modules disponibles sur le simulateur ont été testés : 9 exercices de base non spécifiques, 4 exercices spécifiques d'un temps de la cholécystectomie (ex: dissection du triangle de Calot) et 1 programme reproduisant l'intervention entière.
- Les preuves de validité de chacun des exercices étaient explorées en comparant les performances de 3 groupes d'opérateurs d'expérience clinique différente (internes inexpérimentés, jeunes séniors de niveau intermédiaire, seniors expérimentés).
- Les critères d'évaluation retenus étaient les paramètres métriques automatiquement enregistrés par le simulateur.
- Une analyse de la courbe d'apprentissage était réalisée pour les opérateurs inexpérimentés en faisant 10 répétitions de chaque exercice.

Voici les principaux résultats de cette étude :

- La majorité des modules testés ont montré des preuves de validité (différence significative entre les 3 groupes).
- Les critères d'évaluation retenus étaient des critères purement métriques, comme le temps ou le nombre de mouvements.
- Une courbe d'apprentissage était observée pour tous les critères valides, avec un plateau apparaissant avant la dixième répétition (2nd - 9^{ème}).
- Le résultat est la proposition d'un parcours d'entraînement spécifique de la cholécystectomie, selon la théorie des curricula (exercices valides retenus, du plus simple au plus compliqué, en respectant le principe de séances distribuées dans le temps, jusqu'à atteindre les critères de qualité des opérateurs expérimentés).

Les questions posées par cette étude étaient :

- Ce curriculum est-il valide et applicable pour une plus large population d'internes ?
- Les compétences acquises sur le simulateur sont-elles utiles lors du transfert au BO ?
- Suite à la réalisation de ce curriculum en laboratoire, quel est le gain en termes de courbe d'apprentissage au BO par rapport au compagnonnage classique ?
- Est-ce que les critères métriques retenus sont corrélés à la qualité opératoire ?

Development of a virtual reality training curriculum for laparoscopic cholecystectomy

R. Aggarwal, P. Crochet, A. Dias, A. Misra, P. Ziprin and A. Darzi

Department of Biosurgery and Surgical Technology, St Mary's Campus, Imperial College Healthcare NHS Trust, 10th Floor, Queen Elizabeth the Queen Mother Building, St Mary's Hospital, Praed Street, London W2 1NY, UK

Correspondence to: Mr R. Aggarwal (e-mail: rajesh.aggarwal@imperial.ac.uk)

Background: Training within a proficiency-based virtual reality (VR) curriculum may reduce errors during real surgical procedures. This study used a scientific methodology for development of a VR training curriculum for laparoscopic cholecystectomy.

Methods: Inexperienced (had performed fewer than ten laparoscopic cholecystectomies), intermediate (20–50) and experienced (more than 100) surgeons were recruited. Construct validity was defined as the ability to differentiate between the three levels of experience, based on simulator-derived metrics for nine basic skills, four procedural tasks and full laparoscopic cholecystectomy on a high-fidelity VR simulator. Inexperienced subjects performed ten repetitions for learning curve analysis. Proficiency measures were based on the performance of experienced surgeons.

Results: Thirty inexperienced, 11 intermediate and 16 experienced operators were recruited. Eight of nine basic skills and three of four procedural tasks were found to be construct valid. The full procedure revealed significant intergroup differences for time (1541, 673 and 816 s; $P = 0.002$), movements (1021, 595 and 638; $P = 0.006$) and path length (2038, 1235 and 1303 cm; $P = 0.033$). Learning curves plateaued between the second and ninth sessions.

Conclusion: This study shows that it is possible to define and develop a whole-procedure VR training curriculum for laparoscopic cholecystectomy using structured scientific methodology.

Paper accepted 16 April 2009

Published online in Wiley InterScience (www.bjs.co.uk). DOI: 10.1002/bjs.6679

Introduction

Training in the operating theatre is often unstructured, and occurs by chance encounters dependent on patient and disease variability. A particular facet of surgical practice is the need to train inexperienced individuals to a level of competence in their chosen field. The term 'learning curve' has been used repeatedly to account for longer operating times, higher complication rates and even higher mortality rates during this period of skills acquisition^{1,2}. Although training is supervised, and in accordance with informed consent of the patient, this is probably no longer an ethically or economically viable option for modern medical practice. It is thus necessary to explore, define and implement modes of surgical skills training that do not expose the patient to preventable errors³.

Simulation in the form of virtual reality (VR) and synthetic models has been proposed for technical skills training at the early part of the learning curve⁴. A

recent systematic review has substantiated the transfer of simulation-based training to the operative setting, in the fields of laparoscopy and endoscopy, through provision of a 'safe, effective and ethical way for trainees to acquire surgical skills'⁵. Nonetheless, the authors stated that the included studies were of variable quality, using a number of different strategies for training of participants. It is considered preferable for training to be structured within a standardized curriculum⁶.

The aim of a training curriculum is for an individual to acquire skills to a predetermined level of proficiency before progression to more challenging cases. It is thus not solely the simulator, but also the mode of training on the simulator, that determines the degree of transference of skill to the operative setting. This constitutes knowledge-based learning, a stepwise technical skills pathway, ongoing feedback and progression toward proficiency goals, enabling transfer to the real environment⁷.

The aim of this study was to develop an evidence-based and stepwise VR training curriculum for acquisition of technical skills for laparoscopic cholecystectomy (LC). Although curricula for basic and procedural tasks on VR simulators have been defined previously^{8,9}, this is the first time that a whole-procedure VR simulation tool has been subjected to a structured scientific method for curriculum development. The challenge was to define a hierarchical model of training that is underpinned by objective measures of proficiency, with the eventual aim of demonstrating a significant reduction in length and gradient of the learning curve during LC on real patients.

Methods

The study recruited subjects, divided into experienced (who had performed more than 100 LCs), intermediate (20–50 LCs) and inexperienced (fewer than ten LCs) operators. To be included in the inexperienced group, individuals who had not performed any LCs at all had to have been the primary assistant for a minimum of five LC procedures. Recruitment was solely through personal communication. The only exclusion criterion was previous training experience with a laparoscopic simulator of any kind, including box trainers.

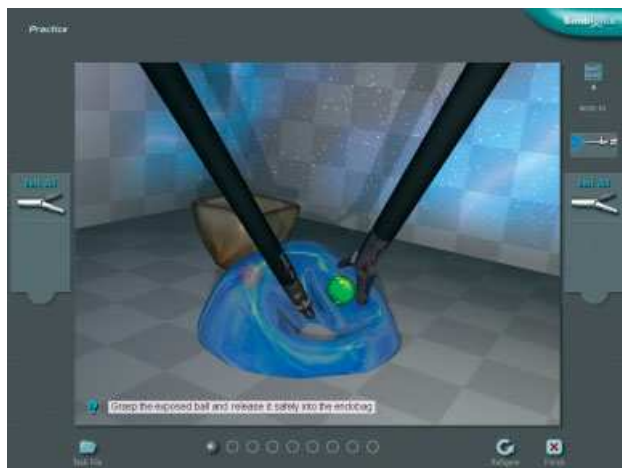
Virtual reality simulation tool

The LAP Mentor™ VR laparoscopic surgical simulator (Symbionix Corporation, Cleveland, Ohio, USA) is divided into laparoscopic basic skills, procedural tasks and full cholecystectomy procedures. There are nine basic skills

(such as two-handed manoeuvres; *Fig. 1a*), four procedural tasks (for example Calot's triangle dissection; *Fig. 1b*) and 18 full LC procedures. The full LC procedures vary in terms of biliary anatomy and degree of inflammation. For the purposes of this study, which aimed to develop a curriculum to train inexperienced laparoscopic surgeons, only the first LC was used with a patient history of biliary colic with standard anatomy. The subject interacts with the simulated tissues with two generic laparoscopic instruments through a haptic feedback device. Instrument type (grasper, scissors, clip applier, diathermy hook, etc.) may be selected through an on-screen menu-driven process. A detailed description of the simulator tasks is provided in *Table 1*.

Tasks performed

Each of the nine basic skills, four procedural tasks and one full procedure were performed for two sessions by experienced and intermediate subjects, and for ten sessions by inexperienced subjects. Before commencing the tasks, each subject was provided with a one-on-one simulator familiarization session, by an experienced operator. Before each task was performed, a full demonstration was provided by the experienced operator, with an opportunity for the subject to ask questions. The familiarization period lasted for approximately 30 min. No assistance was provided during hands-on training. All sessions were completed at least 1 h apart, and those in the inexperienced group completed no more than two sessions per day.



a Two-handed manoeuvres



b Calot's triangle dissection

Fig. 1 Screen shots of **a** two-handed manoeuvres basic skill and **b** Calot's triangle dissection procedural task

Table 1 Description of the modules on the LAP Mentor™ virtual reality simulator

| Skill or task | Description |
|--|--|
| Basic skills | |
| Camera manipulation 0° | Using a 0° camera, locate and snap photographs of ten balls, in an abstract environment |
| Camera manipulation 30° | Using a 30° angled camera, locate and snap photographs of ten balls, in an abstract environment |
| Eye–hand coordination | Locate each flashing ball and touch with the tool of the appropriate colour |
| Clip application | Clip leaking ducts within a specified segment, before the pool fills |
| Clipping and grasping | Safely grasp and clip leaking ducts within a specified segment, before the pool fills |
| Two-handed manoeuvres | Use two graspers to locate the balls within the jelly mass and then place them in the endobag |
| Cutting | Safely cut and separate a circular form using one tool to retract and the scissors to cut accurately |
| Electrocautery | Use a hook to burn the highlighted band, while retracting other bands with an accessory instrument |
| Translocation of objects | Manipulate object with two graspers and, with a minimum number of translocations, place the object into the orientation of the matching transparent object |
| Procedural tasks | |
| Clipping and cutting—retracted gallbladder | Gallbladder already exposed with Hartmann's pouch retracted by a static tool. Clip the cystic artery and duct within a specified area and then cut safely between the clips |
| Clipping and cutting—two hands | With the gallbladder already exposed use a blunt grasper to retract Hartmann's pouch. Once correct retraction is achieved, clip the cystic artery and duct within a specified area and then cut safely between the clips |
| Calot's triangle dissection | Grasp the infundibulum of the gallbladder, retract away from the liver, and dissect the peritoneal coverings to expose the cystic duct and artery |
| Gallbladder separation | Separate the gallbladder from the liver bed with appropriate retraction and dissection of the peritoneal adhesions to the liver bed. Continue dissection until the gallbladder is free from the liver |
| Full procedure | Virtual complete cholecystectomy procedures, based on anatomies created from CT/MRI real patient data. Practise a complete cholecystectomy procedure, with a range of appropriate instruments |

CT, computed tomography; MRI, magnetic resonance imaging.

Performance evaluation

Data for each of the performed tasks were measured instantly and objectively by the VR simulator. This comprised metrics ranging from time taken and economy of movement to accuracy, cautery time and error scores. The data were stored by the simulator software in a Microsoft® Excel™ spreadsheet (Microsoft Corporation, Redmond, Washington, USA).

Construct validity is a test of whether a model can differentiate between different levels of experience, and thus be used to assess performance⁷. Comparison of median performance at the first two sessions between the three groups of surgeons was used to assess whether each simulated task was construct valid and substantiated the use of the defined settings of the simulator to assess laparoscopic technical skill. Statistical analysis of the learning curves for novice laparoscopic surgeons was used to clarify whether repeated practice improved performance toward that of the experienced group. The definition of benchmark criteria to be achieved before progression to the next stage of the curriculum was by calculation of the median score for each parameter during the second

session for all experienced surgeons. Assessment of the experienced surgeons' second repetition was assumed to reduce the effect of their familiarization with the simulator during the first session.

The way novices advanced through these clearly defined steps through comparative measurement of simulator-derived metrics, that is construct validation, learning curve analysis and benchmark definition, enabled the assembly of a curriculum for procedural training, based on data rather than supposition. This provided an evidence- and proficiency-based pathway for novice surgeons to follow.

Statistical analysis

The choice of at least ten subjects per group was based on a two-tailed test, with $\alpha = 0.05$ and power $(1 - \beta) = 0.80$, and an intended reduction of 30 per cent in time taken to complete tasks for experienced *versus* inexperienced operators, based on data from previous studies of VR simulation^{8,9}. This yielded a value of eight subjects per group, which was increased to ten to allow for dropout and technical malfunction of the simulator. Thirty inexperienced individuals were recruited to ensure that ten

Table 2 Construct valid metrics for basic skills

| Task | Time taken (s) | Accuracy rate (%) | Total speed (cm/s) | Total no. of movements | Total path length (cm) |
|--------------------------|----------------|-------------------|--------------------|------------------------|------------------------|
| Camera manipulation 0° | ✓ | ✓ | | | |
| Camera manipulation 30° | ✓ | | | | |
| Eye–hand coordination | | | | | |
| Clip application | | | ✓ | | |
| Clipping and grasping | ✓ | | ✓ | | |
| Two-handed manoeuvres | ✓ | | ✓ | ✓ | ✓ |
| Cutting | ✓ | | | | |
| Electrocautery | ✓ | | | | |
| Translocation of objects | | | ✓ | | |

subjects performed ten repetitions of the nine basic skills, a further ten subjects trained solely on the four procedural tasks, and a final ten practised on the full LC procedure. The intention was to make sure that inexperienced subjects were not trained for ten sessions on one VR module and then performed another module, as this would have influenced the results of the study inappropriately. This was not an issue for the intermediate and experienced subjects as they performed only two repetitions of each task.

The data were analysed with SPSS® version 16.0 (SPSS, Chicago, Illinois, USA) using non-parametric tests. Comparison of performance between experienced, intermediate and inexperienced groups was undertaken using the Kruskal–Wallis test and Mann–Whitney *U* test, as appropriate. Learning curve data were analysed by means of the Friedman test (non-parametric repeated-measures ANOVA). Multiple comparisons were then made to identify when skills had plateaued. $P < 0.050$ was considered statistically significant.

Results

Fifty-seven subjects, comprising 16 experienced, 11 intermediate and 30 inexperienced operators, were recruited. Of the experienced surgeons, nine completed two sessions on the nine basic skills, 11 on the four procedural tasks and ten on the full LC procedure. Similarly, of the intermediate subjects, six completed two sessions on the nine basic skills, seven on the four procedural tasks and seven on the full LC procedure. The 30 inexperienced subjects were divided into three equal groups, the first of which performed ten repetitions of nine basic skills, the second four procedural tasks, and the final group the full LC procedure.

Basic skills

Eight of the nine basic skills demonstrated construct validity, based primarily on time taken and speed of

completion (Table 2). ‘Clipping and grasping’ exhibited significant differences in time taken between inexperienced, intermediate and experienced subjects (161, 120 and 111 s respectively; $P = 0.009$) and total speed (5.6, 7.4 and 7.9 cm/s; $P = 0.011$). The ‘two-handed manoeuvres’ skill revealed group differences for time taken (171, 107 and 102 s; $P = 0.001$) and total speed (6.3, 6.7 and 7.3 cm/s; $P = 0.049$), as well as total number of movements (239, 128 and 128; $P = 0.007$) (Fig. 2) and total path length (673, 484 and 512 cm; $P = 0.016$) of instrument tips. There were also significant differences in these metrics for the two skills between inexperienced and intermediate, and inexperienced and experienced groups, but not between intermediate and experienced groups.

Analysis of learning curves for the inexperienced group revealed significant levels of improvement for ‘clipping and grasping’ and ‘two-handed manoeuvres’ for the aforementioned valid metrics, apart from total speed for the latter skill. All metrics were shown to plateau at between the sixth and ninth repetition for this group, but total

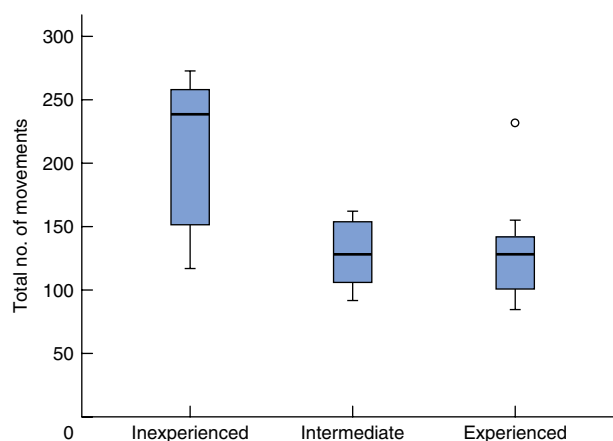


Fig. 2 Total number of movements for two-handed manoeuvres skill. Horizontal lines within boxes, boxes and whiskers represent median, interquartile range and range respectively. Circle represents an outlier ($P = 0.007$, Kruskal–Wallis test)

speed on 'clipping and grasping' plateaued at the second repetition.

Procedural tasks

'Clipping and cutting – retracted gallbladder' did not differ between the three groups based on the metrics recorded by the simulator. The 'clipping and cutting – two hands' task revealed a significant difference only for total speed (3.5, 4.1 and 4.7 cm/s for inexperienced, intermediate and experienced subjects; $P = 0.049$). For 'Calot's triangle dissection', time taken ($P < 0.001$), total number of movements ($P = 0.002$), total path length ($P = 0.011$), total speed ($P = 0.001$), accuracy rate ($P = 0.002$), total cautery time ($P = 0.003$) and total cautery time without tissue contact ($P < 0.001$) were significantly different between inexperienced, intermediate and experienced surgeons over the first session. 'Gallbladder separation' showed significant differences for time taken ($P < 0.001$), total number of movements ($P < 0.001$), total path length ($P = 0.001$) and time taken to extract the gallbladder ($P < 0.001$).

Learning curves for the ten inexperienced subjects were statistically significant ($P < 0.001$) for time taken, total number of movements, total cautery time and total cautery time without tissue contact for 'Calot's triangle dissection', and plateaued at the seventh repetition. For 'gallbladder separation', all construct valid metrics demonstrated significant learning curves for the inexperienced subjects ($P < 0.001$), with a plateau at the fourth session (Fig. 3).

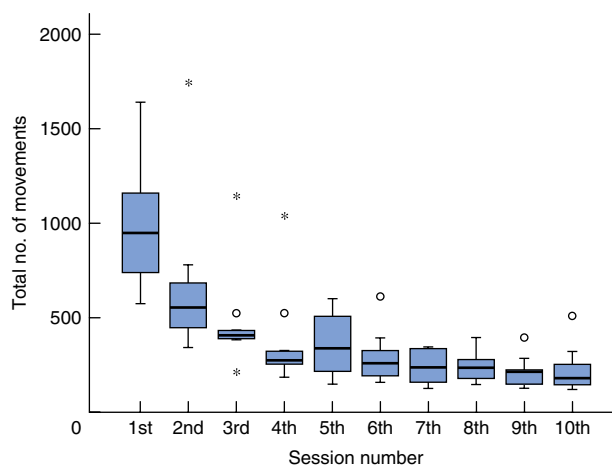


Fig. 3 Learning curve for total number of movements for gallbladder separation. Horizontal lines within boxes, boxes and whiskers represent median, interquartile range and range respectively. Circles and asterisks represent outliers and extreme cases respectively

Full procedure

The full procedure demonstrated significant differences between the three groups for time taken (1541, 673 and 816 s for inexperienced, intermediate and experienced subjects; $P = 0.002$), time to extract the gallbladder (1487, 635 and 768 s; $P = 0.002$), total number of movements (1021, 595 and 638; $P = 0.006$) and total path length (2038, 1235 and 1303 cm; $P = 0.033$). All construct valid

Table 3 Metrics for development of the training curriculum

| Task | Metric | Construct valid | Learning curve | Plateau session | Benchmark level |
|-----------------------------|---------------------------------|-----------------|----------------|-----------------|-----------------|
| Skills | | | | | |
| Clipping and grasping | Time taken (s) | ✓ | ✓ | 9th | 104 |
| | Total speed (cm/s) | ✓ | ✓ | 2nd | 8.3* |
| Two-handed manoeuvres | Time taken (s) | ✓ | ✓ | 6th | 89 |
| | Total no. of movements | ✓ | ✓ | 8th | 106 |
| | Total path length (cm) | ✓ | ✓ | 8th | 440 |
| | | | | | |
| Tasks | | | | | |
| Calot's triangle dissection | Time taken (s) | ✓ | ✓ | 7th | 278 |
| | Total no. of movements | ✓ | ✓ | 7th | 241 |
| | Total cautery time (s) | ✓ | ✓ | 7th | 15 |
| Gallbladder separation | Time taken (s) | ✓ | ✓ | 4th | 312 |
| | Time to extract gallbladder (s) | ✓ | ✓ | 4th | 219* |
| | Total no. of movements | ✓ | ✓ | 4th | 274 |
| | Total path length (cm) | ✓ | ✓ | 4th | 511 |
| Full procedure | Time taken (s) | ✓ | ✓ | 2nd | 548 |
| | Time to extract gallbladder (s) | ✓ | ✓ | 2nd | 521* |
| | Total no. of movements | ✓ | ✓ | 3rd | 481 |
| | Total path length (cm) | ✓ | ✓ | 3rd | 1012 |

*Metrics not included in the curriculum (Fig. 4).

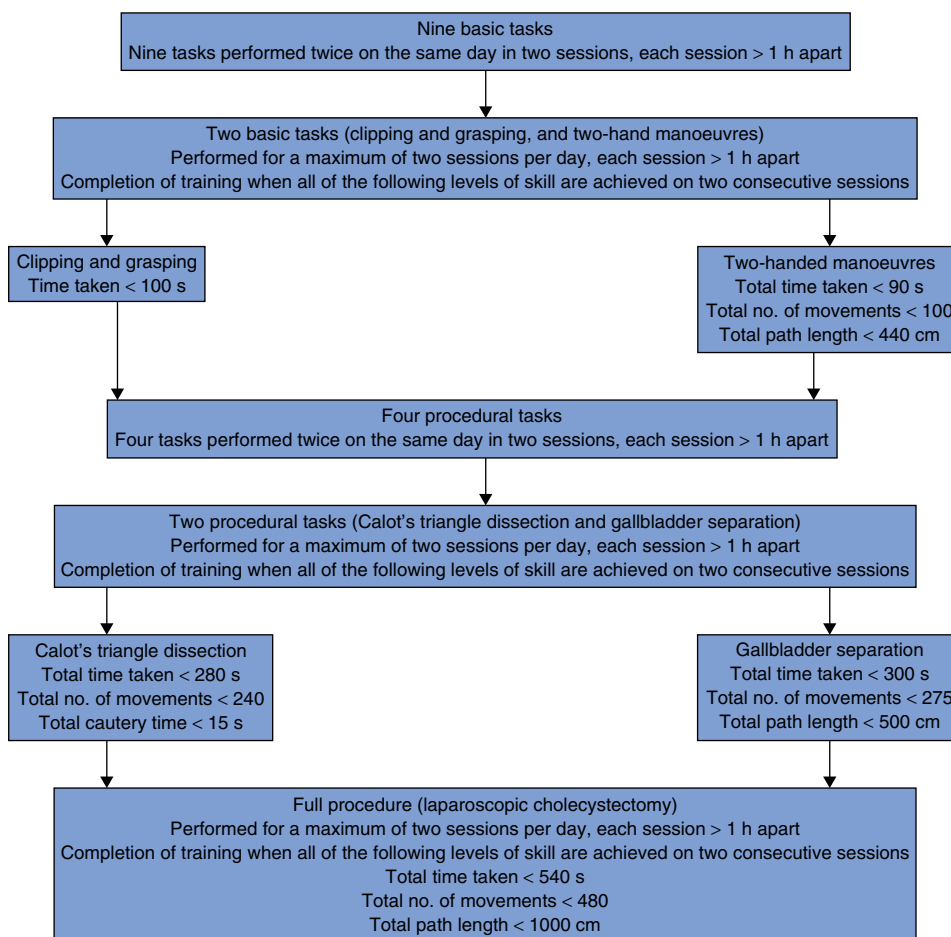


Fig. 4 Evidence-based whole-procedure virtual reality training curriculum for laparoscopic cholecystectomy

metrics again displayed significant learning curves, with a plateau at the second to third repetition ($P < 0.001$).

Curriculum construction

The results summarized in *Table 3* were used in the development of a training curriculum. Two metrics were excluded: total speed for ‘clipping and grasping’ as it was felt that training to speed was inappropriate, and time to extract gallbladder for the full procedure as this was almost identical to time taken and did not add to the measurement of proficiency. The summarized outcome is a proficiency-based VR whole-procedure curriculum for training in laparoscopic cholecystectomy (*Fig. 4*).

Discussion

This study applied a stepwise process to the modules and metrics of a VR simulator, resulting in the development of a whole-procedure training curriculum for LC. The

modules were deemed construct valid through comparison of performance across three levels of surgical experience. Furthermore, learning curve data were established to ensure that repetitive practice indeed improved performance in novices, as measured by the simulator. The technical skills taught by training on the simulator are thus relevant for LC, and should lead to a reduction in the time taken to achieve proficiency in real patients.

Training within the curriculum commences at the basic skills modules, with two repetitions of all nine skills. This is followed by training at the two most challenging skills, evidenced by the fact that they have significant learning curves. Progression to the procedural tasks necessitates achievement of the benchmark proficiency criteria, which are based on scores derived from the performance of experienced laparoscopic surgeons. The structure of the curriculum is identical for the four procedural tasks, leading to the full-procedure module, which again has proficiency criteria for the trainee to achieve before completion of the training period. It is also important to note that the

curriculum adheres to the concept of 'distributed' rather than 'massed' training schedules, with a maximum of two sessions performed per day, each at least 1 h apart^{10,11}. Finally, to confirm acquisition of skill rather than attainment of a good score by chance, all benchmark levels must be achieved at two consecutive sessions.

The data from this study have clearly determined the construct validity and learning curve progression of the simulator-based modules. Interestingly, there was no difference between the performance of intermediate and experienced groups on the simulator. This is an entirely appropriate finding, as those in the intermediate group are approaching the plateau phase of their learning curve for LC. Inexperienced subjects are thus most likely to benefit from training with this curriculum. A further point to note is the increase in consistency of scores both from increasing experience levels and with repeated training of inexperienced surgeons.

Other studies have investigated the construct validity of the basic skills on the LAP Mentor™ VR simulator. McDougall and colleagues¹² analysed performance across four different levels of experience with respect to a summed score for each skill¹². Although the outcomes in terms of construct validity for the basic skills were positive, it is uncertain whether the arbitrary nature of summed scores (as provided by the manufacturer of the simulator) is appropriate for skills measurement. It is necessary to develop the evidence for relative weights of the different components that make up these scores. However, Zhang and co-workers¹³ have provided good evidence for construct validity of the basic skills on the basis of time and overall scores, in concordance with the findings of McDougall *et al.*¹².

Several previous studies have assessed the validity and learning curves of other VR simulator devices^{14,15}. Although a fundamental for use of simulation in clinical training schedules, the organization of such data into a coherent, stepwise and proficiency-based training curriculum has not yet been pursued. The systematic review by Sturm and colleagues⁵ stated that a number of different training strategies was used for VR-based training before transfer to the operative environment. Most studies trained subjects based on a prescribed number of repetitions; those that trained to proficiency used isolated modules based on the performance measures of experienced surgeons^{16,17}. For simulation-based training to be disseminated across surgical residency programmes, it is necessary to provide 'instructions for use' of the simulator, that is a prescribed training curriculum. This will enable a standardized introduction to laparoscopic surgery, which can also ensure that a novice surgeon cannot

operate on a surgical patient until prescribed proficiency criteria have been achieved.

A common complaint is the expense in terms of simulator cost and upkeep, training space and faculty time required for integration of VR curricula into residency programmes¹⁸. With reductions in the learning curve during real operations, it is possible that the total cost of training each surgical resident will be reduced. In terms of training schedules, this curriculum prescribes two sessions per day, at least 1 h apart. The evidence for distributed training schedules is clear, although it is uncertain whether this means practice once a day or once a week^{10,11}. Flexibility in accommodating training sessions will be needed when implementing this curriculum, but this should not detract from acquisition of skill as curriculum completion is based on achievement of proficiency measures.

This training programme is not intended as a substitute for skills acquisition in the operating theatre, but it will allow part of the learning curve to be transferred to the skills laboratory⁴. The curriculum has not taken into account previous procedural or technical knowledge, nor objectively measured this before enlisting trainees into technical practice. A cognitive skills module is essential at the front end of any training programme, such as that available from the Royal College of Surgeons of England for LC¹⁹. Furthermore, completion of this curriculum is based on dexterity, rather than safety scores or clinical outcome measurements. This is an important aspect of research with respect to the use of technical skills rating scales and, furthermore, the integration of such scales into the simulator software²⁰.

It is crucial to disseminate this curriculum to other users of VR simulation, to enable external validation of the curriculum in terms of ease of use and feasibility, and to determine ultimately whether its use does actually lead to the notion of the pretrained novice who can operate with greater dexterity and skill on patients undergoing laparoscopic surgical procedures. It is then only a matter of time until other domains of surgical practice have to follow this lead of simulation-based training, with objective measurement of performance before operative intervention.

Acknowledgements

The authors declare no conflict of interest.

References

- 1 Reichenbach DJ, Tackett AD, Harris J, Camach D, Gravis EA, Dewan B *et al.* Laparoscopic colon resection early

- in the learning curve: what is the appropriate setting? *Ann Surg* 2006; **243**: 730–735.
- 2 Watson DI, Baigrie RJ, Jamieson GG. A learning curve for laparoscopic fundoplication. Definable, avoidable, or a waste of time? *Ann Surg* 1996; **224**: 198–203.
 - 3 Aggarwal R, Darzi A. Technical-skills training in the 21st century. *N Engl J Med* 2006; **355**: 2695–2696.
 - 4 Aggarwal R, Ward J, Balasundaram I, Sains P, Athanasiou T, Darzi A. Proving the effectiveness of virtual reality simulation for training laparoscopic surgery. *Ann Surg* 2007; **246**: 771–779.
 - 5 Sturm LP, Windsor JA, Cosman PH, Cregan P, Hewett PJ, Maddern GJ. A systematic review of skills transfer after surgical simulation training. *Ann Surg* 2008; **248**: 166–179.
 - 6 Anastakis DJ, Wanzel KR, Brown MH, McIlroy JH, Hamstra SJ, Ali J *et al.* Evaluating the effectiveness of a 2-year curriculum in a surgical skills center. *Am J Surg* 2003; **185**: 378–385.
 - 7 Aggarwal R, Grantcharov TP, Darzi A. Framework for systematic training and assessment of technical skills. *J Am Coll Surg* 2007; **204**: 697–705.
 - 8 Aggarwal R, Grantcharov TP, Eriksen JR, Blirup D, Kristiansen VB, Funch-Jensen P *et al.* An evidence-based virtual reality training program for novice laparoscopic surgeons. *Ann Surg* 2006; **244**: 310–314.
 - 9 Aggarwal R, Grantcharov T, Moorthy K, Hance J, Darzi A. A competency-based virtual reality training curriculum for the acquisition of laparoscopic psychomotor skill. *Am J Surg* 2006; **191**: 128–133.
 - 10 Mackay S, Morgan P, Datta V, Chang A, Darzi A. Practice distribution in procedural skills training: a randomized controlled trial. *Surg Endosc* 2002; **16**: 957–961.
 - 11 Moulton CA, Dubrowski A, Macrae H, Graham B, Grober E, Reznick R. Teaching surgical skills: what kind of practice makes perfect?: a randomized, controlled trial. *Ann Surg* 2006; **244**: 400–409.
 - 12 McDougall EM, Corica FA, Boker JR, Sala LG, Stoliar G, Borin JF *et al.* Construct validity testing of a laparoscopic surgical simulator. *J Am Coll Surg* 2006; **202**: 779–787.
 - 13 Zhang A, Hünerbein M, Dai Y, Schlag PM, Beller S. Construct validity testing of a laparoscopic surgery simulator (Lap Mentor): evaluation of surgical skill with a virtual laparoscopic training simulator. *Surg Endosc* 2008; **22**: 1440–1444.
 - 14 Gallagher AG, Satava RM. Virtual reality as a metric for the assessment of laparoscopic psychomotor skills. Learning curves and reliability measures. *Surg Endosc* 2002; **16**: 1746–1752.
 - 15 Grantcharov TP, Bardram L, Funch-Jensen P, Rosenberg J. Learning curves and impact of previous operative experience on performance on a virtual reality simulator to test laparoscopic surgical skills. *Am J Surg* 2003; **185**: 146–149.
 - 16 Grantcharov TP, Kristiansen VB, Bendix J, Bardram L, Rosenberg J, Funch-Jensen P. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *Br J Surg* 2004; **91**: 146–150.
 - 17 Seymour NE, Gallagher AG, Roman SA, O'Brien MK, Bansal VK, Andersen DK *et al.* Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Ann Surg* 2002; **236**: 458–463.
 - 18 MacRae HM, Satterthwaite L, Reznick RK. Setting up a surgical skills center. *World J Surg* 2008; **32**: 189–195.
 - 19 Darzi A. *Laparoscopic Cholecystectomy Course Handbook*. Royal College of Surgeons of England: London, 1996.
 - 20 Aggarwal R, Moorthy K, Darzi A. Laparoscopic skills training and assessment. *Br J Surg* 2004; **91**: 1549–1558.

4.2 La pratique délibérée sur simulateur de RV améliore la qualité opératoire

4.2.1 Présentation

Deliberate practice on a virtual reality laparoscopic simulator enhances the quality of surgical technical skills. Crochet P, Aggarwal R, Dubb SS, Ziprin P, Rajaretnam N, Grantcharov T, Ericsson KA, Darzi A. Publié dans Annals of Surgery. 2011 ;253(6) :1216-22.

L'objectif principal de ce travail était d'évaluer l'impact de la théorie de la pratique délibérée lors d'une formation sur un simulateur de RV pour une intervention coelioscopique complète : la cholécystectomie.

- Les participants étaient de jeunes chirurgiens inexpérimentés, séparés en 2 groupes (témoin et pratique délibérée) réalisant dix séances d'entraînement sur le simulateur de RV (LapMentor*).
- Le groupe « pratique délibérée » recevait un entraînement spécifique pendant les séances : un feedback immédiat par un formateur à l'aide de deux échelles de qualité (une globale et une spécifique de l'intervention) et des exercices à répéter ciblés en fonction de leurs besoins.
- Il y avait une étape de transfert sur tissus réels (cadavre de cochon).
- Les critères d'évaluation étaient les paramètres métriques enregistrés par le simulateur ainsi que deux échelles de qualité.

Voici les principaux résultats de cette étude :

- La formation « pratique délibérée » permettait d'atteindre un meilleur niveau de qualité opératoire sur le simulateur et également lors du transfert sur tissus réels.

- La courbe d'apprentissage est plus courte dans le groupe contrôle. Les automatismes apparaissant dans ce groupe se faisaient au détriment de la qualité opératoire par rapport au groupe « pratique délibérée ».
- Une plus large proportion de participants dans le groupe « pratique délibérée » estimaient ne pas avoir atteint de plateau en termes de compétences techniques à la fin des séances.

Les questions posées par ce travail étaient :

- Quelles sont les mesures les plus pertinentes pour évaluer les gestes techniques dans des conditions de simulation ?
- La pratique délibérée en laboratoire permet-elle une meilleure compétence lors du transfert au BO ?

Deliberate Practice on a Virtual Reality Laparoscopic Simulator Enhances the Quality of Surgical Technical Skills

Patrice Crochet, MD,* Rajesh Aggarwal, PhD, MA, MRCS,* Sukhpreet Singh Dubb, BSc,* Paul Ziprin, MD, FRCS,* Niroshini Rajaretnam, BSc,* Teodor Grantcharov, MD, PhD,† K. Anders Ericsson, PhD,‡ and Ara Darzi, KBE, FMedSci, HonFREng, MD, FRCSI, FRCS, FACS, FRCPSG*

Introduction: Virtual reality (VR) simulation provides unique training opportunities. This study evaluates whether the deliberate practice (DP) can be successfully applied to simulated laparoscopic cholecystectomy (LC) for enhancement of the quality of surgical skills.

Methods: Twenty-six inexperienced surgeons underwent a training program for LC on a VR simulator. Trainees were randomly allocated to 1 of 2 specific protocols of 10 sessions comprising a total of 20 LCs. For each session, the control group performed 2 LCs separated by 30 minutes of occupational activities; the DP group were assigned 30 minutes of DP activities in between 2 LCs. Each participant then performed 2 LCs on a cadaveric porcine model. Quantitative parameters were recorded from the simulator and a motion tracking device; qualitative assessment utilized validated rating scales.

Results: Twenty-two subjects completed training. Learning curves on the VR simulator were significant for time taken and number of movements in both groups. The DP group was slower from the third LC (1373 vs. 872 seconds, $P = 0.022$) and utilized more movements from the seventh (942 vs. 701, $P = 0.033$). Global rating scores improved significantly in both groups over repeated LCs. The DP group revealed higher scores than control from tenth (19.5 vs. 14, $P = 0.014$) until the twentieth LC (22 vs. 16, $P = 0.003$). On the porcine model, the DP group also achieved higher global rating scores (25.5 vs. 19.5, $P = 0.002$).

Conclusions: VR training improved dexterity for both groups, and led to transfer of skill onto a porcine LC model. The DP group achieved higher quality, and demonstrated superior transfer onto real tissues.

(*Ann Surg* 2011;253:1216–1222)

In recent years, there has been increased attention toward patient safety, which aims to report, analyze and ultimately prevent the occurrence of medical adverse events, to build a safer health care system.¹ This emerging field is in correlation with a growing societal pressure for transparency in medical practice and public expectations of improved strategies to reduce medical errors.² Surgical adverse events account for up to two-third of all adverse events, of which one-half are preventable.³ Errors in surgery may stem from failure of perception (cognitive disposition to respond), cognition (knowledge

and decision making), or execution (mistakes).⁴ Technical errors are the most common type of surgical errors reported in the literature.^{5–7} As the reduction of surgical error rate becomes a major concern for health care systems, the need for enhanced surgical competency seems more acute.⁸

Effective methods to reduce errors lead to the concept of expertise, which represents a very high level of skill acquisition and is the result of a gradual improvement through extended experience in a given domain. However, it seems that expert performance is not the result of extensive experience alone, but rather the consequence of engaging in specific training purposely designed to improve the current level of skill.⁹ Ericsson has proposed that individual differences in attained professional performance can be explained by differences in deliberate practice (DP).^{10,11} This framework is based on the assumption that expert performance results from engagement in activities *deliberately* chosen for their ability to improve performance and maintain it at the highest level. DP calls individuals to focus their training on defined effortful tasks, or drills. It involves repeated practice with immediate feedback on performance, often delivered by teachers and coaches.

Attained level of expertise has been closely related to time devoted to DP in different domains requiring a high level of technical performance such as music, chess and sport. In a typical case of DP a coach designs a training activity to improve a specific aspect of an athlete's performance. These kinds of practice tasks permit athletes to stretch their performance beyond their current level through focus on improvement in a given aspect. Furthermore, by receiving immediate feedback and opportunities to reflect on possible refinements, competitive athletes are able to improve with repeated exposures to similar tasks, or drills, over time.¹²

The principles of DP are potentially of great interest in surgical education, though to our knowledge not empirically investigated to date.¹³ Technical competency is indeed a cornerstone in surgery. However, the learning process of surgical technical skills occurs relatively late in a domain that requires high degrees of hand-eye coordination and dexterity, delayed by the time necessary to acquire essential medical knowledge. Within a system that is still largely based on apprenticeship in the operating theatre, a standardized and rigorous approach is difficult to implement. Occurrence of learning opportunities is dependent on the operating schedule and often does not follow a logical path, adapted to the level of the learner.

After initial acquisition of basic skills, DP training should allow surgical trainees to focus on specific, or personalized, training activities to enhance these skills. The lack of in-built objective measures of performance in the surgical domain makes it difficult to provide immediate and informative feedback regarding performance. In the operating room, it is very difficult if not impossible to apply the principles of DP, which regard to, for example, repeated drills. This may explain the lack of DP techniques in surgical training programs. However, the emergence of high fidelity and sophisticated laparoscopic virtual reality (VR) simulators can provide an opportunity for

From the *Department of Biosurgery and Surgical Technology, Imperial College London, United Kingdom; †Department of Surgery, University of Toronto, Division of General Surgery, St. Michael's Hospital, Toronto, Canada; and ‡Department of Psychology, Florida State University, Tallahassee, FL.

Presented at the 95th Annual Clinical Congress of the American College of Surgeons, 64th Annual Sessions of the Owen H Wangenstein Forum on Fundamental Surgical Problems, October 2009.

Published in abstract form in *Journal of the American College of Surgeons*, September 2009 (Vol. 209, Issue 3, Supplement, Page S112).

Reprints: Rajesh Aggarwal, Department of Biosurgery and Surgical Technology, St Mary's Campus, Imperial College Healthcare NHS Trust, 10th Floor, Queen Elizabeth the Queen Mother Building, St. Mary's Hospital, Praed Street, London W2 1NY, United Kingdom. E-mail: rajesh.aggarwal@imperial.ac.uk.

Copyright © 2011 by Lippincott Williams & Wilkins

ISSN: 0003-4932/11/25306-1216

DOI: 10.1097/SLA.0b013e3182197016

application of such techniques into a new model of surgical progression. This is of interest to the surgical community, educationalists and patients alike.^{14–16}

The first aim of this study is to define whether DP can be successfully applied to training for laparoscopic cholecystectomy (LC) on a VR simulator, with the intended outcome of improved quality of surgical skills. The second aim is to investigate the transfer of surgical skills from a VR simulator onto real tissues, after DP training.

METHODS

Subjects

The study recruited 26 participants. All participants were Imperial College Healthcare NHS Trust medical trainees (15 fifth year medical students and 11 junior postgraduate residents) with limited surgical experience (performed 0 but observed >3 LCs in the operating room). Informed consent was provided upon recruitment to the study. The flow of subjects through the study is illustrated in Figure 1.

Baseline Test of Laparoscopic Skills

All participants initially underwent a test of their baseline laparoscopic skills on a rudimentary VR laparoscopic simulator (MIST-VR, Mentice, Gothenburg, Sweden).^{15,17} Performance was measured by dexterity parameters on 3 different tasks (withdraw-insert, diathermy, and manipulate-diathermy) performed at the easy level twice by each participant. The intention was to ensure that there were no differences in baseline laparoscopic skills between participants.

Didactic Training

All participants subsequently attended a laparoscopic skills training seminar to be introduced to the concepts of LC as previously described.¹⁸ After completion of preliminary common training, each of the 26 novice laparoscopic surgeons was randomly allocated to 1 of 2 groups for VR simulator-based training (control or DP group) by the closed envelope technique.

Virtual Reality Laparoscopic Simulator Training Device

The LapMentor VR laparoscopic simulator (Symbionix Corporation, Cleveland, OH) comprises LC modules which for the purposes of this study enabled training on 4 procedural tasks (clip and cut—retracted gallbladder, clipping, and cutting—2 hands, Calot's triangle dissection, and gallbladder dissection) and a full LC procedure. Construct validity has been previously demonstrated for each of the procedural tasks and full procedures used in this study.¹⁹

Control Group Training Sessions

All subjects in the control group participated in 10 training sessions comprising a total of 20 simulated LCs. They were due to complete 10 sessions within a 6-week period on the recommended frequency of 2 sessions per week. The same protocol was used for each session, as described. Subjects performed a virtual LC on the VR simulator. They subsequently viewed an online 30-minute surgical tutorial, on a topic unrelated to laparoscopy or cholecystectomy, and finally completed another LC on the VR simulator. The subjects did not receive commentary or advice during or after their performance on the simulator, with the laparoscopic camera held by an assistant solely under direction of the operating subject.

Deliberate Practice Sessions

All subjects in the DP group participated in 10 training sessions following a similar pattern to those in the control group, ie, 2 LC on

TABLE 1. Correspondence Between Rating Scale Criteria and Training Task Assigned on Lapsim (LS) or LapMentor (LM) Simulator

| Scale Item | Assigned task |
|--------------------------------------|----------------------------------|
| Generic rating scale | |
| Respect for tissue | Clipping and Cut (LS) |
| Time and motion | Lift and Grasp (LS) |
| Instrument handling | Grasping (LS) |
| Use of assistants | Coordination (LS) |
| Procedure specific rating scale | |
| Cystic duct dissection | Calot's triangle dissection (LM) |
| Cystic duct clipping and transection | Clipping and Cutting (LM) |
| Gallbladder fossa dissection | Gallbladder separation (LM) |

the LapMentor separated by a 30-minute period of activities specific to this group. This period was structured to include fundamental criteria of DP theory, as described. During the first VR-based full LC, a qualified observer (P.C.) rated performance of the subject on 2 previously validated rating scales of surgical technical skill (the OSATS-derived global rating scale, and a procedure-specific rating scale for LC).²⁰ Immediate feedback was provided by the observer to the subject of marks awarded, with reasons explained. A specific task, or drill, was then assigned, explained, demonstrated and then performed repeatedly by the subject, to focus the training efforts to the weakest skills of the subject.

The rules of assignment of the drills were as follows: if a participant was found to score less than 3 of 5 on any of the generic technical skills (of 4 categories), then the category marked with the lowest score was assigned and this was linked to a specific training exercise to be performed and repeated over a period of 20 minutes. Four tasks were chosen for their ability to train on each of these 4 generic skills from the basic skills module of the Lapsim VR simulator (Surgical Science, Gothenburg, Sweden)¹⁶ (Table 1). In case the subject had 2 or more equally low scores, they were given a choice on which of the corresponding basic skills to train on. If the participants were marked with a 3 or greater on all 4 generic technical skills, they were then assigned to a training exercise based on the lowest rated criteria obtained on the procedure-specific rating scale, ie, 1 of the 3 procedural tasks on the LapMentor simulator (Table 1). The specific training exercise was similarly performed and repeated over a period of 20 minutes. Finally, participants had to perform a second LC on the LapMentor simulator in the same manner as for the first.

Cadaveric Porcine Laparoscopic Cholecystectomy

After completion of the 10 VR training sessions, all subjects performed 2 LCs on a porcine model, over a maximum period of 1 week. Procedures were implemented as described in our previous work, including use of diathermy.¹⁸ Motion tracking with synchronous video capture utilizing previously described in-house software was employed, to record the procedure from initial entry of the endoscope into the box trainer through to complete detachment of the gallbladder from the liver bed.

Posttraining Session

During a final session, all participants were requested to perform the same tasks they had initially completed on the MIST-VR simulator during the baseline tests. This enabled comparison between control and DP groups in terms of the laparoscopic dexterity levels achieved at completion of the study, independent of any task, which had been used during the training sessions. Subjects also were requested to complete a questionnaire reporting thoughts regarding VR

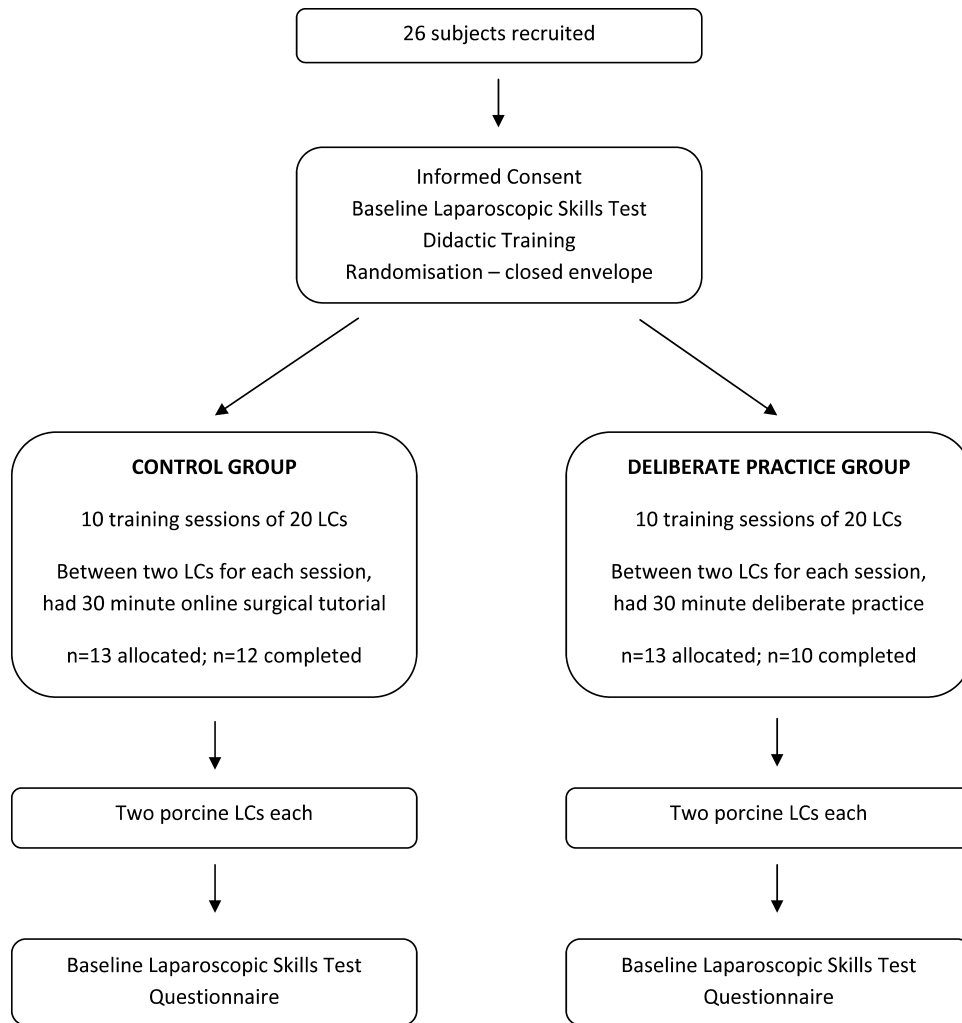


FIGURE 1. Flow of participants through study.

training on the simulators, and their opinion regarding progression of their own laparoscopic skills during the study.

Assessment Tools

Baseline and posttest performances on the MIST-VR simulator were assessed for the task manipulate-diathermy on the basis of previously validated quantitative parameters, ie, time taken, economy of movement for each hand and total error score.¹⁵ Data of the second attempt was only considered to account for the familiarization process of the simulator.

All participants performed a total of 20 full LC procedures each. These performances were analyzed quantitatively based upon 3 previously validated parameters derived automatically and objectively from the simulator, ie, time taken, total number of movements and total path length.¹⁹ All full LC procedures were also recorded with FRAPS real time video capture software (Beepa Pty Ltd, Melbourne, Victoria, Australia), and performance numbers 1, 5, 10, 15, and 20 were arbitrarily chosen for qualitative analysis, in the form of blinded video review of the procedure. This was carried out by 2 experienced laparoscopic surgeons (R.A. and T.G.) blinded to the identity, group and trial number of each subject. Procedures were marked on the

aforementioned OSATS-derived generic global rating scale (of 30), and the procedure-specific rating scale (of 25).^{20,21,22}

The cadaveric porcine LC procedures were assessed in terms of dexterity by the motion tracking system on the basis of time taken, total number of movements and total path length.^{23,24} Qualitative analysis of videos was carried out by the same 2 experienced laparoscopic surgeons blinded to the identity, group and trial of each subject on the OSATS-derived generic global rating scale (of 35) and the procedure-specific rating scale (of 25).^{20,21,22}

Statistical Analysis

The choice of at least 10 subjects per group was based on a 2-tailed test, with $\alpha = 0.05$ and power $(1 - \beta) = 0.80$, and an intended reduction of 30% in time taken to complete tasks for intervention versus control groups, based on data from previous studies of VR simulation.^{15,16} The nature of data were non parametric, thus the Mann-Whitney U test was employed for intergroup analysis regarding baseline data on the MIST, full LC procedures on the LapMentor and porcine model LCs, and the Wilcoxon signed-rank test to analyse pair-wise data. Data on learning curves for the LapMentor LCs was analyzed by the Friedman test (nonparametric repeated measures ANOVA). Multiple comparisons were then made to identify when

skills had plateaued. Cronbach's alpha test statistic was used to ascertain the interrater reliability of the video-based scoring system. The χ^2 test was used to analyse parametric data when appropriate in the questionnaire. The Statistical Package for the Social Science version 15.0 (SPSS, Chicago, IL) was used, with results reported as median values and a level of $P < 0.05$ considered statistically significant.

RESULTS

Subjects

Twenty-two (85%) of the 26 subjects recruited completed the study, the remainder citing timetabling constraints for their inability to fulfill the demands of the study. There were 12 subjects within the control group (6 medical student and 6 junior residents), and 10 subjects within the DP group (5 medical students and 5 junior residents). The median age was 24 years, and the male:female ratio was 10:12. Participants were right-handed, except 2 left-handed subjects within the control group.

Baseline Test of Laparoscopic Skills

Baseline scores of laparoscopic skill on the MIST-VR simulator did not demonstrate any significant differences between control and DP groups on the manipulate-diathermy task for time taken (47 vs. 52.2 seconds, $P = 0.879$) right economy of movements (2.2 vs. 2.3, $P = 0.970$) left economy of movement (2.3 vs. 2.1, $P = 0.909$) and total error scores (70.5 vs. 76, $P = 0.518$).

Virtual Reality Laparoscopic Simulator Training Sessions

For the control group, analysis of learning curves for quantitative parameters revealed significant results for time taken with plateau at the 6th LC (756 seconds, $P = 0.053$), total number of movements with plateau at the 9th LC (580, $P = 0.079$) and total path length with plateau at the 6th LC (1566 cms, $P = 0.150$). For the DP group, learning curves were statistically significant for time taken with plateau at the 9th LC (1318 seconds, $P = 0.073$), total number of movements with plateau at the 10th LC (821, $P = 0.153$) though there was no significant learning curve found for total path length (Fig. 2A–C). Step-by-step comparisons between the 2 groups revealed the DP group to be significantly slower to complete a full procedure from the 3rd LC (1373 vs. 872 seconds, $P = 0.022$) and also utilized a greater number of movements from the 7th LC (976 vs. 701, $P = 0.033$) until the end of the training. There was however no significant difference between the groups for total path length.

In terms of video analysis, the control group achieved significant improvements over repeated sessions on both the generic and procedure-specific rating scales, with plateau at the 5th LC ($P = 0.405$ and 0.180, respectively). The DP group achieved significant improvements too on global and procedure-specific rating scales, with plateau at the 10th LC ($P = 0.309$ and 0.080, respectively). Step-by-step comparisons between control and DP groups revealed significantly higher scores for DP group at the 10th, 15th, and 20th LCs for both global and procedure-specific rating scales (Table 2 and Fig. 3).

Porcine Model Performance

Comparison of motion tracking parameters revealed the control group to be faster than the DP group to complete a full LC during the first (2736 vs. 4178 seconds, $P = 0.031$) and second LC (1989 vs. 2834 seconds, $P = 0.026$). However, there were no significant differences between the 2 groups regarding the total number of movements and total path length. With regard to video-based analyses between control and DP groups during the first LC, there were no statistically significant differences on the OSATS-derived generic global scale (18

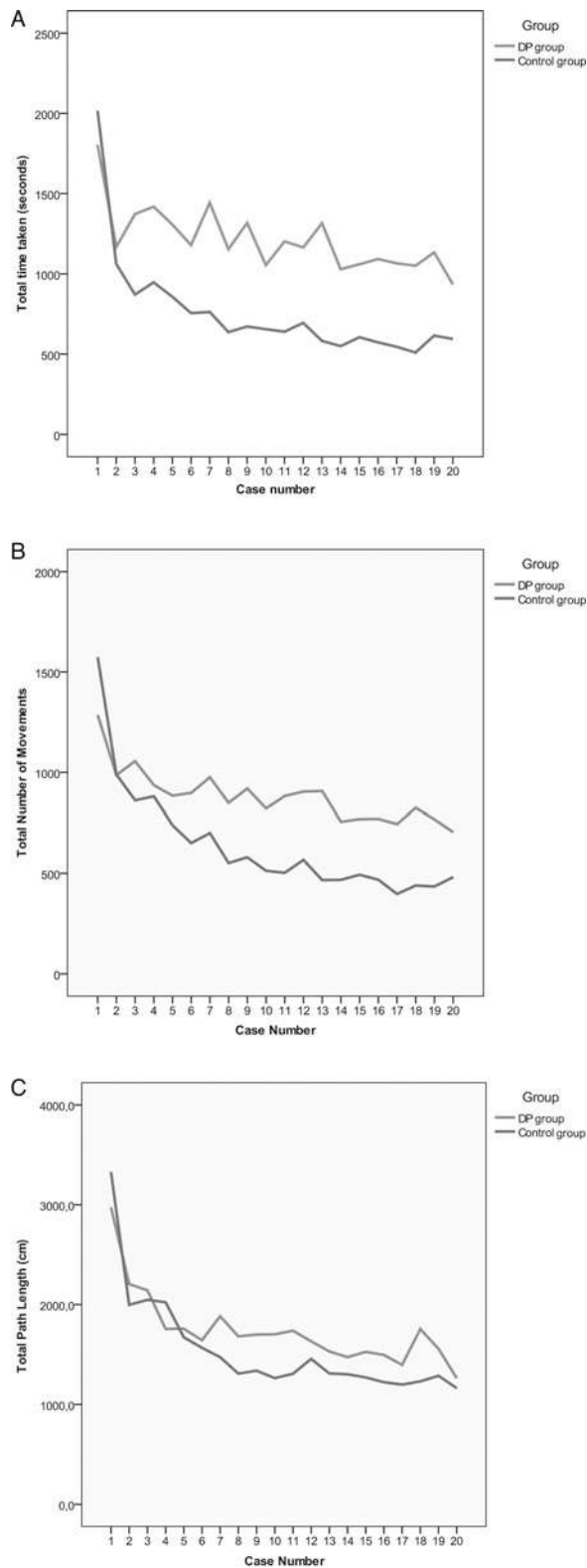


FIGURE 2. (A) Total time taken for full simulated LC (procedures 1–20). (B) Number of movements for full simulated LC (procedures 1–20). (C) Total path length for full simulated LC (procedures 1–20).

TABLE 2. Comparison of Video-Based Scores Between Control and Deliberate Practice Groups for Simulated LCs

| LC Number | Global Rating Scale (of 30) | | | Procedural Rating Scale (of 25) | | |
|-----------|-----------------------------|----------|--------------|---------------------------------|----------|------------------|
| | Control Group | DP Group | P | Control Group | DP Group | P |
| 1 | 11 | 13 | 0.185 | 12 | 10 | 0.774 |
| 5 | 14.5 | 15.5 | 0.340 | 13 | 15 | 0.129 |
| 10 | 14 | 19.5 | 0.014 | 14 | 17 | 0.027 |
| 15 | 12.5 | 19.5 | 0.004 | 13 | 19.5 | 0.003 |
| 20 | 16 | 22 | 0.003 | 14 | 20.5 | <0.001 |

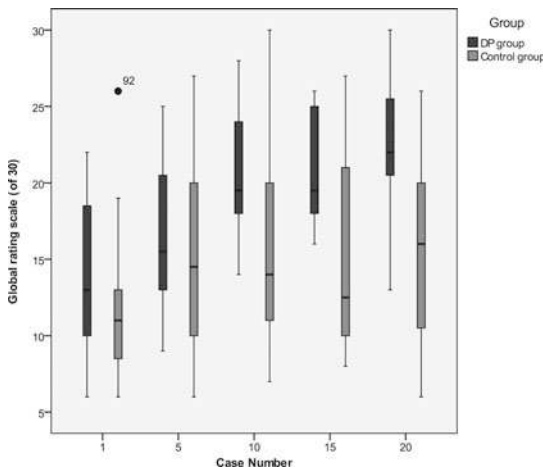


FIGURE 3. Generic global rating score for full simulated LCs (procedures 1, 5, 10, 15, and 20).

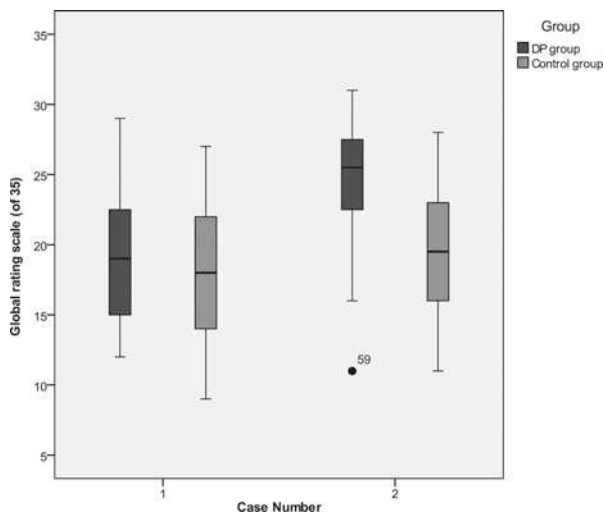


FIGURE 4. Generic global rating scale for porcine LC (procedures 1 and 2).

vs. 19, $P = 0.335$), though the DP group were significantly superior on the procedure-specific scale (10 vs. 12, $P = 0.019$). During the second LC, the control group were significantly inferior to the DP group for generic (19.5 vs. 25.5, $P = 0.002$) and procedural (11 vs. 15, $P = 0.008$) rating scales (Fig. 4).

Video-based Interrater Reliability

The interrater reliability for video-based assessment of the VR and porcine LCs was measured for the global and procedure-specific rating scales separately, revealing values of $\alpha = 0.684$ and 0.691 , respectively.

Posttraining Session

The posttraining performance of control and DP groups for the manipulate-diathermy task on the MIST-VR simulator did not reveal significant differences in terms of time taken (28.4 vs. 34.6 seconds, $P = 0.231$), right economy of movements (1.8 vs. 1.75, $P = 0.915$), left economy of movements (2.0 vs. 1.85, $P = 0.571$) or total error score (32 vs. 35, $P = 0.888$). Furthermore, the control and DP groups had significantly improved their performance in comparison to the baseline test session for time taken (47 vs. 28.4 seconds, $P = 0.005$ and 52.2 vs. 34.6 seconds, $P = 0.021$, respectively), and total error score (70.5 vs. 32, $P = 0.009$ and 76 vs. 35, $P = 0.015$, respectively).

Questionnaire

All study subjects answered a questionnaire at completion of the study period that reported a high degree a satisfaction (median: 5 of 5) and confidence (median: 7 of 10) in the use of VR simulation-based training before surgical practice. However, the subjects differed in their perception of the level of skill achieved with greater numbers of participants in the control versus DP group stating they had achieved a plateau in their skills after completion of the training program (73 vs. 20%, $P = 0.016$).

DISCUSSION

The importance of DP in attainment of expert performance was first described in a study of expert musicians studying in Berlin. Those in the best groups were found to spend more time in solitary practice, concentrating upon the improvement of specific aspects of the music performance, as directed by their music teachers. The best experts spent 4 hours/d including weekends on this type of solitary practice. By the age of 20, the best expert musicians had spent over 10,000 hours of practice. Since then, several studies and reviews have reported similar associations in chess, typing, athletics and tennis. Commentators have also used the concept of DP to understand skill acquisition in medicine, and described this for pulmonary auscultation, mammogram interpretation and medical diagnosis. Though surgery is an obvious area for the need to acquire expert performance, there has to our knowledge been no empirical study to date that has tested this hypothesis.

The aims of this study were to assess the effect of DP on the learning curve for LC on a VR simulator, and subsequently whether this transferred to enhanced outcomes on realistic tissues. At baseline, subjects in both groups were similar in terms of technical laparoscopic skills. After training on the LapMentor VR simulator, skills improved with regard to dexterity parameters in both groups. However, members of the control group completed the procedures in a shorter time

and with fewer numbers of movements than the DP group, with a greater rate and final level of plateau. Conversely, the DP group achieved significantly higher scores on the video-based rating scales, which are a measure of quality rather than dexterity alone. For the 10th, 15th, and 20th LCs on the simulator, the DP group achieved global rating scores of 39%, 56%, and 50% greater than those of the control group, respectively, with similar results achieved for the procedural rating scale.

It can thus be seen that those in the DP group were trained to a higher level of skill with respect to quality of performance, though this was at the expense of dexterity. It is likely that DP led to greater care and precision of the operative procedure, which during skills acquisition takes longer and may result in greater numbers of movements. It is though likely that further practice beyond the 20 procedures prescribed in this study for the DP group would lead to enhanced dexterity toward the level of the control group. The converse with regard to improvements in quality of the procedure of the control group may not be so clearly stated.

This finding raises the question about the most appropriate method to achieve high quality surgical skills. The control group participants who did not receive any feedback implicitly generated their own instructions toward completion of the task. On the contrary, the DP group participants aimed to perform the procedures in a meticulous manner. This echoes the Fitts and Posner 3 stage theory of motor skill acquisition, widely accepted in both the motor skills literature and the surgical literature.²⁷ After the first *cognitive* stage, the learner needs to advance through an *integrative* stage during which knowledge is translated into appropriate motor behavior, leading to fluidity and rapidity of execution by which time the last *autonomous* stage is attained. In this study, all participants progressed through a cognitive stage after theoretical training. The control group seems to have neglected the second integrative stage and seemingly advanced toward an autonomous stage. On the contrary, the DP group remains at the integrative stage of motor skill acquisition, continuing to refine the quality of the surgical procedure. Advancement onto the autonomous stage is yet to be achieved.

Nonetheless, the literature on expert performance has questioned this model of expertise, with the proposition that expert performers never automate their performance but gain increasing control over their actions. This enables them to retain an ability to make intentional adjustments required for unexpected situations.^{28,29} Research in music, ballet, and ice-skating, where a specific sequence of precise motor action must be learnt, shows that the key object is not to play a piece as fast as possible but rather doing so with fundamentally correct movements and control so they can deliver their own interpretation of the piece. Only when sufficient control is attained can the execution of movements be increased in speed although preserving control. Following this logic, it can be assumed that, in case of further training, the DP group participants would improve their speed and reduce their number of movements in keeping the benefits of higher quality surgical skills. The responses to the questionnaire completed by all subjects at the end of the study support the previously developed hypothesis, in showing that fewer participants in the DP group thought they had achieved plateau in their laparoscopic surgical skills, and thus are aware that further skills improvement needs to occur.

The results of this study also raise questions about the relevance of quantitative parameters to assess surgical performance. Their ability to assess dexterity at an early stage on basic tasks remains. This is illustrated by the results on the MIST-VR simulator. Both groups improved similarly during the study on these tasks. It clearly seems that quantitative parameters cannot be used alone to evaluate of the qualitative skills in the context of a complex full procedure.

With regard to transfer of the learned skills to real tissues, in this case a cadaveric porcine gallbladder model, similar results were

noted to performance on the VR simulator. The DP group was once again slower than the control group, though there was no difference in the other dexterity parameters. However, those who had undergone the process of DP once again had higher scores than the control group on both the global and procedure-specific rating scales of the order of 29% and 36%, respectively. This aspect of the study is important, as it confirms that skills acquired during DP on a VR simulator are transferable to a realistic model. It is thus highly likely that DP will be transferable onto real procedures in the operating room.

This is the first study that has investigated the path toward acquisition of expert performance in surgical technical skills. Training was provided on a VR laparoscopic simulator with regard to the basic tenets of DP, ie, immediate feedback of performance, provided by a coach or teacher, leading to repeated and focused practice, on specific drills which relate to aspects of overall performance. It is at this moment unclear whether each tenet of DP contributes an equal amount to skills gain in the surgical arena, or if this is a compound effect of feedback, repeated practice and drills, following surgical procedures. Further, would ad hoc and nonformalized exposure to the skills laboratory in between patient cases yield similar benefits? These important questions are the basis of future research in this area.

Whether this study provides enough evidence to lead to a new paradigm in the training of future surgical specialists is unknown. However, it is absolutely clear that repeated performance, whether on a simulator or a real case, is not adequate to lead to the development, and indeed maintenance of expertise. The analogy to surgical training programs is the need to disregard numbers of procedures performed to achieve independent practice, and insist upon a measure of quality of the procedure performed. Current charting of learning curves for surgical trainees adheres to the former model whereby plateau in time taken, conversion rates or morbidity are deemed to be satisfactory evidence for achievement of expertise. This may only be a measure of experience, which may not directly correlate with expertise.

In conclusion, this study has demonstrated enhanced quality of surgical skills can be achieved with DP, both on simulated and realistic tissues. Transposition of these findings to the clinical environment suggests that surgical trainees should be provided with immediate objective feedback on a rating scale following a specific procedure, then repeatedly train on an appropriate task on a simulator, ideally to levels of proficiency, and only then progress to perform a second procedure within the operating room, and so on. This is eminently workable, though as well as a shift in learning technical skills outside the operating room, must be driven by trainers to adopt this approach for their trainees, and indeed their patients. It is then possible that the technical basis of operating room performance can be enhanced through DP within the skills laboratory, leading to higher quality surgical skills and perhaps reduction in preventable error.

ACKNOWLEDGMENTS

Rajesh Aggarwal is funded by a Clinician Scientist Award from the National Institute for Health Research, London, United Kingdom.

REFERENCES

1. Institute of Medicine. *To Err is Human: Building a Safer Health System*. Washington, DC: National Academies Press; 2000.
2. Vincent C, Young M, Phillips A. Why do people sue doctors? A study of patients and relatives taking legal action. *Lancet*. 1994;25:343(8913):1609–1613.
3. Gawande AA, Thomas EJ, Zinner MJ, et al. The incidence and nature of surgical adverse events in Colorado and Utah in 1992. *Surgery*. 1999;126:66–75.
4. Champion HR, Meglan DA, Shair EK. Minimizing surgical error by incorporating objective assessment into surgical education. *J Am Coll Surg*. 2008;207(2):284–291.

5. Regenbogen SE, Greenberg CC, Studdert DM, Lipsitz SR, Zinner MJ, Gawande AA. Patterns of technical error among surgical malpractice claims: an analysis of strategies to prevent injury to surgical patients. *Ann Surg.* 2007;246(5):705–711.
6. Leape LL, Brennan TA, Laird N, et al. The nature of adverse events in hospitalized patients. Results of the Harvard Medical Practice Study II. *N Engl J Med.* 1991;324(6):377–384.
7. Thomas EJ, Studdert DM, Newhouse JP, et al. Costs of medical injuries in Utah and Colorado. *Inquiry.* 1999;36(3):255–264.
8. Cuschieri A, Francis N, Crosby J, Hanna GB. What do master surgeons think of surgical competence and revalidation? *Am J Surg.* 2001;182(2):110–116.
9. Ericsson KA. The influence of experience and deliberate practice on the development of superior expert performance. In Ericsson KA, Charness N, Feltovich P, et al., eds. *Cambridge Handbook of Expertise and Expert Performance.* Cambridge, UK: Cambridge University Press; 2006:685–706.
10. Ericsson KA, Krampe RTh, Tesch-Romer C. The role of deliberate practice in the acquisition of expert performance. *Psychol Rev.* 1993;100(3):363–406.
11. Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Acad Med.* 2004;79(10 Suppl):S70–S81.
12. Ericsson KA. Deliberate practice and the modifiability of body and mind: toward a science of the structure and acquisition of expert and elite performance. *Int J Sport Psychol.* 2007;38:4–34.
13. Reznick RK, MacRae H. Teaching surgical skills—changes in the wind. *N Engl J Med.* 2006;355(25):2664–2669.
14. Aggarwal R, Darzi A. Technical-skills training in the 21st century. *N Engl J Med.* 2006;355(25):2695–2696.
15. Aggarwal R, Grantcharov T, Moorthy K, et al. A competency-based virtual reality-training curriculum for the acquisition of laparoscopic psychomotor skill. *Am J Surg.* 2006;191(1):128–133.
16. Aggarwal R, Grantcharov TP, Eriksen JR, et al. An evidence-based virtual reality-training program for novice laparoscopic surgeons. *Ann Surg.* 2006;244(2):310–314.
17. Taffinder N, Sutton C, Fishwick RJ, et al. Validation of virtual reality to teach and assess psychomotor skills in laparoscopic surgery: results from randomised controlled studies using the MIST VR laparoscopic simulator. *Stud Health Technol Inform.* 1998;50:124–130.
18. Aggarwal R, Ward J, Balasundaram et al. Proving the effectiveness of Virtual Reality for training in laparoscopic surgery. *Ann Surg.* 2007;246(5):771–779.
19. Aggarwal et al. A Scientific method for development of a whole procedure virtual reality training curriculum for laparoscopic cholecystectomy. *Br J Surg.* 2009;96(9):1086–1093.
20. Sarker SK, Chang A, Vincent C, et al. Development of assessing generic and specific technical skills in laparoscopic surgery. *Am J Surg.* 2006;191(2):238–244.
21. Martin JA, Regehr G, Reznick R, et al. Brown Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg.* 1997;84(2):273–278.
22. Aggarwal R, Grantcharov T, Moorthy K, et al. Toward feasible, valid, and reliable video-based assessments of technical surgical skills in the operating room. *Ann Surg.* 2008;247(2):372–379.
23. Dosis A, Aggarwal R, Bello F, et al. Synchronized video and motion analysis for the assessment of procedures in the operating theater. *Arch Surg.* 2005;140(3):293–299.
24. Aggarwal R, Grantcharov T, Moorthy K, et al. An evaluation of the feasibility, validity, and reliability of laparoscopic skills assessment in the operating room. *Ann Surg.* 2007;245(6):992–999.
25. Seymour NE, Gallagher AG, Roman SA, et al. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Ann Surg.* 2002;236(4):458–463.
26. Grantcharov TP, Kristiansen VB, Bendix J, et al. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *Br J Surg.* 2004;91(2):146–150.
27. Fitts PM, Posner MI, *Human Performance.* Belmont, CA: Brooks/Cole; 1967.
28. Ericsson KA, Kintsch W. Long-term working memory. *Psychol Rev.* 1995;102(2):211–245.
29. Ericsson KA, Patel V, Kintsch W. How experts' adaptations to representative task demands account for the expertise effect in memory recall: comment on Vicente and Wang (1998). *Psychol Rev.* 2000;107(3):578.

5. DÉVELOPPEMENT D'UN CURRICULUM SUR SIMULATEUR DE RV POUR L'HYSTÉRECTOMIE TOTALE PAR VOIE CŒLIOSCOPIQUE (HC)

5.1 Présentation

Development of an evidence-based training program for laparoscopic hysterectomy on a virtual reality simulator. Crochet P, Aggarwal R, Knight S, Berdah S, Boubli L, Agostini A.

Publié dans Surgical Endoscopy en septembre 2016.

L'objectif principal de ce travail était de construire, selon une méthode scientifique, un curriculum d'entraînement sur un simulateur de RV (LapMentor*) pour une intervention cœlioscopique complète : l'HC

- Les 2 modules disponibles sur le simulateur ont été testés : 1 programme avec des guides de couleur en 10 étapes successives, reproduisant l'intervention après la mise en place du manipulateur utérin jusqu'à la fin de la colpotomie. 1 programme reproduisant la même intervention sans guides de couleur.
- Les preuves de validité de chacun des exercices étaient explorées en comparant les performances de 3 groupes d'opérateurs d'expérience clinique différente (externes inexpérimentés, internes de niveau intermédiaire, seniors expérimentés).
- Les critères d'évaluation étaient les paramètres métriques automatiquement enregistrés par le simulateur, et une évaluation de la qualité opératoire via une échelle validée (OSATS).
- Une analyse de la courbe d'apprentissage était réalisée pour le groupe inexpérimenté par la répétition 10 fois du même exercice.

Voici les principaux résultats de cette étude :


- Les 2 modules testés ont montré des preuves de validité pour discerner les compétences techniques entre les 3 groupes selon des critères à la fois quantitatifs et qualitatifs.
- Les différences observées au niveau de la qualité opératoire étaient significatives pour chacun des groupes d'opérateurs, globalement et comparés deux à deux.
- Une courbe d'apprentissage était observée pour tous les critères métriques valides, avec un plateau apparaissant entre la 2^{ème} et la 6^{ème} répétition. Une courbe d'apprentissage était observée selon l'échelle de qualité opératoire, sans plateau à la 10^{ème} répétition.
- Le résultat est la proposition d'un parcours d'entraînement spécifique de l'HC, selon la théorie des curricula.

Les questions posées par cette étude et ses limites étaient :

- Peut-on confirmer la validité externe et la faisabilité de ce curriculum appliqué à une plus large population ?
- Les compétences acquises sur le simulateur persistent-elles lors du transfert au BO (niveau 3 de Kirkpatrick) ? Si oui, quel est le gain réalisé grâce au curriculum en termes de courbe d'apprentissage au BO par rapport au compagnonnage classique ?
- À quel moment de la formation ce curriculum devrait-il être utilisé ? Au début de la formation (comme le curriculum sur pelvi-trainer) lorsque les internes ont encore peu d'opportunités opératoires ou dans la deuxième partie de la formation après acquisition des compétences de base ?
- La pratique délibérée intégrée à ce programme d'entraînement aurait-elle une influence positive sur les résultats ?

- Un élément essentiel à la réalisation de l'intervention - la suture vaginale - n'a pas pu être étudiée car le programme correspondant à ce temps opératoire n'était pas disponible au moment de la collecte des données. Il devra donc faire l'objet d'un travail futur.

Development of an evidence-based training program for laparoscopic hysterectomy on a virtual reality simulator

Patrice Crochet¹  · Rajesh Aggarwal² · Sophie Knight¹ · Stéphane Berdah³ · Léon Boubli¹ · Aubert Agostini¹

Received: 26 April 2016 / Accepted: 12 September 2016
© Springer Science+Business Media New York 2016

Abstract

Background Substantial evidence in the scientific literature supports the use of simulation for surgical education. However, curricula lack for complex laparoscopic procedures in gynecology. The objective was to evaluate the validity of a program that reproduces key specific components of a laparoscopic hysterectomy (LH) procedure until colpotomy on a virtual reality (VR) simulator and to develop an evidence-based and stepwise training curriculum.

Methods This prospective cohort study was conducted in a Marseille teaching hospital. Forty participants were enrolled and were divided into experienced (senior surgeons who had performed more than 100 LH; $n = 8$), intermediate (surgical trainees who had performed 2–10 LH; $n = 8$) and inexperienced ($n = 24$) groups. Baselines were assessed on a validated basic task. Participants were tested for the LH procedure on a high-fidelity VR simulator. Validity evidence was proposed as the ability to differentiate between the three levels of experience. Inexperienced subjects performed ten repetitions for learning curve analysis. Proficiency measures were based on experienced surgeons' performances. Outcome measures were simulator-derived metrics and

Objective Structured Assessment of Technical Skills (OSATS) scores.

Results Quantitative analysis found significant inter-group differences between experienced intermediate and inexperienced groups for time (1369, 2385 and 3370 s; $p < 0.001$), number of movements (2033, 3195 and 4056; $p = 0.001$), path length (3390, 4526 and 5749 cm; $p = 0.002$), idle time (357, 654 and 747 s; $p = 0.001$), respect for tissue (24, 40 and 84; $p = 0.01$) and number of bladder injuries (0.13, 0 and 4.27; $p < 0.001$). Learning curves plateaued at the 2nd to 6th repetition. Further qualitative analysis found significant inter-group OSATS score differences at first repetition (22, 15 and 8, respectively; $p < 0.001$) and second repetition (25.5, 19.5 and 14; $p < 0.001$).

Conclusions The VR program for LH accrued validity evidence and allowed the development of a training curriculum using a structured scientific methodology.

Keywords Laparoscopic hysterectomy · Simulation · Education · Virtual reality

Following a larger trend to improve patient safety in healthcare systems [1], surgical education has shifted from an exclusively companionship-based training to one integrating simulation. The use of simulators for training is well supported by evidence on technical and non-technical aspects of surgical practice [2, 3]. Although much has been written in the literature relating to effective frameworks for technical skills training programs, other than the FLS, few evidence-based comprehensive curricula have been adopted for minimally invasive surgery (MIS) [4]. Furthermore, most of them focus on basic technical skills. Box trainers have proven to be highly effective for training on

✉ Patrice Crochet
pcrochet.marseille@gmail.com

¹ Department of Obstetrics and Gynecology, Assistance Publique - Hôpitaux de Marseille, La Conception Hospital, Aix Marseille Université, 147 Boulevard Baille, 13005 Marseille, France

² Department of Surgery, Arnold and Blema Steinberg Medical Simulation Centre, McGill University, Montreal, Canada

³ CERC, IFSTTAR, LBA UMR_T 24, Aix Marseille Université, 13916 Marseille, France

laparoscopic tasks like knot tying [2] but are not able to reproduce a sequence of tasks specific of a given laparoscopic procedure. Due to technology improvement, programs reproducing entire MIS procedures are increasingly available on virtual reality (VR) simulators. However, very few of them have been tested for the development of curricula so far [5, 6]. Filling this gap in procedure specific evidence-based curricula is one of the keys to enable wider incorporation of simulation in surgical training programs [7]. In gynecology, there is no curriculum available regarding laparoscopic hysterectomy (LH), and training remains largely based on companionship [8].

The aim of this study was to evaluate the validity of a VR program reproducing key specific components of a LH procedure until colpotomy, in terms of quantitative and qualitative parameters, and to develop an evidence-based and stepwise training curriculum for surgical residents.

Materials and methods

Participants

Participants were divided into three groups based on their levels of experience. Forty participants were enrolled, comprising 8 experienced, 8 intermediates and 24 inexperienced.

The experienced group was composed of senior surgeons who had performed more than 100 LH. The intermediate group consisted of surgical trainees (5th year residents) who had performed 2–10 full LH as first operator and >25 LH as first assistant. The inexperienced group was composed of 5th year medical students having done 3-month rotation in a gynecologic operating room (OR) and assisted ≥ 3 LH procedures. Inexperienced participants were given a lecture on pelvic anatomy, fundamentals in laparoscopy and a step-by-step description of the LH procedure [9–11].

Virtual reality simulation tool

The LAP Mentor™ VR laparoscopic surgical simulator (Symbionix-3D Systems, Cleveland, Ohio, USA) provides with 9 basic skill tasks and one LH program (presenting as a 6 to 8 week size regular uterus consistent with adenomyosis). This program includes bilateral salpingo-oophorectomy. It starts after port insertion, and once the uterine manipulator is in place, and ends with circumferential colpotomy. The same procedure can be conducted with a “guidance” mode, i.e., 10 consecutive steps with colored signs to be completed. The subject interacts with the simulated tissues using two generic laparoscopic instruments through a haptic feedback device (Fig. 1).

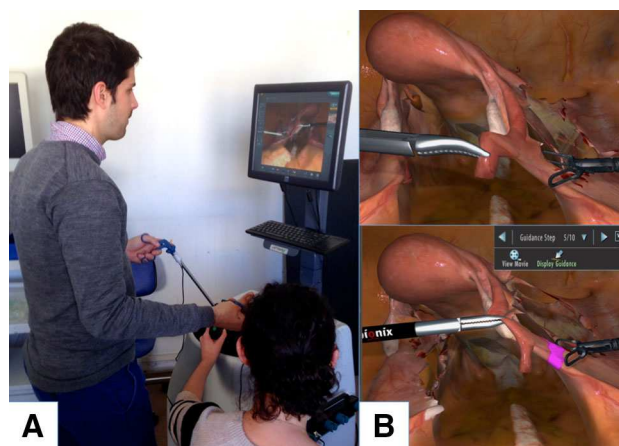


Fig. 1 **A** Virtual reality LapMentor simulator. **B** Description of the infundibulopelvic ligament division step, with colored guidance (*bottom*) and without colored guidance (*top*)

Instruments may be selected through an on-screen menu. The operator can select an optional third grasping instrument for exposure. In addition, the uterine manipulator is controlled by the operator, with six degrees of freedom.

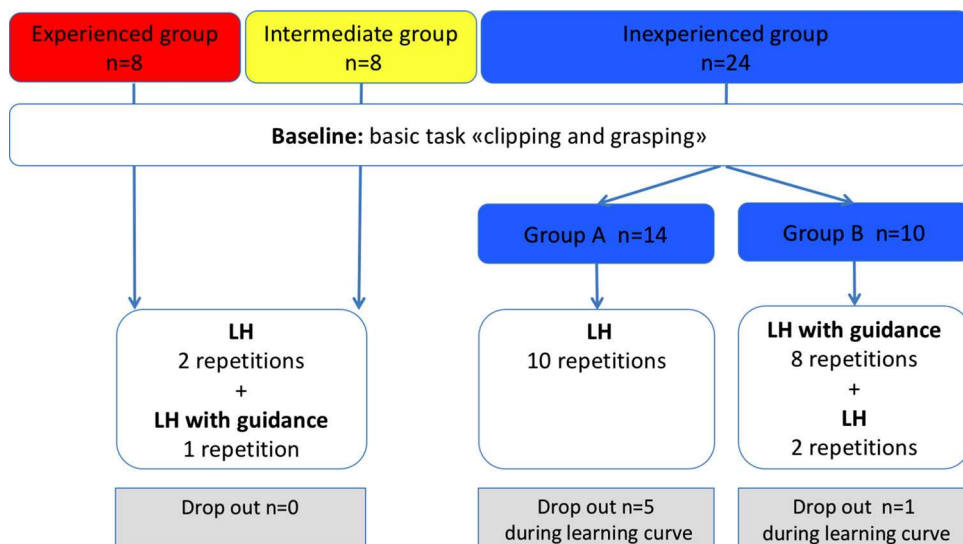
Tasks performed

Each participant initially performed two repetitions of a validated basic task (“clipping and grasping”) [6] to measure baseline skills. Experienced and intermediate subjects performed LH for two sessions, followed by LH with guidance for one session. Inexperienced participants were randomly allocated to two groups: group A ($n = 14$) performed LH for ten sessions and group B ($n = 10$) performed LH with guidance for eight sessions followed by LH for two sessions (Fig. 2). Before each task was performed, a senior surgeon performed a demonstration of the task. No assistance was provided during hands-on training, and participants were provided with a passive assistant holding the camera at their command. No feedback was provided to any of the participants, but they were all given access to the didactic training presentation and the expert’s video. All sessions were completed at least 1 h apart, and those in the inexperienced group completed no more than two sessions per day.

Performance evaluation

The VR simulator measured data for each of the performed tasks instantly and objectively. All parameters automatically recorded by the simulator were taken into account for validity analysis. This comprised metrics ranging from time taken and number of movements to colpotomy incision accuracy and idle time (total time for which instruments are in movement but are not in contact with tissue).

Fig. 2 Study design (LH laparoscopic hysterectomy)



In addition, the VR simulator recorded the occurrence of organ injuries (i.e., bladder, colon or vascular injuries).

Validity was considered a unitary concept in which construct validity is the singular validity type, following consensus standards of the American Educational Research Association [12]. Validity evidence was based upon whether the model could differentiate between three different levels of experience [13]. Comparison of performance at the first session between the three groups of operators was used to assess whether LH and the simulator’s defined settings possessed validity evidence.

Statistical analysis of the learning curves for inexperienced group was used to determine whether repeated practice improved performance toward that of the experienced group. Benchmark levels were defined by calculation of the experienced group’s median scores for each parameter displaying validity evidence, during the second LH repetition. Assessment of experienced surgeons’ second repetition was assumed to reduce the effect of their familiarization with the simulator during the first repetition. Same calculations were made for LH with guidance at first session, as only one LH with guidance was performed for this group. The way novices advanced through these clearly defined steps through comparative measurement of simulator-derived metrics that are construct valid, learning curve analysis and benchmark definition enabled the development of a curriculum for procedural training, based on data rather than supposition. This provided an evidence- and proficiency-based pathway for novice surgeons to follow.

All simulated procedures were recorded for quality assessment with a real-time video capture software: FRAPS (Beepa Pty Ltd, Melbourne, Australia). The assessment tool was the Objective Structured Assessment of Technical Skills (OSATS)-derived generic global rating scale [14, 15]. Due to difficulty in differentiating the 2 last

components of the scale on videos (“Flow of operation and forward planning” and “Knowledge of specific procedure”), they were evaluated as a single component giving a score out of 30 instead of 35. Only LH without guidance was assessed. The following procedures were arbitrarily chosen for analysis: n° 1 and 2 for experienced and intermediate participants, n° 1, 2, 4, 6, 8 and 10 for group A participants and n° 10 for group B participants. Two experienced laparoscopic surgeons (A.A and P.C) blinded to identity, group and trial number of each subject assessed independently the videos.

Statistical analysis

The sample size of at least eight subjects per group was based on a two-tailed test, with $\alpha = 0.05$ and power $(1 - \beta) = 0.80$, and an intended reduction of 30 per cent in time taken to complete tasks for experienced *versus* inexperienced operators, based on data from previous studies of VR simulation [16, 17]. Group A participants trained solely on the LH, and group B participants practiced on the LH with guidance for the 8 first repetitions. The intention was to make sure that inexperienced subjects did not switch from one VR module to the other, as this would have influenced the results of the study inappropriately. This was not an issue for the intermediate and experienced surgeons, as they performed no more than two repetitions of each task. Nonetheless, intermediate and experienced subjects were asked to perform the LH without guidance first, as starting with the guidance procedure could have given them an insight of what is considered correct by the VR simulator. Data were analyzed with SPSS version 20.0 (IBM Corp., Armonk, NY 2011, USA) using nonparametric tests. Comparison of performance between groups was made using the Kruskal–Wallis test

and Mann–Whitney U test, as appropriate. Learning curve data were analyzed by means of the Friedman test. Multiple comparisons were then made to identify when skills had plateaued. Inter-rater reliability for qualitative assessment (OSATS scale) was calculated using the intraclass correlation coefficient (ICC). Results were reported as median values (except for organ injuries reported as mean values), and a level of $p < 0.05$ was considered statistically significant. Ethical approval was obtained from the Institutional Review Board of the French College of Obstetricians and Gynecologists (CEROG 2015-GYN-0504).

Results

Demographics and surgical experience are presented in Table 1. Withdrawal before the end of learning curve analysis occurred for 6 inexperienced participants due to time constraints. A total of 34 participants completed the simulator sessions as initially planned.

Baselines

“Clipping and grasping” task performances were significantly different between experienced, intermediate, group A and group B subjects in terms of time (84, 109, 167 and 165 s, respectively; $p < 0.001$).

LH metrics

LH with guidance revealed significant differences at first repetition between experienced, intermediate and inexperienced group B participants for time (1100, 1818 and 2662 s, respectively; $p < 0.001$), path length (2267, 3078 and 3846 cm; $p = 0.004$), number of movements (1275, 2128 and 2880; $p < 0.001$) and idle time (314, 562 and 801 s; $p < 0.001$). No significant difference was found in terms of organ injuries.

LH without guidance at first repetition revealed significant differences between experienced, intermediate and

inexperienced group A participants for time (1369, 2385 and 3370 s, respectively; $p < 0.001$), path length (3390, 4526 and 5749 cm; $p = 0.002$), number of movements (2033, 3195 and 4056; $p = 0.001$), idle time (357, 654 and 747 s; $p = 0.001$) and respect for tissue (24, 40 and 84; $p = 0.01$). A significant difference was also found in terms of number of bladder injuries (0.13, 0 and 4.27; $p < 0.001$) and ureter injuries (0, 0 and 0.55; $p = 0.018$). No significant difference was found in terms of recorded injuries of other organs.

Performance comparison between experienced and intermediate at second session and inexperienced groups at tenth session revealed a significant difference only for time (937, 1614, 1268 and 1279 s, respectively; $p = 0.012$).

All metrics that demonstrated validity and learning curves analysis are presented in Table 2. Figure 3 illustrates the outcomes regarding “total number of movements.”

Quality

Video-based inter-rater reliability

Video-based assessment of a random sample ($n = 20$) of all eligible procedures was undertaken by the two independent raters. Inter-rater reliability coefficient for OSATS total score was $ICC = 0.848$ ($p < 0.001$). Inter-rater reliability coefficients for individual components of the OSATS scale ranged from $ICC = 0.674$ to 0.870 ($p < 0.01$). Given the high inter-rater reliability obtained, the rest of procedures were blindly assessed by one of the two raters (PC) and are presented as results.

LH

There was a significant difference in OSATS scores between experienced, intermediate and inexperienced group A participants at first repetition (22, 15 and 8, respectively; $p < 0.001$) and second repetition (25.5, 19.5 and 14; $p < 0.001$). The between-group analysis revealed

Table 1 Demographics and surgical experience of participants in the operating room

| | Experienced group (senior surgeons) | Intermediate group (5th year surgical residents) | Inexperienced group (5th year medical students) |
|-----------------------------|-------------------------------------|--|---|
| No. of participants | 8 | 8 | 24 |
| Age (years) | 41.5 (35–52) | 30 (28–31) | 24.5 (23–36) |
| Gender (female, male) | 2, 6 | 7, 1 | 14, 10 |
| Full TLH as primary surgeon | 100+ | 5 (2–10) | NA |
| TLH as first assistant | 100+ | 29 (25–60) | 4 (3–5) |

Data are as n or median (range)

TLH total laparoscopic hysterectomy, NA not applicable

Table 2 Metrics for development of the training curriculum

| Modules | Metrics | Construct valid | <i>p</i> Learning curve* | Plateau session | Benchmark level |
|-------------------------------|---------------------------------|-----------------|--------------------------|-----------------|-----------------|
| LH ¹ with guidance | Time (s) | ✓ | <0.001 | 6th | 1100 |
| | Total path length (cm) | ✓ | <0.001 | 4th | 2267 |
| | Total number of movements | ✓ | <0.001 | 5th | 1275 |
| | Idle time ¹ (s) | ✓ | <0.001 | 5th | 314 |
| LH ¹ | Time (s) | ✓ | <0.001 | 5th | 937 |
| | Total path length (cm) | ✓ | <0.001 | 5th | 1853 |
| | Total number of movements | ✓ | <0.001 | 5th | 1256 |
| | Idle time ² | ✓ | <0.001 | 5th | 244 |
| | Respect for tissue ³ | ✓ | 0.008 | 2nd | 16** |
| | Bladder injuries | ✓ | 0.016 | 6th | 0 |
| | Ureter injuries | ✓ | 0.135 | – | 0 |

¹ Laparoscopic hysterectomy

² Idle “non-productive” time: total time the movements of the instruments do not make tissue contact

³ Respect for tissue: number of occurrences unnecessarily force was applied by instruments on the bladder colon right and left ureters right and left arteries

* The *p* values indicating significant learning curves are italicized

** Metrics not included in the curriculum

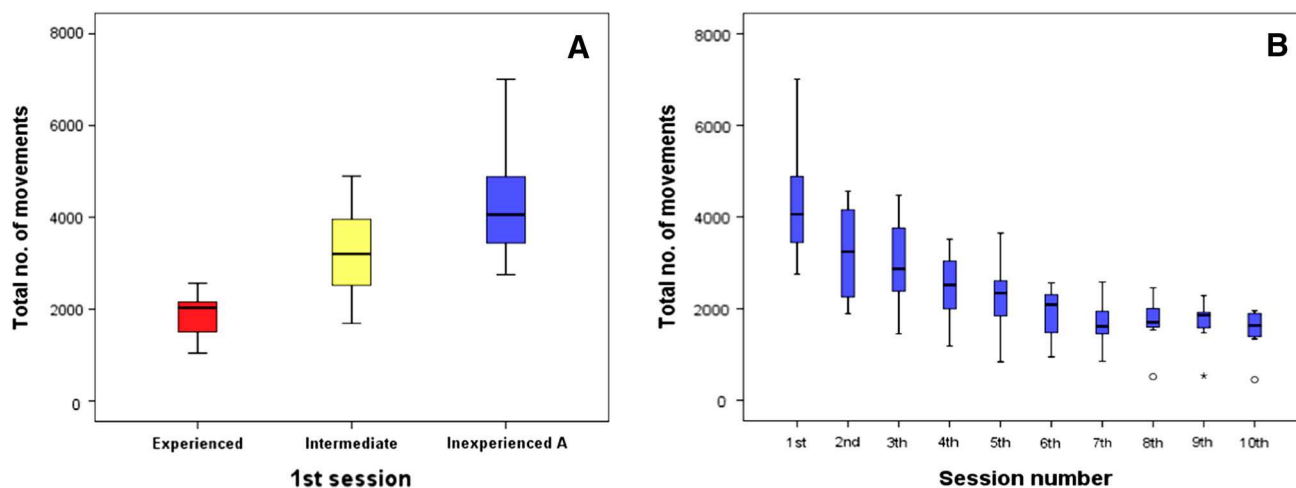


Fig. 3 Total number of movements for laparoscopic hysterectomy without guidance: **A** Outcomes at first session ($p = 0.001$; Kruskal–Wallis test). **B** Learning curve for inexperienced group A participants ($p < 0.001$; Friedman test; plateau at 5th session). Horizontal lines

within boxes, boxes and whiskers represent median, inter-quartile range and range, respectively. Circles and asterisks represent outliers and extreme cases, respectively

significant differences between inexperienced and intermediates, inexperienced and experienced, and intermediate and experienced groups (Mann–Whitney test; $p < 0.05$).

A comparison between experienced and intermediates scores at second session and inexperienced scores (groups A and B) at tenth session did not reveal any significant difference (25.5, 19.5, 21 and 22, respectively; $p = 0.07$). Outcomes regarding OSATS scores (validity and learning curve analysis) are presented in Fig. 4.

Curriculum construction

The results presented in Table 2 and Fig. 4 were used for the development of a training curriculum. Basic tasks proficiencies retained were published previously [6]. Respect for tissue was excluded as learning curves plateaued at second session, and it does not add to the measurement of proficiency. Qualitative OSATS assessment is undertaken once metric proficiency is achieved, in order to

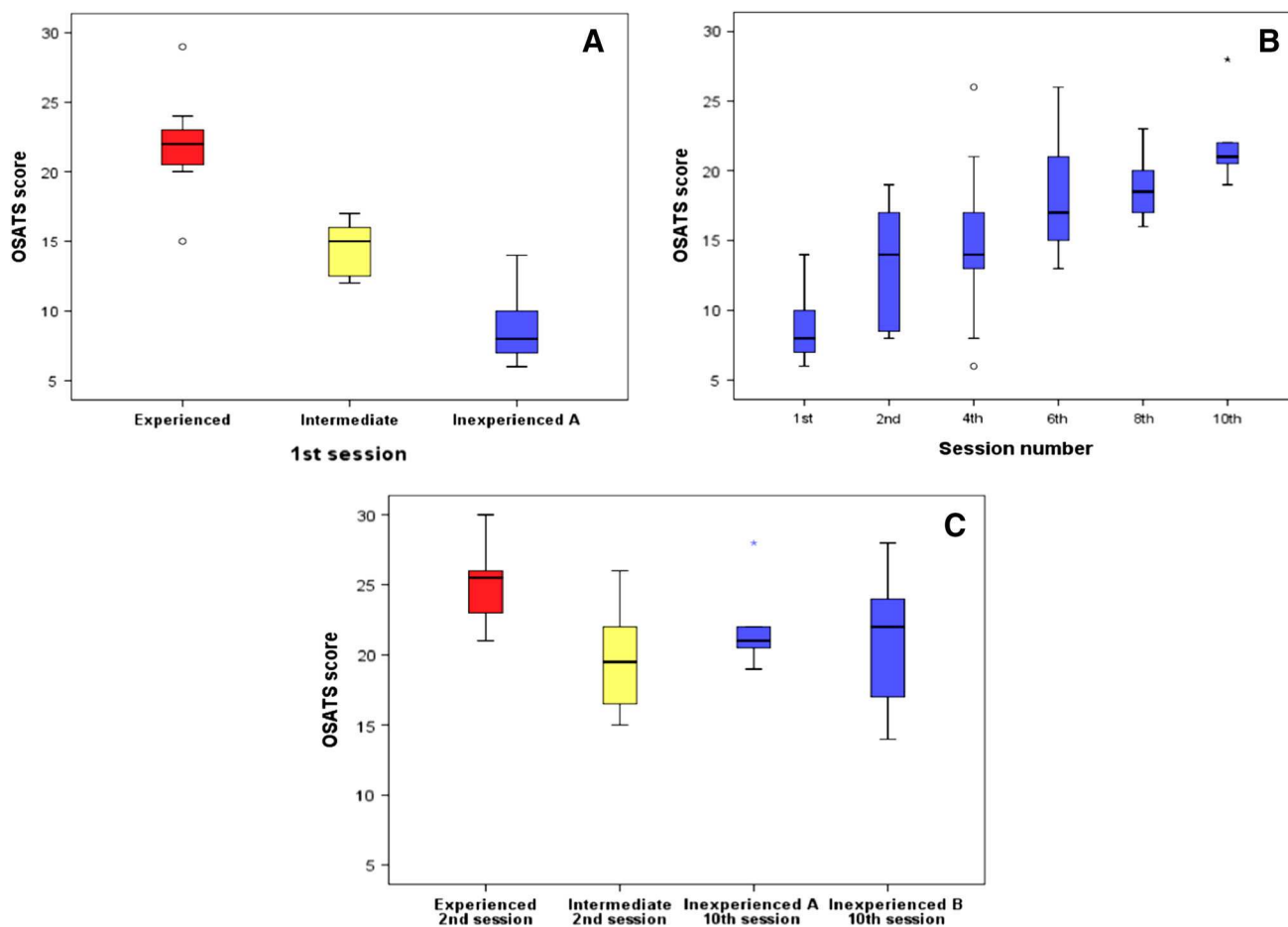


Fig. 4 OSATS global rating scores for laparoscopic hysterectomy without guidance: **A** Outcomes at first session ($p < 0.001$; Kruskal–Wallis test). **B** Learning curve for inexperienced group A participants ($p = 0.034$; Friedman test; no plateau). **C** Outcomes at 2nd session for experienced and intermediate groups, and at 10th session for

inexperienced groups (A and B) ($p = 0.07$, Kruskal–Wallis test). Horizontal lines within boxes, boxes and whiskers represent median, inter-quartile range and range, respectively. Circles and asterisks represent outliers and extreme cases, respectively

minimize assessment time required from a senior rater. The summarized outcome is a proficiency-based VR curriculum for training in LH (Fig. 5).

Discussion

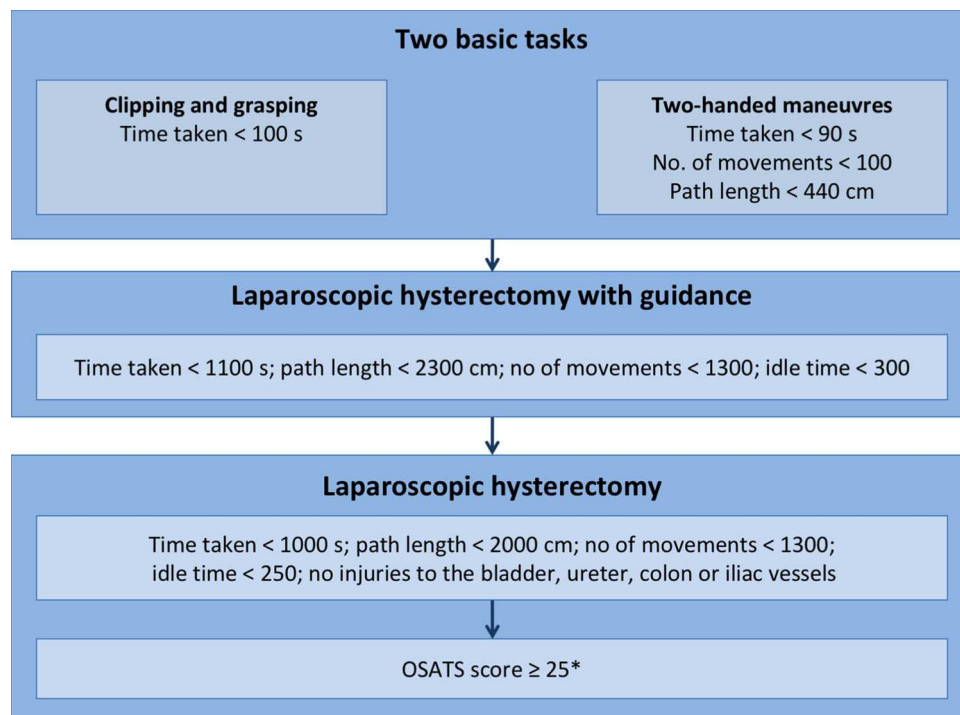
This study presents the development of a proficiency-based curriculum for laparoscopic hysterectomy on a VR simulator. Early novices are most likely to benefit from this curriculum, which is based on a VR program reproducing a simple case. Offered to trainees during their first years of residency, this training program should allow them to acquire the status of a pretrained novice who can operate with greater skills on patients undergoing LH procedure.

The LH program displayed validity evidence on number of the automatically recorded parameters. In addition, OSATS score analysis revealed a high degree of construct validity with regard to quality. Performances of the three

different groups were distributed as expected according to their level of experience. Learning curve analysis of the inexperienced group supported the effectiveness of the program on the curriculum's quantitative metrics with a plateau occurring at 4th–6th repetition. Interestingly, quality scores improved throughout the ten sessions with no plateau reached, suggesting that the simulator enables trainees to further refine qualitative skills with extra sessions. OSATS scores obtained by inexperienced groups at 10th repetition were not significantly different from those of the experienced and intermediates, although it can be pointed out that the absolute value of median scores for the inexperienced groups leveled with intermediates but remained lower than experienced group.

This study has limitations. The suturing of the vaginal cuff is not included in this VR program and consequently could not be integrated into the present curriculum. Cuff closure low-fidelity models have been reported [18, 19], revealing a good validity. Further research is needed to

Fig. 5 Evidence-based virtual reality training curriculum for laparoscopic hysterectomy. *Asterisk* video-based assessment by an experienced rater. All tasks should be performed for a maximum of two sessions a day, each session >1 h apart; completion of training at proficiency (achieved on 2 consecutive sessions)



comparatively assess available cuff closure models, before integration in a whole-procedure LH curriculum. The second limitation relates to the ability of this VR program to reproduce all aspects of real operative conditions: This curriculum was developed using a standardized case presenting with a small uterus. Thus, the operators' ability to react to anatomical distortion or challenging exposure was not assessed. Thirdly, groups were defined by experience; as experience does not necessarily correlate with expertise, this could have undermined the "quality" of the benchmark levels. The intermediate group was deliberately composed of advanced residents with moderate experience as first operator in the OR, given the relatively simple presentation of the case. Their reported level of experience for this procedure was consistent with the findings of a recent survey addressed to French gynecologic residents [20]. Finally, it is important to underline that costs associated with VR simulators are significant and could limit the widespread use of the proposed curriculum.

Two categories of automatically recorded parameters were found valid: one purely quantitative (ex: time taken) and another closely linked to surgical quality, like bladder injuries or idle time. Bladder injury is the most frequent intraoperative complication during the learning curve on real patients [21]. Other organ injuries occurred less frequently and were not statistically different between the groups (data not presented). These encouraging findings may result from advance in technology that has enabled the development of a complex high-fidelity program closer to real operating situations, although some anatomical

features are lacking such as uterine veins. To date, previously published curricula on VR simulators validated purely quantitative parameters [6]. In the case of VR LH, the control of the uterine manipulator that is a crucial component of this procedure for exposure and planning process probably adds to the ability of the program to measure quality. This assumption is supported by results on OSATS scores. Ultimately, a step forward in VR simulation would allow the integration of components of OSATS scale to be automatically measured by the simulator.

Regarding the impact of the LH with guidance, it can be noticed that quality of operative skills obtained after the proposed training was not different between group A and group B. However, the interquartile range of OSATS score seemed surprisingly wider in the group trained with the guidance program. A hypothesis is that the LH program without guidance was difficult for trainees, so all of them had an incentive to improve because they realized early their lack of knowledge through their mistakes. On the contrary, participants in the guidance group improved very differently depending on their personal profile: The guidance signs could have been a way to refine skills for some, leading to improved quality, or a too secure situation for others that did not engaged in an active learning process and consequently had lower quality scores when tested without guidance. The guidance mode was included at the beginning of the curriculum, as it provides an opportunity for the trainees to follow the safety principles of this procedure.

A proficiency-based training program on a VR simulator was developed in gynecological laparoscopy for

salpingectomies [5], demonstrating a positive impact on operative skills when transferred to the OR [22]. The present work is the first to focus on LH. This complex laparoscopic procedure requires a range of surgical competencies, both technical and nontechnical [3] that are tested with the simulator's program. Subjects control the camera, the uterine manipulator and a second assisting instrument. These simultaneous actions highlight the operator's skills in terms of knowledge of the procedure, forward planning and decision making. The proficiency time required to complete a LH procedure on the VR simulator (16 min) seems shorter than operative time reported in clinical studies (80–90 min) [21]. However, this VR module does not include port insertion, vaginal cuff and skin closure. Thus, comparison with average OR operative time reported in the literature should be interpreted with caution.

This study demonstrates that it is possible to build a proficiency-based curriculum for VR laparoscopic hysterectomies with specific quantitative and qualitative goals to be achieved. Further work, with a multicenter setting, should confirm its value on a larger scale. Future refinement through interventions, like deliberate practice or mental practice [23–25], shall also be tested to optimize learning curve. A next step will be to evaluate whether this VR training program leads to improved operative skills when transferred to the OR and ultimately improves patient's safety.

Acknowledgments The authors thank Marine Alessandrini for her contribution to this project (Service de santé publique et d'information médicale, centre hospitalier universitaire Timone, 13385 Marseille, Aix Marseille université, France).

Funding CERC (Centre d'Enseignement et de Recherche Chirurgicale), Aix Marseille Université, Assistance Publique - Hôpitaux de Marseille. Rajesh Aggarwal was funded by "Fonds de recherche santé Québec" and "Montréal General Hospital Foundation."

Compliance with ethical standards

Disclosure Rajesh Aggarwal is a consultant for Applied Medical, Patrice Crochet, Sophie Knight, Stéphane Berdah, Léon Boubli and Aubert Agostini have no conflicts of interest or financial ties to disclose.

References

- Kohn LT, Corrigan JM, Donaldson MS (2000) To err is human: building a safer health system. Institute of Medicine (US) Committee on Quality of Health Care in America. National Academies Press (US). <http://www.ncbi.nlm.nih.gov/books/NBK225182/>
- Zendejas B, Brydges R, Hamstra SJ, Cook DA (2013) State of the evidence on simulation-based training for laparoscopic surgery: a systematic review. *Ann Surg* 257(4):586–593
- Dedy NJ, Bonrath EM, Zevin B, Grantcharov TP (2013) Teaching nontechnical skills in surgical residency: a systematic review of current approaches and outcomes. *Surgery* 154(5):1000–1008
- Fraser SA, Klassen DR, Feldman LS, Ghitulescu GA, Stanbridge D, Fried GM (2003) Evaluating laparoscopic skills: setting the pass/fail score for the MISTELS system. *Surg Endosc* 17(6):964–967
- Aggarwal R, Tully A, Grantcharov T, Larsen CR, Miskry T, Farthing A, Darzi A (2006) Virtual reality simulation training can improve technical skills during laparoscopic salpingectomy for ectopic pregnancy. *BJOG* 113(12):1382–1387
- Aggarwal R, Crochet P, Dias A, Misra A, Ziprin P, Darzi A (2009) Development of a virtual reality training curriculum for laparoscopic cholecystectomy. *Br J Surg* 96(9):1086–1093
- Stefanidis D, Sevdalis N, Paige J, Zevin B, Aggarwal R, Grantcharov T, Jones DB (2015) Simulation in surgery: what's needed next? *Ann Surg* 261(5):846–853
- Igwe E, Hernandez E, Rose S, Uppal S (2014) Resident participation in laparoscopic hysterectomy: impact of trainee involvement on operative times and surgical outcomes. *Am J Obstet Gynecol* 211(5):484.e1–484.e7
- Reich H (2007) Total laparoscopic hysterectomy: indications, techniques and outcomes. *Curr Opin Obstet Gynecol* 19(4):337–344
- Donnez O, Jadoul P, Squifflet J, Donnez J (2009) A series of 3190 laparoscopic hysterectomies for benign disease from 1990 to 2006: evaluation of complications compared with vaginal and abdominal procedures. *BJOG* 116(4):492–500
- Fundamentals of Laparoscopic Surgery—the definitive laparoscopic skills enhancement and assessment module. <http://www.flsprogram.org/>
- Joint Committee on Standards for Educational and Psychological Testing APA, American Educational Research Association and National Council on Measurement in Education (1999) Standards for educational and psychological testing. Washington, DC
- Ahmed K, Miskovic D, Darzi A, Athanasiou T, Hanna GB (2011) Observational tools for assessment of procedural skills: a systematic review. *Am J Surg* 202(4):469–480
- Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, Brown M (1997) Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg* 84(2):273–278
- Aggarwal R, Grantcharov T, Moorthy K, Milland T, Darzi A (2008) Toward feasible, valid, and reliable video-based assessments of technical surgical skills in the operating room. *Ann Surg* 247(2):372–379
- Aggarwal R, Grantcharov T, Moorthy K, Hance J, Darzi A (2006) A competency-based virtual reality training curriculum for the acquisition of laparoscopic psychomotor skill. *Am J Surg* 191(1):128–133
- Aggarwal R, Grantcharov TP, Eriksen JR, Blirup D, Kristiansen VB, Funch-Jensen P, Darzi A (2006) An evidence-based virtual reality training program for novice laparoscopic surgeons. *Ann Surg* 244(2):310–314
- Tunitsky-Bitton E, Propst K, Muffly T (2016) Development and validation of a laparoscopic hysterectomy cuff closure simulation model for surgical training. *Am J Obstet Gynecol* 214(3):392.e1–392.e6
- King CR, Donnellan N, Guido R, Ecker A, Althouse AD, Mansuria S (2015) Development and validation of a laparoscopic simulation model for suturing the vaginal cuff. *Obstet Gynecol* 126(Suppl 4):27S–35S
- Crochet P, Aggarwal R, Berdah S, Yaribakht S, Boubli L, Gannerre M, Agostini A (2014) Current and future use of surgical skills simulation in gynecologic resident education: a French

- national survey. *J Gynecol Obstet Biol Reprod (Paris)* 43(5): 379–386 (in French)
21. Wattiez A, Soriano D, Cohen SB, Nervo P, Canis M, Botchorishvili R, Mage G, Pouly JL, Mille P, Bruhat MA (2002) The learning curve of total laparoscopic hysterectomy: comparative analysis of 1647 cases. *J Am Assoc Gynecol Laparosc* 9(3):339–345
 22. Larsen CR, Soerensen JL, Grantcharov TP, Dalsgaard T, Schouenborg L, Ottosen C, Schroeder TV, Ottesen BS (2009) Effect of virtual reality training on laparoscopic surgery: randomised controlled trial. *BMJ* 338:b1802
 23. Crochet P, Aggarwal R, Dubb SS, Ziprin P, Rajaretnam N, Grantcharov T, Ericsson KA, Darzi A (2011) Deliberate practice on a virtual reality laparoscopic simulator enhances the quality of surgical technical skills. *Ann Surg* 253(6):1216–1222
 24. Palter VN, Grantcharov TP (2014) Individualized deliberate practice on a virtual reality simulator improves technical performance of surgical novices in the operating room: a randomized controlled trial. *Ann Surg* 259(3):443–448
 25. Arora S, Aggarwal R, Moran A, Sirimanna P, Crochet P, Darzi A, Kneebone R, Sevdalis N (2011) Mental practice: effective stress management training for novice surgeons. *J Am Coll Surg* 212(2):225–233

6. DÉVELOPPEMENT D'UNE ECHELLE D'ÉVALUATION SPÉCIFIQUE DE L'HC :

H-OSATS

6.1 Présentation

Development of an objective assessment tool for total laparoscopic hysterectomy: a Delphi method among experts and evaluation on a virtual reality simulator. Knight S, Aggarwal R, Agostini A, Loundou A, Berdah S, Crochet P.

Ce travail a été présenté en communication orale lors du congrès de l' European Society of Gynecologic Endoscopy, à Bruxelles le 3 Octobre 2016. Il est soumis pour publication au journal PloS one.

L'objectif principal de ce travail était de construire une échelle d'évaluation spécifique de l'HC (H-OSATS) et de tester sa validité sur un simulateur de RV.

- Une analyse séquentielle de l'HC a été réalisée à partir de vidéos opératoires.
- Les étapes à inclure dans l'échelle pour l'évaluation des compétences techniques ont été sélectionnées par un groupe d'experts internationaux grâce à la méthode Delphi.
- L'échelle créée a été évaluée lors d'une HC sur le simulateur de RV (LapMentor*), par l'analyse de la performance de 3 groupes d'opérateurs d'expérience clinique différente (externes inexpérimentés, internes intermédiaires, seniors expérimentés).

Voici les principaux résultats de cette étude :

- 64 items ont été retenus par le groupe Delphi. Ils sont présentés dans l'échelle en 14 temps opératoires qui comprennent l'installation et l'abord, le temps opératoire intra-péritonéal et la fermeture.

- Le caractère détaillé de l'échelle est compatible avec une utilisation dans le cadre de la formation et comme instrument de feedback.

Les questions posées par cette étude et ses limites étaient :

- La validité et la faisabilité d'H-OSATS doivent être confirmés pour une HC complète au BO.
- L'utilisation de cette échelle comme outil de pratique délibérée doit être testée.
- Cette échelle peut-elle être utilisée pour l'évaluation dans le cadre de la certification ?

6.2 Article soumis

Development of an objective assessment tool for total laparoscopic hysterectomy:
a Delphi method among experts and evaluation on a virtual reality simulator.

Short running title:

Laparoscopic hysterectomy assessment scale

Authors:

1. Sophie Knight, MD, Department of Obstetrics and Gynecology, Assistance Publique - Hôpitaux de Marseille, La Conception Hospital, 13005 Marseille, Aix Marseille Université, France.
2. Rajesh Aggarwal, PhD, FRCS, Arnold and Blema Steinberg Medical Simulation Centre, Department of Surgery, McGill University, Montreal, Canada.
3. Aubert Agostini, MD, Department of Obstetrics and Gynecology, Assistance Publique - Hôpitaux de Marseille, La Conception Hospital, 13005 Marseille, Aix Marseille Université, France.
4. Anderson Loundou, PhD, Public Health Laboratory, Aix-Marseille University, Marseille, France.
5. Stéphane Berdah, MD, PhD, Aix Marseille Université, CERC, IFSTTAR, LBA UMR_T 24, 13916, Marseille, France.
6. Patrice Crochet, MD, Department of Obstetrics and Gynecology, Assistance Publique - Hôpitaux de Marseille, La Conception Hospital, 13005 Marseille, Aix Marseille Université, France.

Corresponding author:

Address: 147 boulevard Baille, 13005 Marseille

Tel: 00 33 688384803

Email: pcrochet.marseille@gmail.com

Sources of funding:

CERC (Centre d'Enseignement et de Recherche Chirurgicale), Aix Marseille Université, Assistance Publique - Hôpitaux de Marseille.

Rajesh aggarwal is funded by « Fonds de recherche santé Québec » and « Montréal General Hospital Foundation ».

Abstract:

Background: Laparoscopic hysterectomy (LH) requires an advanced level of operative skills and training. The aim of this study was to develop an objective scale specific for the assessment of technical skills for LH (H-OSATS) and to demonstrate feasibility of use and validity in a virtual reality (VR) setting.

Methods:

The scale was developed using a hierarchical task analysis (HTA) and a panel of international experts. A Delphi method obtained consensus among experts on relevant steps that should be included into the H-OSATS scale for assessment of operative performances. Feasibility of use and validity of the scale were evaluated by reviewing video recordings of LH performed on a virtual reality (VR) laparoscopic simulator that reproduces key specific components of a LH procedure until colpotomy. Three groups of operators of different levels of experience were assessed (10 novices, 8 intermediates and 8 experienced surgeons).

Results:

A total of 76 discrete steps were identified by the HTA. 14 experts completed the first and second round of the Delphi questionnaire. 64 steps reached consensus after the second round and were integrated in the scale. During the validation process, median time to rate each video recording was 25 minutes. There was a significant difference between the novice, intermediate and experienced group for total H-OSATS scores (133, 155.9 and 178.25 respectively; $p=0.002$). H-OSATS scale demonstrated high inter-rater reliability (intraclass correlation coefficient [ICC] =0.930; $p<0.001$) and test retest reliability (ICC=0.877; $p<0.001$). High correlations were found between total H-OSATS scores and total OSATS scores ($\rho=0.928$; $p<0.001$).

Conclusion:

The H-OSATS scale is a feasible and valid instrument for assessment of technical performances for LH performed on a VR simulator. The implementation of this scale is expected to facilitate deliberate practice. Next steps should focus on validating the scale in the operating room.

Keywords: Laparoscopic hysterectomy. Assessment scale. Technical skills. Consensus. Expert

Introduction:

Hysterectomy is the second most frequently performed surgical procedure on women, after cesarean section [1]. More than 300 000 inpatient hysterectomies were performed in the US in 2012 [2]. Surgical routes were distributed as follows: abdominal (52.8%), vaginal (14.7%) and laparoscopic (32.4%). Two fifths of laparoscopic hysterectomies performed for benign indications were robotically assisted. The majority of hysterectomies were performed abdominally, though benefits of the vaginal and laparoscopic routes have largely been demonstrated in terms of speedier return to normal activities, lower intraoperative blood loss, reduced wound infections and enhanced cosmetic results [3]. For benign indications, the Cochrane database recommends to favor the vaginal approach over the abdominal approach, and to attempt laparoscopic hysterectomy (LH) when vaginal hysterectomy is not possible [3]. LH has notable advantages over vaginal hysterectomy, such as allowing an optimal exploration of the abdominal cavity. This aspect is particularly interesting in oncological gynecology and for certain benign indications. Furthermore the laparoscopic route offers the possibility to perform additional procedures such as lymphadenectomy. Thus, LH has become the procedure of choice in surgical oncology for treatment of endometrial cancer [4, 5].

Although Harry Reich performed the first LH more than 25 years ago [6], this approach is not fully exploited in current gynecologic practice. It has often been suggested that this could be attributed to difficulty to train operators to this procedure [7]. LH requires an advanced level of surgical skills, with a learning curve estimated between 30 and 75 cases depending on the surgeon's laparoscopic experience [8, 9]. There is no clear definition of proficiency level which is currently based on the number of cases performed as a primary surgeon and on the subjective opinion of a senior preceptor.

This method is known to be imprecise and unreliable [10]. Improved training and competence assessment should allow optimal incorporation of this technique into the surgical armamentarium.

Evaluation of technical performance requires objective measurement tools. Generic global rating scales are well validated [11, 12] but do not provide trainees with information on the specific parts of the procedure that require improvement. For this purpose, procedure specific scales have been developed for advanced laparoscopic procedures [13, 14, 15, 16]. These procedure-specific scales have number of potential applications including assessment of trainee's operative skills and validation of competences. They can be used to facilitate constructive feedback and deliberate practice.

The objective of the study was to develop an objective scale for assessment of technical skills for LH (H-OSATS) and to demonstrate feasibility and validity in a virtual reality (VR) setting.

Materials and methods:

- Hierarchical task analysis (HTA)

A HTA was conducted in order to deconstruct the procedure into its component steps [17]. The purpose was to identify the successive discrete steps that are required to complete a LH. This process can be carried out including the intervention of other members of the team (anaesthetist, scrub nurse). Because the objective of the scale was to focus on surgical performances, it was chosen to select steps that referred to the surgeon only.

Literature was searched in order to identify the different operative techniques [18, 19, 20, 21, 22, 23, 24]. A panel of video recordings of extrafascial LH performed by expert

laparoscopic gynecologists illustrating these techniques were selected from 2 online databases (AAGL, WeB Surg). The aim was to develop a scale that could assess operative performances regardless of the different approaches to the procedure. Two experienced laparoscopic surgeons (AA and PC) reviewed independently the videos. Each reviewer listed all consecutive discrete steps required for completion of each operation. An in-person meeting was organised in order to pool together the results and elaborate a joint list of steps.

- Creation the H-OSATS scale

The list of consecutive steps generated by the HTA was submitted to a panel of international expert laparoscopic gynecologists using a Delphi method. The Delphi method is a systematic and interactive forecasting method used to obtain consensus among a panel of experts, who are consulted over several rounds. After each round answers are collected, analysed and submitted back in an iterative fashion to the group. Over the successive rounds, group opinion should converge towards consensus [25, 26]. Eligible experts were identified based the following criteria: having prior publications on LH; being key opinion leaders in the field of gynecologic laparoscopy; having active involvement among international endoscopic societies. They were invited to participate in the project via email or at the occasion of an international congress. Experts from different geographic zones were recruited (US, Canada, Europe and Australia) in order to develop an internationally relevant scale. Two rounds of the Delphi survey were conducted via an online questionnaire (docs.google.com). Experts were asked to rate, using a Likert scale from 1 “strongly disagree” to 5 “strongly agree”, each discrete step based on its relevance for assessment of operative skills. They were invited to comment their answers in order to modify or add steps during the second round. Mean and standard deviation obtained

for each steps during the first round were presented to the experts during the second round. A total of three email reminders were sent during both rounds.

A rate of agreement (RoA) was calculated as a measure of consensus among the experts: $[(\text{Agreement} - \text{Disagreement}) / (\text{Agreement} + \text{Disagreement} + \text{Indifferent})] \times 100$

A RoA greater than 70% was chosen as a measure of consensus [27, 28] . In case of missing data, RoA was calculated by replacing the missing answer by 3, mean and mode. Steps that reached consensus during the second round were included into the final scale. Finally, a numerical scoring scale ranging from 1 to 5 was assigned to each selected step and an arbitrary description of the attribution of points was established [16]. Steps that are performed bilaterally were to be rated for each side.

- Determination of feasibility of use and validity evidence of the H-OSATS scale

This part of the study used a single blinded observational study design. Video recordings of LH performed on a VR simulator (LAP Mentor™ VR; Simbionix-3D Systems, Cleveland, Ohio, USA) were collected. The simulator provides with a LH program that begins after trocar insertion and once the uterine manipulator is in place, and ends after circumferential colpotomy. Patient positioning, trocar insertion, vaginal vault and skin closure could therefore not be assessed. This program displayed good validity evidence during a previous study according to quantitative and qualitative parameters [29].

LH performed by three groups of operators of different levels of experience were evaluated: novices, intermediates (had performed 2-10 LH) and experienced (had performed over 100 LH). The group of novices comprised 5th year medical students. Two LH were performed in each group: it was chosen to assess the second LH. Each video was recorded using FRAPS real time video capture software (Beepa Pty Ltd,

Melbourne, Victoria, Australia). Videos were anonymized and assessed independently by two trained raters (PC and SK) using the H-OSATS scale and the OSATS scale. Due to difficulty in differentiating the 2 last components of the OSATS scale on videos (“Flow of operation and forward planning” and “Knowledge of specific procedure”), they were evaluated as a single component giving a score out of 30 instead of 35 [11, 30].

➤ Feasibility

Raters were asked to record the time required to score LH performed on the simulator, using the H-OSATS scale.

➤ Validity evidence

The contemporary meaning of validity is a unitary concept with multiple aspects that considers construct validity as the whole of validity [31]. The following aspects of construct validity were evaluated: Evidence of content validity was provided by comparing scores between the 3 groups of different levels of experience. Total scores and component scores from “round ligament” to “colpotomy” were analysed. Evidence of internal structure was provided by evaluating reliability of the scores obtained: Inter-rater reliability was assessed by correlating total and component scores for the H-OSATS scale obtained between 2 independent raters using the intraclass correlation coefficient (ICC: 2-way mixed-effects model, absolute agreement). Test-retest reliability was assessed by correlating scores for the 10 first videos assessed at two different time points by the same rater, using the ICC. Evidence to support relationship to other variables was provided by correlating total scores obtained with the H-OSATS scale with those obtained with the OSATS scale.

- Statistics:

Data was analysed using SPSS version 20.0 (SPSS, Chicago, Illinois, USA). Results were reported as median values and a level of $p < 0.05$ was considered statistically significant. Comparison of scores was undertaken using non-parametric tests, Kruskal-Wallis and Mann-Whitney U tests. Correlation between both scales were analysed using Spearman's coefficient. Ethical approval was obtained from the Institutional Review Board of the French College of Obstetricians and Gynecologists (CEROG 2015-GYN-0801).

Results:

- Hierarchical task analysis

Eight video recordings of LH performed by expert laparoscopic gynecologists were reviewed and a total of 69 consecutive steps were identified. As the videos collected recorded the intra-abdominal camera view, certain aspects of the full procedure were not visualized. Therefore the reviewers listed steps evaluating patient positioning, abdominal access and skin closure without video support.

This list of steps was distributed into 14 tasks: from "patient positioning" (task 1) to "port removal and skin closure" (task 14). Seven additional points on the order in which tasks should be performed were addressed and added to the list of steps (**table 1**).

- Delphi process

The 76 steps were submitted to a panel of experts. 14 of the 17 experts who accepted to participate in the project completed the first round of the Delphi process. The composition of the panel of experts is detailed in **table 2**. A total of 20/76 steps did not reach the predefined level of consensus. 22 discrete steps were reformulated based on comments made by the experts, 2 steps were added and 2 were deleted and a second questionnaire was developed. This questionnaire was submitted to the 14

experts who had completed the first round and all of them completed the second round. The Delphi survey was conducted between April and November 2015. 12/76 steps did not reach consensus level during the second round. Results are detailed in **table 1**.

- H-OSATS scale

A total of 64/76 steps were selected for inclusion into the final H-OSATS scale. Thus H-OSATS maximum score for the full scale was 370 points. Steps that are not applicable in most cases, i.e. “If anatomy is distorted: Identify the ureter prior to division of the uterine vessels” were assigned a negative score (**Appendix**).

- Determination of feasibility of use and validity evidence of the H-OSATS scale

A total of 26 video recordings were reviewed and scored: 10 LH performed by novices, 8 by intermediates and 8 by experienced operators. Steps evaluated on the VR program ranged from “round ligament division” to “colpotomy”, except for division of cardinal ligaments which cannot be distinguished from the vaginal cul-de-sac on the VR program, giving a total score out of 210 points.

- Feasibility:

Median operative time for the novice, intermediate and experienced group were 34, 27 and 16 minutes respectively ($p < 0.001$) Median time to rate each video recording was 25 minutes (37, 32 and 20 minutes for the novice, intermediate and experienced group respectively).

- Validity evidence:

There was a significant difference between the novice, intermediate and experienced group for total H-OSATS scores (133, 155.9 and 178.25 respectively; $p = 0.002$). Significant differences in scores were observed for the following components of the scale: task 7 (division of the IP ligament), task 8 (creation of the bladder flap), task 9 (opening the posterior peritoneum), task 10 (division of uterine vessels) and task

11 (colpotomy) (**table 3**). Regarding task 8 (creation of the bladder flap), scores for the intermediate group were, in absolute value, higher than those of the experienced group. A post hoc analysis found a significant difference between novices and intermediates and between novices and experts ($p=0.026$ and $p=0.006$ respectively). There was a significant difference between the novice, intermediate and experienced group for total OSATS scores (15, 21.75 and 26 respectively; $p<0.001$). Significant differences in scores were observed for each individual component of the OSATS scale (data not presented).

Inter-rater reliability for total H-OSATS score was $ICC=0.930$ ($p<0.001$). Inter-rater reliability coefficients for individual tasks ranged from 0.717 ($p<0.001$) for task 9 (opening of the posterior peritoneum) to 0.940 ($p<0.001$) for task 7 (division of the infundibulo-pelvic ligaments). Test retest reliability for total H-OSATS scores was $ICC=0.877$ ($p<0.001$).

A correlation was found between H-OSATS and OSATS total scores ($\rho=0.923$; $p<0.001$) (**figure**).

Discussion:

This study enabled the development of a scale for objective assessment of operative performances for LH, using a systematic approach. The H-OSATS scale is composed of three main parts: patient positioning and exposure (task 1 to 5), the core part of the procedure (tasks 6 to 12) and closure (task 13 and 14). Feasibility of use and sources of evidence to support validity of the H-OSATS scale were demonstrated for the core part of the procedure using a high-fidelity VR laparoscopic simulator.

There are two types of procedure specific assessment tools available: scales and checklists; both these tools have the advantages of identifying specific areas of the

procedure that require improvement. Rating scales tend to be less rigid than checklists that oblige operators to perform a series of steps in a predefined order [32]. Moreover, rating scales are more reliable and appropriate than checklists for summative assessments [33]. The H-OSATS scale belongs to the procedure-specific assessment tool category. Although these scales are intended to be relatively flexible regarding the chronological sequence of steps, the notion of order in which the main tasks are to be performed was addressed during the Delphi survey. Unsurprisingly, a consensus was achieved on the order in which the first tasks of the procedure (“patient positioning” to “inspection of the pelvis”) should be performed. Regarding the core part the procedure, the only point for which a consensus was obtained was that “colpotomy” should not be started before the completion of “uterine vessels division”. Logically, experts agreed that the last tasks of the procedure should be performed successively in the same order (“uterine vessels division” to “port removal”).

The advantages of the use of a Delphi method in order to obtain consensus among experts are well described: the anonymous nature of the process prevents a dominant member of the group from influencing the group’s opinion. Furthermore the questionnaire is completed by email and does not require for the experts to physically meet, allowing members from different geographic zones to take part in the process, thus improving feasibility. The Delphi method has been used in number of studies in order to establish consensus among a group of experts but there is no clear recommendation on the most suitable method to measure consensus [34]. Von der Gracht et al. suggested that a RoA is an appropriate measure of consensus particularly when Likert scales are used [35]. The notable limitations of the Delphi method include the fact that the selection of questions submitted to the experts are in part controlled by the Delphi facilitators and that interest of experts can diminish with consecutive

rounds. A two round Delphi was conducted in order to limit the number of non-responders and to avoid forced consensus. Indeed repeated rounds may in some cases lead to an artificial consensus when a question is submitted repeatedly, and “dissensus”, i.e the absence of consensus may be as important as consensus [34].

Validity evidence refers to data collected in order to assign a meaningful interpretation of assessment scores. Validity requires multiple source of evidence to support or refute this interpretation [36]. Messick identifies five sources of evidence to support construct validity: content, response process, internal structure, relationship to other variables and consequences [37]. This study evaluated the H-OSATS scale using a VR laparoscopic simulator and various aspects of validity evidence were evaluated. The H-OSATS displayed good validity evidence with regards to the three tested sources of evidence (content validity, internal structure and relationship to other variables). The VR setting allowed evaluation of novices in safe conditions, with no intervention from a senior surgeon that could influence results; thus avoiding the inherent bias of operating room (OR) assessment of supervised inexperienced operators. The H-OSATS scale was construct valid for each individual task of the scale assessed by the VR simulator except for task 6 i.e. “division of the round ligaments”. This is probably due to the fact that this task does not require a high degree of technical expertise and is relatively easy to perform on the simulator. Operators were allocated to a group based on case volume criteria, although experience does not necessarily correlate with expertise. Comparison of OSATS scores between the three groups confirmed the correct allocation to each group.

Regarding feasibility, average time to assess each video was 25 minutes. This relatively short time may be longer for LH performed in the OR. Indeed operative time for LH performed on the simulator appears to be shorter than operative time reported

in clinical studies [8]. However the program reproduces a relatively simple case and it does not include port insertion and vaginal cuff suture.

Some steps that did not reach consensus illustrate the difficulty to standardize some parts of the procedure whose approach can differ depending on surgeons surgical habits but also on the anatomical presentation. For example two steps evaluating fenestration of the broad ligament did not reach consensus during the first round and it appeared from the experts comments that those who usually do not fenestrate the broad ligament rated these steps poorly. Experts were reminded during the second round that the objective was to select steps that best assess operative performances and that the scale should allow assessment of operators regardless of the chosen approach. Therefore, an optional assessment of this part of the procedure was included in the scale. Another issue concerned the uterine vessels division. The level at which the uterine vessels should be coagulated was left with a vague description, as this step often varies depending on the anatomy and operative findings, and reached a high consensus during the first round. However the formulation regarding the level at which the uterine artery should be sectioned was modified based on repeated expert' comments that encouraged a precise description. The formulation "in the ascending portion, at the level of the colpotomizer" reached consensus at the second round. "Division of the utero-sacral (US) ligaments" did not reach consensus among the experts. There was however a consensus regarding the need to include US ligaments into the vaginal cuff suture. This can be explained by the fact the insersion of these ligaments on the posterior vaginal cul-de-sac remains in most cases . However, the panel of experts considered important that when US ligaments are divided they should be sutured to the vaginal cuff to strengthen pelvic support. Due to the current debate on morcellation it was chosen not to address this matter, as the

purpose was to select steps that best assess operative quality, not to obtain consensus on a controversial subject [38].

The ability to optimize the role of assistants during laparoscopic procedures is an important component of surgical competency [22]. All steps relating to management of the uterine manipulator were included into the scale, confirming that it is a key element for optimal exposure and closely linked to operative safety [39]. Four other steps regarding exposure were selected. Thus, the use of assistant is considered an operative skill all along the successive parts of the procedure. This implies that the assistant should remain passive during assessment, acting only on instructions from the operators. This point appears particularly important when the assistant is a supervisor: indeed too much self-initiative from the assistant could lead to over-rating. The scale did not include assessment of the quality of camera holding because this was considered as a generic skill and not directly assessed by this procedure specific scale.

Two other procedure specific scales have been developed for LH. Tremblay et al. published a scale developed from a list of steps arbitrarily chosen with no preliminary HTA and included items that are usually used in generic global rating scales such as « use of assistant to facilitate exposure » [40]. There was no assessment of the validity of this scale, neither in a simulated setting nor in the operating room. Frederick et al. recently published an assessment scale for robotically-assisted laparoscopic hysterectomy [41]. Consensus was obtained between 5 experts, based on a list of arbitrarily chosen steps. This study included a validation process in the OR that found significant difference between experts, advanced beginners, and supervised novices. Median scores were surprisingly close (75.6 vs 71.3 vs 69, out of 80), due to the interference of OR supervision according to the authors. Finally,

Weizman et al. developed a checklist for laparoscopic suturing of the vaginal cuff using a boxtrainer model [42].

The very detailed nature of the H-OSATS scale makes it an interesting tool for training purposes. By identifying specific areas of the procedure that require improvement, the H-OSATS should facilitate constructive feedback and thus deliberate practice. It may also be used for research in surgical education. At present, additional studies are required to explore its potential for use in summative assessment, and to generate specific proficiency cut-off values.

This study has limitations. The VR simulator did not enable the evaluation of tasks 1 to 5 and tasks 12 to 14. However, tasks that are the most specific of this procedure were evaluated. Validity of the entire scale needs to be confirmed in the OR. The validation process will have to ensure that the degree of verbal instructions given by a senior surgeon does not interfere with the evaluation of surgical trainees.

This study demonstrated that the H-OSATS scale is a valid instrument for assessment of technical performances for LH performed on a VR simulator. It is a step towards a more objective assessment of technical performances for this advanced laparoscopic procedure. Future research will focus on evaluating the H-OSATS scale in the OR.

Acknowledgment:

The authors thank the members of the panel of experts for their expertise and the time invested in answering the Delphi survey (alphabetically): Olivier Donnez, Jon I. Einarsson, Anna Fagotti, Andrew Kent, Philippe Koninckx, Philippe Laberge, Mario Malzoni, Camran Nezhat, Michelle Nisolle, Resad Pasic, Benoit Rabischong, Andreas Thurkow, Thierry Vancaillie and Arnaud Wattiez.

Disclosure:

Rajesh AGGARWAL is a consultant for Applied Medical. Sophie KNIGHT, Aubert AGOSTINI, Anderson LOUNDOU, Stéphane BERDAH and Patrice CROCHET have no conflicts of interest or financial ties to disclose.

References:

1. Mistry K, Fingar KR, Elixhauser A (2006) Variation in the Rate of Cesarean Section Across U.S. Hospitals, 2013: Statistical Brief #211.
2. Desai VB, Xu X (2015) An update on inpatient hysterectomy routes in the United States. *Am J Obstet Gynecol* 213:742–743
3. Aarts JWM, Nieboer TE, Johnson N, Tavender E, Garry R, Mol BWJ, Kluivers KB (2015) Surgical approach to hysterectomy for benign gynaecological disease. *Cochrane Database Syst Rev* 8:CD003677. doi: 10.1002/14651858.CD003677.pub5.
4. Beck TL, Morse CB, Gray HJ, Goff BA, Urban RR, Liao JB (2015) Route of hysterectomy and surgical outcomes from a statewide gynecologic oncology population: is there a role for vaginal hysterectomy? *Am J Obstet Gynecol*. doi: 10.1016/j.ajog.2015.10.006
5. Morice P, Leary A, Creutzberg C, Abu-Rustum N, Darai E (2015) Endometrial cancer. *Lancet* 387(10023):1094-108
6. Reich H (1992) Laparoscopic hysterectomy. *Surg Laparosc Endosc* 2:85–88
7. Wattiez A, Soriano D, Cohen SB, Nervo P, Canis M, Botchorishvili R, Mage G, Pouly JL, Mille P, Bruhat MA (2002) The learning curve of total laparoscopic hysterectomy: comparative analysis of 1647 cases. *J Am Assoc Gynecol Laparosc* 9:339–345

8. Mäkinen J, Johansson J, Tomás C, Tomás E, Heinonen PK, Laatikainen T, Kauko M, Heikkinen a M, Sjöberg J (2001) Morbidity of 10 110 hysterectomies by type of approach. *Hum Reprod* 16:1473–1478
9. Terzi H, Biler A, Demirtas O, Guler OT, Peker N, Kale A (2016) Total laparoscopic hysterectomy: Analysis of the surgical learning curve in benign conditions. *Int J Surg* 35:51–57
10. Reznick RK, MacRae H (2006) Teaching surgical skills--changes in the wind. *N Engl J Med* 355:2664–2669
11. Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, Brown M (1997) Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg* 84:273–278
12. Vassiliou MC, Feldman LS, Andrew CG, Bergman S, Leffondré K, Stanbridge D, Fried GM (2005) A global assessment tool for evaluation of intraoperative laparoscopic skills. *Am J Surg* 190:107–113
13. Larsen CR, Grantcharov T, Schouenborg L, Ottosen C, Soerensen JL, Ottesen B (2008) Objective assessment of surgical competence in gynaecological laparoscopy: development and validation of a procedure-specific rating scale. *BJOG* 115:908–916
14. Peyre SE, Peyre CG, Hagen JA, Sullivan ME, Lipham JC, Demeester SR, Peters JH, Demeester TR (2009) Laparoscopic Nissen fundoplication assessment: task analysis as a model for the development of a procedural checklist. *Surg Endosc* 23:1227–1232
15. Palter VN, MacRae HM, Grantcharov TP (2011) Development of an objective evaluation tool to assess technical skill in laparoscopic colorectal surgery: a Delphi methodology. *Am J Surg* 201:251–259

16. Zevin B, Bonrath EM, Aggarwal R, Dedy NJ, Ahmed N, Grantcharov TP, ATLAS group (2013) Development, feasibility, validity, and reliability of a scale for objective assessment of operative performance in laparoscopic gastric bypass surgery. *J Am Coll Surg* 216:955-965-1031, 1033
17. Sarker SK, Chang A, Albrani T, Vincent C (2008) Constructing hierarchical task analysis in surgery. *Surg Endosc* 22:107–111
18. Reich H (2007) Total laparoscopic hysterectomy: indications, techniques and outcomes. *Curr Opin Obstet Gynecol* 19:337–344
19. Wood C, Maher PJ (1997) Laparoscopic hysterectomy. *Baillieres Clin Obstet Gynaecol* 11:111–136
20. Clayton RD (2006) Hysterectomy. *Best Pract Res Clin Obstet Gynaecol* 20:73–87
21. Einarsson JI, Suzuki Y (2009) Total laparoscopic hysterectomy: 10 steps toward a successful procedure. *Rev Obstet Gynecol* 2:57–64
22. Velemir L, Azuar A-S, Botchorishvili R, Canis M, Jardon K, Rabischong B, Pouly J-L, Mage G (2009) [Optimizing the role of surgeons assistants during a laparoscopic hysterectomy]. *Gynécologie, Obs Fertil* 37:74–80. doi: 10.1016/j.gyobfe.2008.09.013
23. Thoma V, Salvatores M, Mereu L, Chua I, Wattiez A (2007) [Laparoscopic hysterectomy: technique, indications]. *Ann Urol (Paris)* 41:80–90
24. Nezhat C, Nezhat F, Admon D, Nezhat AA (1995) Proposed classification of hysterectomies involving laparoscopy. *J Am Assoc Gynecol Laparosc* 2:427–429
25. Fink A, Kosecoff J, Chassin M, Brook RH (1984) Consensus methods: characteristics and guidelines for use. *Am J Public Health* 74:979–983
26. Jones J, Hunter D (1995) Consensus methods for medical and health services research. *BMJ* 311:376–380

27. Blikkendaal MD, Twijnstra ARH, Stiggelbout AM, Beerlage HP, Bemelman WA, Jansen FW (2013) Achieving consensus on the definition of conversion to laparotomy: a Delphi study among general surgeons, gynecologists, and urologists. *Surg Endosc* 27:4631–4639
28. Janssen PF, Brölmann H a. M, Huirne J a. F (2011) Recommendations to prevent urinary tract injuries during laparoscopic hysterectomy: a systematic Delphi procedure among experts. *J Minim Invasive Gynecol* 18:314–321
29. Crochet P, Aggarwal R, Knight S, Berdah S, Boubli L, Agostini A (2016) Development of an evidence-based training program for laparoscopic hysterectomy on a virtual reality simulator. *Surg Endosc*. doi: 10.1007/s00464-016-5249-3
30. Aggarwal R, Grantcharov T, Moorthy K, Milland T, Darzi A (2008) Toward feasible, valid, and reliable video-based assessments of technical surgical skills in the operating room. *Ann Surg* 247:372–379
31. Cook DA, Beckman TJ (2006) Current concepts in validity and reliability for psychometric instruments: theory and application. *Am J Med* 119:166.e7-16
32. Regehr G, MacRae H, Reznick RK, Szalay D (1998) Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination. *Acad Med J Assoc Am Med Coll* 73:993–997
33. Sarker SK, Kumar I, Delaney C (2010) Assessing operative performance in advanced laparoscopic colorectal surgery. *World J Surg* 34:1594–1603
34. Diamond IR, Grant RC, Feldman BM, Pencharz PB, Ling SC, Moore AM, Wales PW (2014) Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. *J Clin Epidemiol* 67:401–409

35. Gracht HA Von Der (2012) Consensus measurement in Delphi studies Review and implications for future quality assurance. *Technol Forecast Soc Change* 79:1525–1536
36. Downing S (2003) Validity: on the meaning ful interpretation of assessment data. *Med Educ* 37:830–837
37. Messick S (1995) Standards of Validity and the Validityof Standardsin PerSormance Assessment. *Educ Meas Issues Pract* 14:5–8
38. Raine-Bennett T, Tucker L-Y, Zaritsky E, Littell RD, Palen T, Neugebauer R, Axtell A, Schultze PM, Kronbach DW, Embry-Schubert J, Sundang A, Bischoff K, Compton-Phillips AL, Lentz SE (2016) Occult Uterine Sarcoma and Leiomyosarcoma: Incidence of and Survival Associated With Morcellation. *Obstet Gynecol* 127:29–39
39. Nassif J, Wattiez A (2010) Clermont Ferrand uterine manipulator. *Surg Technol Int* 20:225–231
40. Tremblay C, Grantcharov T, Urquia ML, Satkunaratham A (2014) Assessment tool for total laparoscopic hysterectomy: a Delphi consensus survey among international experts. *J Obstet Gynaecol Canada JOGC = J d'obstétrique gynécologie du Canada JOGC* 36:1014–1023
41. Frederick PJ, Szender JB, Hussein AA, Kesterson JP, Shelton JA, Anderson TL, Barnabei VM, Guru K (2016) Surgical Competency for Robot-Assisted Hysterectomy: Development and Validation of a Robotic Hysterectomy Assessment Score (RHAS). *J Minim Invasive Gynecol*. doi: 10.1016/j.jmig.2016.10.004.
42. Weizman NF, Manoucheri E, Vitonis AF, Hicks GJ, Einarsson JI, Cohen SL (2015) Design and validation of a novel assessment tool for laparoscopic suturing of the vaginal cuff during hysterectomy. *J Surg Educ* 72:212–219

Figure: Spearman’s rank correlation between H-OSATS scores and OSATS scores for laparoscopic hysterectomy (tasks “division of the round ligament” to “colpotomy”) on a Virtual Reality simulator. H-OSATS, Hysterectomy Objective Structured Assessment of Technical Skill; OSATS, Objective Structured Assessment of Technical Skills

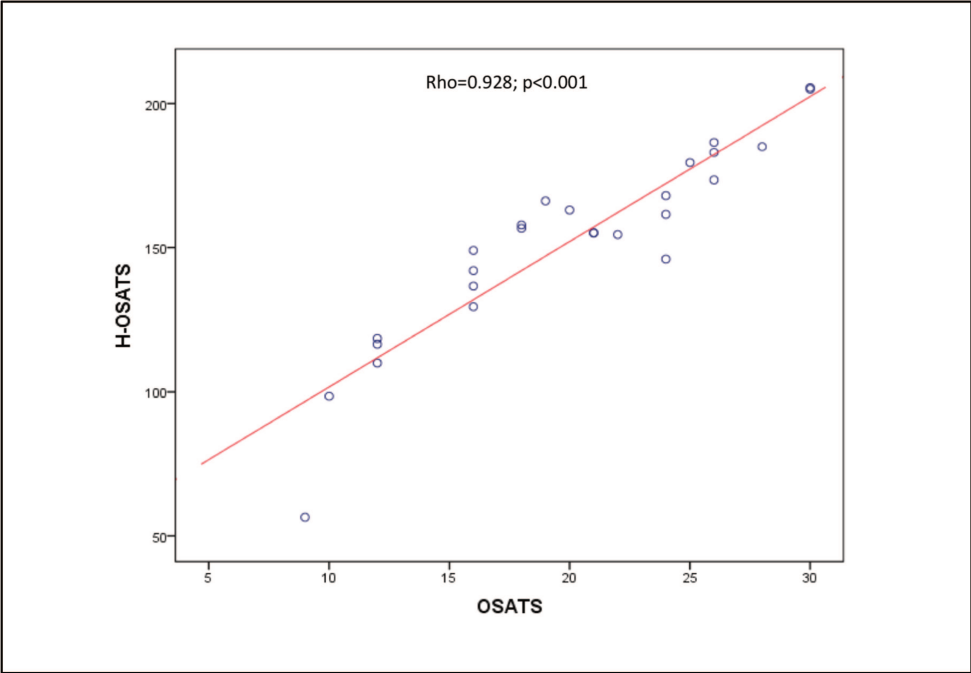


Table 1. Agreement among panel members for first and second round of online Delphi questionnaire on the list of surgical steps

| Tasks | Steps | RoA* 1st round | RoA 2 nd round |
|---|---|----------------------|------------------------------|
| 1. Patient positioning | Legs spread apart (<u>with very little flexion from the abdomen</u>) | 93% | 100% |
| | Both arms tucked along side | 86% | 79% |
| | Buttocks close (<u>slightly over the edge of the operating table</u>) to the edge of the operating table | 93% | 100% |
| 2. Abdominal access | Achieve intraperitoneal access using a recognized method (Veress needle, open technique, etc.) | 100% | 100% |
| | <u>If open technique:</u> Check the optical viewing trocar is placed in the peritoneal cavity before insufflation | 21% | 57% |
| | Create appropriate pneumoperitoneum | 100% | 100% |
| 3. Inspection of the peritoneal cavity | Perform diagnostic laparoscopy (<u>including liver and diaphragm</u>) | 93% | 93% |
| | Patient put in <i>steep</i> Trendelenburg position <u>allowing appropriate exposure</u> | 71% | 100% |
| 4. Trocar insertion | Avoid epigastric vessels | 100% | 100% |
| | Insertion of three operating trocars | 71% | 79% |
| | Ergonomic trocar placement | 100% | 100% |

| | | | |
|---|---|------|------|
| | Look for injuries from port placement | 100% | 93% |
| 5. Inspection of the pelvis | Expose pelvis: retract small bowel and sigmoid colon, perform adhesiolysis if necessary | 93% | 100% |
| | Inspection of uterus and adnexas | 93% | 92% |
| | Insert the uterine manipulator | 69% | 79% |
| | Check that uterine manipulator allows appropriate exposure (i.e: it has its 6 degrees of freedom) | 86% | 92% |
| | Check access to pouch of Douglas and sub-ovarian fossas | 93% | 79% |
| | Visualise ureters as they cross over the iliac vessels and travel downwards to the lateral pelvic walls <u>Check ureter's path in the pelvis</u> | 92% | 100% |
| 6. Division of the round ligaments | Manipulator: push uterus cranially and laterally towards the opposite side, <i>and maintain in a medial position</i> | 78% | 93% |
| | Put round ligament into <u>moderate</u> tension | 71% | 57% |
| | Coagulation and transection of the round ligament | 78% | 86% |
| | Individualize the front and back fold of the anterior leaf of the broad ligament | 78% | 71% |
| | Manipulator: push uterus cranially, laterally towards the opposite side <i>while being maintained in a medial position</i> | 78% | 100% |

| | | | | |
|---|--|---|-----|------|
| 7. Division of the IP ligament or the utero-ovarian ligament | Expose IP ligament or utero-ovarian ligament | | 93% | 100% |
| | If fenestration of the broad ligament is performed | Open the anterior leaf of the broad ligament backwards (<u>parallel with IP ligament</u>) | 50% | 50% |
| | | Expose the posterior leaf of the broad ligament in its grey area | 61% | 86% |
| | | Open a peritoneal window in the broad ligament and enlarge | 71% | 71% |
| | | Check the ureter has been put at a distance | 78% | 78% |
| | If fenestration of the broad ligament is not performed | Identify the ureter by transperitoneal visualisation | 69% | 78% |
| | Coagulate the IP ligament (if adnexectomy) or the utero-ovarian ligament (if interadnexal hysterectomy) using an appropriate energy source or suture | | 93% | 86% |
| | Section the IP ligament (if adnexectomy) or the utero-ovarian ligament (if interadnexal hysterectomy) | | 86% | 86% |
| 8. Creation of the bladder flap | Manipulator: push uterus cranially, towards the opposite side <i>while being maintained in a medial position or pushed posteriorly</i> | | 93% | 93% |

| | | | |
|---|--|---------------------|------|
| | <i>Apply tension to the external fold of the prevesical peritoneum in order to open the plan</i> | 50% | - |
| | <i>Open of the vesicouterine plan on the lateral side</i> | 50% | - |
| | <u>Open the anterior fold of the broad ligament on both sides down to the level of the vesico-uterine reflexion</u> | - | 79% |
| | Manipulator: push uterus cranially and posteriorly <i>while being maintained in a median position in the transverse plan</i> | 85% | 86% |
| | Section of the anterior peritoneum down to the lower uterine segment | 85% | 86% |
| | The bladder is grasped at the midline, applying an anterior-superior traction | 58% | 71% |
| | Opening of the vesicouterine space at the midline <i>to expose the cervicovaginal margin</i> <u>until the cervico-vaginal margin is exposed</u> | 61% | 79% |
| 9. Opening of the posterior peritoneum | Manipulator: push uterus <u>anteriorly and</u> cranially, towards the opposite side, <i>while being kept in a medial position or pushed anteriorly</i> | 71% | 93% |
| | Dissection and section of the posterior leaf of the broad ligament downwards and towards the insertion of the utero-sacral ligaments <u>on each side</u> | 77% | 93% |
| 10. Division of the uterine vessels | Manipulator: push uterus towards the opposite side, cranially <i>while being kept in a medial position</i> | 92% | 100% |
| | Optimize exposure of the uterine vessels | 100% | 100% |
| | Skeletonize uterine vessels at the ascending portion of the uterine artery | 64%** 71% 71% | 100% |

| | | | |
|---|--|------|------|
| | If anatomy is not distorted: Identify the ureter prior to division of the uterine vessels | 42% | 50% |
| | If anatomy is distorted: Identifying the ureter prior to division of the uterine vessels | 100% | 93% |
| | Coagulate the uterine vessels using an appropriate energy source or suture | 93% | 100% |
| | Section uterine vessels <i>at the appropriate level in the ascending portion, at the level of the colpotomizer</i> | 78% | 71% |
| | Divide <i>the distal</i> cervical attachment of the cardinal ligament | 64% | 85% |
| | Divide the uterosacral ligament | 21% | 36% |
| 11. Colpotomy: | Manipulator: push uterus cranially | 92% | 100% |
| | Identify the cervico-vaginal delineation from the colpotomizer | 93% | 100% |
| | Check that there are no interposed elements around the vaginal fornices and complete dissection if necessary | 78% | 100% |
| | Identify the ureter prior to proceeding to the colpotomy | -14% | 36% |
| | Proceed to circumferential colpotomy using an appropriate energy source | 93% | 100% |
| 12. Uterus retrieval and vault closure | Specimen retrieval vaginally <i>or laparoscopically by morcellation</i> | 71% | 100% |
| | <u>If the specimen is not retrieved in once piece through vaginal route, this step will not be assessed.</u> | | |

| | | | |
|---------------------------------------|---|------|------|
| | Occlude vagina to restore pneumoperitoneum | 78% | 86% |
| | Identify the ureter prior to proceeding to the suture | 14% | 43% |
| | Suture the vaginal vault angles separately | 0% | 43% |
| | Suture the <i>remaining</i> vaginal vault with interrupted or continuous sutures | 93% | 93% |
| | Vaginal suture including sufficient width of vaginal mucosa and fascia | 93% | 100% |
| | Suture includes the US ligaments <i>to restore pericervical ring for pelvic support</i> | 50% | 71% |
| 13. Haemostasis and inspection | Irrigation <u>and aspiration</u> of the pelvis | 78% | 93% |
| | Check vascular pedicles, bladder reflection and vaginal cuff <i>if needed</i> <u>under low abdominal pressure</u> : secure hemostasis is needed | 100% | 86% |
| | Check there is no damage to surrounding structures | 84% | 93% |
| | <u>Perform cystoscopy or Indigo Carmin test if ureteral integrity is of concern.</u> | - | 85% |
| | Perform cystoscopy or bleu test if bladder integrity or ureteral patency is of concern | 71% | 86% |
| 14. Port removal | Remove trocars under direct visualisation and inspecting port sites for haemostasis | 78% | 79% |
| | Evacuate pneumoperitoneum | 84% | 100% |

| | | | |
|---|---|-----|------|
| | Suture fascia for trocars ≥ 10 mm | 93% | 100% |
| | Close skin incisions with any acceptable method | 93% | 85% |
| Order in which tasks should be performed | Perform tasks 1 (patient positioning) to 5 (inspection of the pelvis) in a chronological order | 78% | 100% |
| | Perform task 6 (division of the round ligament) and 7 (division of the IP ligament/utero-ovarian) for each side in any order | 21% | 57% |
| | Perform task 8 (creation of the bladder flap) and 9 (opening of the posterior peritoneum) for each side in any order | 7% | 57% |
| | Perform tasks 10 (division of the uterine vessels) to 14 (port removal) in a chronological order | 64% | 86% |
| Penalty if tasks performed in an incorrect order | Errors: task 9 (opening posterior peritoneum) before task 7 (IP ligament or utero-ovarian section) | 36% | 21% |
| | Errors: task 10 (division of the uterine vessels) before task 8 (creation of the bladder flap) AND 9 (opening posterior peritoneum) are performed | 21% | 57% |
| | Errors: task 11 (colpotomy) before task 10 (division of the uterine vessels) | 57% | 86% |

* RoA: Rate of agreement

** Missing data replaced by 3, mean and mode, respectively. Reformulations between round 1 and 2 are underlined. Deletion between round 1 and 2 are in *italic*.

Table 2. Composition of the international expert panel for the online delphi questionnaire

| Location | Participants in round 1 | | Participants in round 2 | |
|-----------------|-------------------------|-----------|-------------------------|-----------|
| | Contacted | Responded | Contacted | Responded |
| Australia | 1 | 1 | 1 | 1 |
| Belgium | 3 | 3 | 3 | 3 |
| Canada | 1 | 1 | 1 | 1 |
| France | 2 | 2 | 2 | 2 |
| Italy | 2 | 2 | 2 | 2 |
| The Netherlands | 1 | 1 | 1 | 1 |
| Turkey | 1 | 0 | 0 | 0 |
| UK | 1 | 1 | 1 | 1 |
| USA | 5 | 3 | 3 | 3 |
| Total | 17 | 14 | 14 | 14 |

Table 3. Comparison of scores for different components of Laparoscopic Hysterectomy as assessed by Hysterectomy Objective Structured Assessment of technical Skill scale, using case-volume criteria for definition of novice intermediate and experienced operators (results are presented as median scores).

| Tasks | Novices n=10 | Intermediates n=8 | Experts n=8 | p | Maximum possible score | |
|-------|---|----------------------|----------------|--------|---------------------------|-----|
| 6 | Division of the round ligaments | 22.75 | 24.25 | 25.75 | 0.091 | 30 |
| 7 | Division of the IP ligament or the utero-ovarian ligament | 28.75 | 36.60 | 42.5 | 0.001 | 50 |
| 8 | Creation of the bladder flap | 26.75 | 34.75 | 32.25 | 0.035 | 40 |
| 9 | Opening of the posterior peritoneum | 10.75 | 13.25 | 17 | 0.049 | 20 |
| 10 | Division of the uterine vessels* | 34 | 36.75 | 46 | 0.004 | 50 |
| 11 | Colpotomy | 11.5 | 16.50 | 18 | <0.001 | 20 |
| | Total | 133 | 155.90 | 178.25 | 0.002 | 210 |

* Steps « If anatomy is distorted: Identifying the ureter prior to division of the uterine vessels » and « Divide cervical attachment of the cardinal ligaments » were not assessed as the LH program provided by the simulator does not allow assessement of these steps.

Appendix: H-OSATS scale

| Tasks | | Steps | | 1 | 2 | 3 | 4 | 5 | Score | |
|-------|--|-------|---|---|---|---|---|--|-------|-----|
| 1 | Patient positioning (once patient fully equipped with Foley catheter inserted) | 1.1 | Legs spread apart (with very little flexion from the abdomen) | Inadequately performed | | Performed adequately (access to the perineum is not optimal) | | Performed adequately (optimal access to the perineum) | /5 | /15 |
| | | 1.2 | Both arms tucked along side | Not performed | | | | Performed adequately, allowing an easy access to the patient | /5 | |
| | | 1.3 | Buttocks slightly over the edge of the operating table | Not performed adequately | | Performed adequately, moderate sliding up while in Trendelenburg position | | Performed adequately | /5 | |
| 2 | Abdominal access | 2.1 | Achieve intraperitoneal access using a recognized method (Veress needle, open technique, etc) | Performed with difficulty or inadequately | | Performed adequately with hesitation but in a safe fashion | | Performed efficiently and skillfully | /5 | /10 |
| | | 2.2 | Create appropriate pneumoperitoneum | Failed to perform | | Performed with an incorrect pressure and slow to correct | | Performed efficiently and skillfully | /5 | |
| 3 | Inspection of the peritoneal cavity | 3.1 | Perform diagnostic laparoscopy (including liver and diaphragm) | Not performed | | Moderately detailed inspection, not all 4 quadrants visualized | | Careful and thorough inspection (including liver and diaphragm) | /5 | /10 |
| | | 3.2 | Patient put in Trendelenburg position allowing appropriate exposure | Not performed | | Performed after delay. Suboptimal exposure | | Patient sufficiently tilted allowing optimal exposure | /5 | |
| 4 | Trocar insertion | 4.1 | Avoid epigastric vessels | Location of the epigastric vessels not checked | | | | Anatomical landmarks identified + transillumination performed | /5 | /20 |
| | | 4.2 | Insertion of 3 operating trocars | Performed inadequately; without intraperitoneal visualization | | Performed adequately in a safe fashion | | Performed smoothly and skillfully | /5 | |
| | | 4.3 | Ergonomic trocar placement | Port placement does not allow appropriate triangulation | | Port placement allows triangulation, but awkward posture of the surgeon | | Port placement allowing appropriate triangulation and ergonomic operator posture | /5 | |
| | | 4.4 | Look for injuries from port placement | Not performed | | Injury cannot be definitively excluded | | Injury properly excluded | /5 | |

| | | | | | | | | |
|----------|---|-----|--|---|---|--|-----|-----|
| 5 | Inspection of the pelvis | 5.1 | Expose pelvis: retract small bowel and sigmoid colon, perform adhesiolysis if necessary | Failed to expose pelvis adequately; use of traumatic grasper; adhesiolysis causing damage to surrounding structures | Exposed pelvis clumsily. Adhesiolysis performed adequately with some hesitation | Pelvis exposed adequately and smoothly using an atraumatic grasper | /5 | /30 |
| | | 5.2 | Inspection of uterus and adnexas | Not performed | Inspection done but some areas not adequately visualized | Careful and thorough inspection | /5 | |
| | | 5.3 | Insertion of the uterine manipulator | Failed to insert the manipulator or done with trauma to tissues | Inserted adequately, but with difficulty | Performed smoothly and skillfully | /5 | |
| | | 5.4 | Check that uterine manipulator allows appropriate exposure (i.e.: it has its 6 degrees of freedom) | Not performed | Not all directions checked | Performed adequately | /5 | |
| | | 5.5 | Check access to pouch of Douglas and sub-ovarian fossas | Not performed | Performed after delay without mobilizing manipulator | Performed adequately | /5 | |
| | | 5.6 | Check ureter's path in the pelvis | Not performed | Position not clearly established, only one ureter identified | Identified clearly and skillfully | /5 | |
| 6 | Division of the round ligaments (left & right) | 6.1 | Manipulator: push uterus cranially and laterally towards the opposite side | Not performed | Uterus mobilized, but suboptimal exposure | Performed adequately, allowing optimal exposure | /10 | /30 |
| | | 6.2 | Coagulation and transection of the round ligament | Performed to close to uterine horn; caused significant bleeding or tissue trauma | Performed adequately, with minor bleeding | Performed skillfully and efficiently | /10 | |
| | | 6.3 | Individualize the front and back fold of the anterior leaf of the broad ligament | Not performed | Performed adequately, but the plan was not easily found | Plan developed skillfully with minimal bleeding and tissue trauma | /10 | |

| | | | | | | | | | |
|---|--|-----------------------------------|--|--|--|---|--|-----|-----|
| 7 | Division of IP ligament or utero-ovarian ligament (left & right) | 7.1 | Manipulator: push uterus cranially and laterally towards the opposite side | Not performed | Uterus mobilized, but suboptimal exposure | Performed adequately allowing optimal exposure | /10 | /50 | |
| | | 7.2 | Expose IP ligament or utero-ovarian ligament | Inadequate exposure putting surrounding structures at risk | Performed adequately | Performed adequately, allowing optimal exposure | /10 | | |
| | | If fenestration is performed | 7.3 | Expose the posterior leaf of the broad ligament in its grey area | Unable to expose the posterior leaf | Performed adequately, with minor bleeding | Performed skillfully and efficiently | | /10 |
| | | | 7.4 | Open a peritoneal window in the broad ligament and enlarge | Performed inadequately: did not check underlying structures or caused damage | Performed adequately, in the right direction, with minor bleeding | Performed skillfully and efficiently, peritoneal window enlarged by divergent traction | | |
| | | | 7.5 | Check the ureter has been put at a distance | Not performed | Position not clearly established | Clearly identified | | |
| | | If fenestration is not performed: | 7.6 | Identify the ureter by transperitoneal visualization | Not performed | Position not clearly established | Clearly identified | | /10 |
| | | | 7.7 | Coagulate using an appropriate energy source or suture | Coagulated too close to surrounding structures or insufficiently | Coagulated at the appropriate level, with some hesitation | Coagulated safely and efficiently at a 90° angle | | /10 |
| | | 7.8 | Section the IP ligament or the utero-ovarian ligament | Section causing significant bleeding | Section causing minor bleeding rapidly controlled | Performed skillfully at a 90° angle, causing no bleeding | /10 | | |

| | | | | | | | | |
|----------|---|-----|---|---|--|---|-----|-----|
| 8 | Creation of the bladder flap | 8.1 | Manipulator: push uterus cranially and laterally to the opposite side (left and right) | Not performed | Uterus mobilized but required repositioning, suboptimal exposure | Performed adequately allowing optimal exposure | /10 | /40 |
| | | 8.2 | Open the anterior fold of the broad ligament on both sides down to the level of the vesico-uterine reflection | Performed inadequately: in the wrong plan, causing excessive bleeding, in the wrong direction | Performed clumsily but in the right plan. Slow to control bleeding | Performed smoothly in the right plan | /10 | |
| | | 8.3 | Manipulator: push uterus cranially and posteriorly | Not performed | Uterus mobilized but required repositioning, suboptimal exposure | Performed adequately allowing optimal exposure | /5 | |
| | | 8.4 | Section the peritoneum down to the lower uterine segment | Not performed, or caused significant bleeding or tissue trauma | Performed adequately, with hesitation | Performed adequately in the right plan, with no bleeding | /5 | |
| | | 8.5 | The bladder is grasped at the midline, applying an anterior-superior traction | Not performed | Performed with insufficient traction or bladder not properly grasped | Grasped with an atraumatic forceps applying the right amount of traction | /5 | |
| | | 8.6 | Opening of the vesico-uterine space at the midline until the cervico-vaginal margin is exposed | Opening in the wrong plan, caused significant bleeding or tissue trauma | Moderate difficulty to identify correct plan, occasional bleeding | Opening in the correct plan down to the cervico-vaginal margin; no bleeding | /5 | |
| 9 | Opening of the posterior peritoneum (left & right) | 9.1 | Manipulator: push uterus anteriorly and cranially towards the opposite side | Not performed | Uterus mobilized but required repositioning, suboptimal exposure | Performed adequately allowing optimal exposure | /10 | /20 |
| | | 9.2 | Dissection and section of the posterior leaf of the broad ligament towards the insertion of the utero-sacral ligaments on each side | Performed in the wrong direction or in the wrong plan. caused significant bleeding | Performed adequately (Moderate difficulty to identify correct plan, occasional bleeding) | Performed smoothly and skillfully | /10 | |

| | | | | | | | | |
|-----------|---|------|--|---|---|--|-------------|-----|
| 10 | Division of the uterine vessels (left & right) | 10.1 | Manipulator: push uterus cranially and to the opposite side | Not performed | Uterus mobilized but required repositioning, suboptimal exposure | Performed adequately allowing optimal exposure | /10 | /60 |
| | | 10.2 | Optimize exposure of the uterine vessels | Not performed | Insufficient exposure | Excellent exposure (retracting fundus in the opposite direction) | /10 | |
| | | 10.3 | Skeletonize uterine vessels at the ascending portion of the uterine artery | Insufficient dissection of uterine artery; caused bleeding or tissue damage | Sufficient exposure of uterine artery but done with difficulty or causing moderate bleeding | Uterine vessels perfectly exposed, no bleeding | /10 | |
| | | 10.4 | If anatomy is distorted: Identify the ureter prior to division of the uterine vessels | Not performed | Position not clearly identified | Position clearly identified | /- 10 or NA | |
| | | 10.5 | Coagulate the uterine vessels using an appropriate energy source or suture | Performed inadequately: at the wrong level, incorrect angle or inappropriate energy source. | Performed adequately but with hesitation | Performed skillfully and efficiently | /10 | |
| | | 10.6 | Section uterine vessels in the ascending portion, at the level of the colpotomizer | Section at the wrong level occasioning important bleeding or tissue trauma | Performed adequately but with hesitation; occasioned bleeding but was able to control | Performed efficiently and skillfully, at a 90° angle | /10 | |
| | | 10.7 | Divide cervical attachments of the cardinal ligaments | Not performed, or performed with an incorrect angle, or causing significant bleeding or tissue trauma | Performed efficiently but required some time | Performed skillfully and efficiently | /10 | |
| 11 | Colpotomy | 11.1 | Manipulator: push uterus cranially | Not performed | Uterus mobilized but required repositioning, suboptimal exposure | Performed adequately allowing optimal exposure | /5 | /20 |
| | | 11.2 | Identify the cervico-vaginal delineation from the colpotomizer | Cervico-vaginal delineation poorly identified | Identified with some difficulty or incompletely | Cervico-vaginal delineation clearly identified | /5 | |
| | | 11.3 | Check that there are no interposed elements around the vaginal fornices and complete dissection if necessary | Not performed | Performed partially and hastily | Careful inspection, clear visualization of the distance from bladder, rectum and ureters | /5 | |
| | | 11.4 | Proceed to circumferential colpotomy using an appropriate energy source | Performed in wrong location, not over the colpotomizer; caused excessive bleeding or tissue trauma; poor exposure | Performed adequately but with difficulty, causing moderate bleeding | Performed skillfully over the colpotomizer on the full circumference of the vaginal fornices | /5 | |

| | | | | | | | | | |
|------|--|--|--|---|--|---|----|-----|----------|
| 12 | Uterus retrieval and vault closure | 12.1 | Specimen retrieval vaginally | Failed to perform or caused tissue damage | Performed adequately but with difficulty | Performed efficiently and safely (checking integrity of surrounding structures) | /5 | /20 | |
| | | If the specimen is not retrieved in once piece through vaginal route, this step will not be assessed | | | | | | | |
| | | 12.2 | Occlude vagina to restore pneumoperitoneum | Looses pneumoperitoneum entirely | Performed after delay or partial leak of pneumoperitoneum which is not corrected | Pneumoperitoneum fully restored, performed rapidly allowing good exposure | /5 | | |
| | | 12.3 | Suture the vaginal vault with interrupted or continuous sutures | Inadequate knot tying technique, damage to surrounding structures or poor quality of suture | Performed adequately, tissues generally well approximated | Performed skillfully and smoothly; Tissues very well approximated | /5 | | |
| | | 12.4 | Vaginal suture including sufficient width of vaginal mucosa and fascia | Suture not including both layers; Insufficient bites | Performed adequately, but uneven width of vagina | Performed efficiently with regular good bites | /5 | | |
| 12.5 | Suture includes US ligaments for pelvic support (if US insertion divided from the vagina during colpotomy) | Not performed, or caused damage to surrounding structures | Performed adequately and clumsily | Performed skillfully and efficiently | /- 5 or NA | | | | |
| 13 | Hemostasis and inspection | 13.1 | Irrigation and aspiration of the pelvis | Not performed | Performed incompletely or required some time to perform | Performed skillfully and efficiently | /5 | /15 | |
| | | 13.2 | Check vascular pedicles, bladder reflection and vaginal cuff under low abdominal pressure: secure hemostasis if needed | Not performed | Performed adequately but not all areas inspected | Careful inspection of all areas and meticulous hemostasis if needed | /5 | | |
| | | 13.3 | Check there is no damage to surrounding structures | Not performed | Moderately detailed inspection | Careful and thorough inspection | /5 | | |
| | | 13.4 | Perform cystoscopy or Indigo Carmin test if ureteral integrity is of concern. | - 5 if not performed | | | | | NA or -5 |
| | | 13.5 | Perform cystoscopy or bleu test if bladder integrity is of concern. | - 5 if not performed | | | | | NA or -5 |

| | | | | | | | | |
|-----------|--------------------------------------|------|--|---|---|--------------------------------------|----------|-----|
| 14 | Port removal and skin closure | 14.1 | Remove trocars under direct visualization and inspect port sites for hemostasis | Not performed | Performed incompletely | Performed smoothly and skillfully | /5 | /20 |
| | | 14.2 | Evacuate pneumoperitoneum | Not performed | Incompletely evacuated | Completely evacuated | /5 | |
| | | 14.3 | Suture fascia for trocars \geq 10 mm | Not performed or inadequately | Required some time to perform but performed correctly | Performed efficiently and skillfully | /5 | |
| | | 14.4 | Close skin incisions with any acceptable method | Performed inadequately | Required some time to perform but performed correctly | Performed efficiently and skillfully | /5 | |
| 15 | Order | 15.1 | Perform tasks 1 (patient positioning) to 5 (inspection of the pelvis) in a chronological order | | | | /5 | /10 |
| | | 15.2 | Perform tasks 10 (division of the uterine vessels) to 14 (port removal) in a chronological order | | | | /5 | |
| | | 15.3 | Task 11 (colpotomy) started before task 10 (division of the uterine vessels) | - 5 if colpotomy started before division of the uterine vessels completed | | | NA or -5 | |

DISCUSSION ET PERSPECTIVES

7.1 Intérêt d'un programme de simulation pour l'HC

Cette thèse présente le premier programme d'entraînement dédié à l'HC sur un simulateur. À l'exception de la suture vaginale qui peut être travaillée sur pelvi-trainer, les gestes spécifiques de cette intervention sont actuellement enseignés exclusivement au BO dans le cadre du compagnonnage. À partir d'une base de données nord-américaine (National Surgical Quality Improvement Program), Igwe et al. ont évalué l'impact sur la santé des patientes de la réalisation de cette intervention par des internes.²⁹ Sur les 3441 patientes opérées entre 2008 et 2011, l'opérateur principal était un interne dans 46% des cas. Le temps opératoire était alors allongé par rapport à celui du sénior (179 vs 135 minutes ; $p < 0.001$), sans différences concernant le taux de complications et le taux de mortalité à 30 jours. Cependant une petite augmentation du taux de transfusions, de ré interventions et de réadmissions était observée lorsque l'interne avait été opérateur principal. Malheureusement, aucune précision n'était disponible à propos des gestes effectivement réalisés par l'interne. Cette publication est une des rares concernant les effets de l'apprentissage de l'HC dans le cadre du compagnonnage classique. Ces résultats ne sont qu'en partie rassurants pour la sécurité des patientes, et illustrent le flou qui existe concernant l'enseignement de l'HC exclusivement au BO. Le principe « ne jamais faire un geste la première fois sur une patiente » devrait être applicable pour cette intervention.

La courbe d'apprentissage au bloc est classiquement évaluée par le nombre de cas nécessaires pour atteindre un temps opératoire et un taux de complications considéré comme acceptable.²⁶ Les chiffres publiés varient entre 30 et 75 HC selon les auteurs et les critères retenus.^{85,86} Ces chiffres sont issus de séries de chirurgiens séniors, et aucune étude à ce jour

n'a estimé le nombre minimum d'interventions nécessaires à réaliser par les internes dans le cadre du compagnonnage. Ce critère de jugement serait néanmoins imparfait, car il dépend trop du contexte du compagnonnage et ne donne que peu d'informations objectives sur la qualité opératoire et le degré d'autonomie obtenue. Selon une enquête française réalisée en 2013 (**annexe 2**), les opportunités opératoires semblent en réalité très limitées pour les internes puisque 60% de ceux de 5^{ème} année déclaraient avoir réalisé moins de 5 HC au cours de leur internat.⁷

En dehors du BO, les outils actuels déjà décrits qui participent à la formation des internes sont les supports didactiques et les vidéos pédagogiques comme celles disponibles sur le site Websurg. La préparation « technique » aux gestes qui seront nécessaires à la réalisation d'une HC repose sur les simulateurs comme les pelvi-trainers et les workshops sur tissu vivant, pour s'entraîner à la gestuelle non spécifique ou bien à la suture.^{21,66} Un modèle spécifique à l'intervention pourrait mieux préparer les internes à bénéficier des quelques opportunités opératoires qu'ils auront au cours de leur formation.

Les simulateurs de RV constituent un support intéressant pour l'apprentissage de l'HC. Tout d'abord car il n'y a pas actuellement d'autre modèle immersif permettant de reproduire la procédure entière. Le modèle animal vivant n'a jusqu'ici pas été développé, probablement car la morphologie de l'utérus du cochon n'est pas assez semblable à celui de la femme (fig F). Par ailleurs, les paramètres automatiquement enregistrés par le simulateur permettent une évaluation objective de la courbe d'apprentissage. Le potentiel de ces outils, pour des interventions gynécologiques simples comme la salpingectomie, a déjà été montré avec un impact positif pour la formation des internes, y compris lors du transfert au BO.⁸⁴

Offrant la possibilité de monitorer les performances et de répéter les essais sans surcoût, les simulateurs de RV représentent une opportunité pour mettre en œuvre des

stratégies éducatives pour des interventions entières en chirurgie mini-invasive. Le curriculum pour l'entraînement à la cholécystectomie sur le LapMentor*, présenté dans les travaux préliminaires en est un exemple, en permettant la répétition de certains temps opératoires jusqu'à atteindre un palier de compétence. Depuis ce travail initial, le curriculum a été intégré dans le cadre d'une stratégie de formation plus large, incluant une phase de « cognitive training » et un scénario de crise simulé pour travailler les compétences non techniques. Cette formation a permis de réduire la courbe d'apprentissage lors du transfert au BO comparativement au compagnonnage classique.⁸⁷ D'autres stratégies éducatives sont applicables pour des interventions complètes grâce aux simulateurs de VR, notamment la pratique délibérée.¹⁰ Depuis la première publication présentée dans les travaux préliminaires, d'autres études ont confirmé que la pratique délibérée permet d'améliorer la qualité opératoire et de dépasser le plateau qui advient avec la seule répétition des exercices.¹¹ L'impact positif de la pratique délibérée en utilisant un simulateur de VR pour le transfert au BO a également été confirmé pour la cholécystectomie.¹² Concernant le curriculum présenté pour l'hystérectomie, la méthode choisie était de ne pas faire d'intervention de type pratique délibérée lors de la courbe d'apprentissage ; le but étant de ne pas confondre la mesure de l'intervention et celle de la courbe d'apprentissage permise par la seule répétition des exercices sur le simulateur.

L'HC est une intervention complexe et difficile à reproduire. Le curriculum développé lors de cette thèse nécessite un support de RV reproduisant fidèlement les composantes de l'intervention. Le programme du LapMentor* utilisé pour le curriculum illustre les considérables avancées en terme de modélisation réalisées dans ce domaine au cours des 10 dernières années. Un des aspects inclus dans le programme qui semble contribuer aux résultats présentés dans le premier travail est la possibilité de régler le manipulateur utérin, y

compris dans le plan cranio-caudal. En effet le positionnement correct du manipulateur est fondamentalement associé à la compétence technique, ce qui est confirmé par la sélection de tous les items liés à ce sujet dans l'échelle H-OSATS. La capacité du programme à distinguer la qualité opératoire du groupe intermédiaire et celle des senior expérimentés est un élément qui a des implications concernant la population cible. En effet, le programme de VR sur lequel est basé le curriculum reproduit un cas simple, mais l'HC reste une intervention qui est complexe. Ce curriculum peut donc probablement être proposé à de jeunes internes mais aussi à des internes en deuxième partie de cursus qui sont les plus à même d'avoir des opportunités au BO par la suite. Ceci devra être confirmé par des études à venir.

Il y a plusieurs freins à l'utilisation des simulateurs de RV pour la formation chirurgicale. Tout d'abord, ces instruments sont peu appréciés parmi les modèles disponibles. Selon une enquête réalisée en 2013 auprès d'internes de 2 centres nord-américains utilisant beaucoup les supports de simulation, les simulateurs de VR était classés en quatrième position derrière le modèle animal vivant, le cadavre humain et les pelvi-trainers.⁸⁸ De plus, cette opinion peu favorable était davantage marquée chez les internes en fin de cursus, ce qui constituerait un frein à leur utilisation. Les deux critères qui sont souvent évoqués sont le manque de réalisme graphique ainsi que le peu de finesse du système de retour de force.⁸⁹ Pourtant, selon une autre enquête, les enseignants canadiens croyaient au potentiel de ces simulateurs et les classaient en deuxième position pour leur utilité pédagogique derrière l'expérience au BO.⁹⁰ Si on juge les progrès considérables réalisés par les simulateurs de RV depuis 10 ans, ce modèle est amené à jouer un rôle majeur dans l'éducation chirurgicale à l'avenir. À cet égard, l'exemple du curriculum pour la cholécystectomie est illustratif. Publié en 2009, ce programme qui reproduit une intervention fréquente en chirurgie générale a depuis fait la preuve de son efficacité lors du transfert au bloc opératoire⁸⁷ et a été cité 180 fois dans la

littérature. Le curriculum présenté dans ce travail concerne une intervention tout aussi importante pour la chirurgie gynécologique et complexe à modéliser. Ses applications éducatives doivent être encore explorées. Le deuxième frein est lié au prix élevé de ces simulateurs à l'achat (80 à 100 000 € pour le LapMentor™). Cependant, ils offrent une palette d'enseignement plus large que les simulateurs simples et évitent la question éthique que soulève le recours aux animaux vivants et aux cadavres. Une fois acquis, leur entretien est peu onéreux et leur utilisation est possible par plusieurs spécialités (urologie, gynécologie, chirurgie digestive pour le LapMentor* utilisé pour ce travail). Ils sont facilement accessibles et permettent les répétitions multiples des gestes sans nécessiter de consommables comme pour les modèles synthétiques. De plus le coût des simulateurs de VR doit être relativisé par rapport à celui de l'entraînement « classique » au BO qui est loin d'être négligeable. Bridges et al. ont estimé celui-ci à \$ 47 000 par interne, sur la base du temps supplémentaire au BO lorsque l'interne est l'opérateur sur une période de 4 ans.⁹¹ Si l'on ajoute les coûts liés à l'enseignement par la faculté, ce montant s'élèverait entre 62 810 € et 172 682 €.⁹²

Les curricula préparant les internes à réaliser des interventions complètes grâce aux simulateurs de VR sont des outils prometteurs pour l'enseignement de la chirurgie. Ils n'ont pas vocation à remplacer les autres modèles qui gardent des avantages propres et peuvent être utilisés de manière complémentaire.²⁵ Ainsi les modèles plus simples comme les pelvi-trainers ont fait la preuve d'une grande efficacité à un faible coût pour l'entraînement à la suture et pourraient être intégrés à ce curriculum pour l'HC, afin de préparer au temps de fermeture vaginale.⁶⁹ Le curriculum récemment publié par Shore et al. pour les internes de gynécologie n'est pas spécifique d'une procédure mais inclut ces deux types de modèles - RV et pelvi-trainers.⁸³ Enfin, le modèle porcin, même s'il ne permet pas un entraînement spécifique à l'HC, reste le modèle le plus fidèle pour reproduire les sensations opératoires de

la coelioscopie : perception du retour de force, dissection avec pneumopéritoine et gestion des saignements.

7.2 Intégration du curriculum de RV dans la formation globale des internes

Les recherches en pédagogie chirurgicale ont mis en lumière l'importance des autres compétences que celles liées aux gestes techniques. Les connaissances théoriques constituent le premier pilier de la formation. Cette connaissance est acquise en 3 phases successives : conceptualisation, visualisation et verbalisation.⁹³ Pendant la phase de conceptualisation, l'interne apprend l'anatomie, les indications, les temps opératoires clefs d'une intervention, les erreurs à éviter ainsi que la gestion du matériel. Ces éléments sont enseignés par des cours didactiques et idéalement des supports complémentaires comme la vidéo.⁹⁴ Les phases de visualisation et verbalisation permettent à l'interne de décrire l'intervention, et ainsi de renforcer ses acquis. Il est en effet important de vérifier ce prérequis théorique avant d'aborder les phases suivantes de la formation. Le «blended learning» est une approche plus récente qui combine cours magistraux classiques et e-learning.⁹⁵ La plupart des études portant sur ce type d'enseignement sont de faible qualité méthodologique⁹⁶, mais le support web facilite le contrôle des connaissances. D'autres approches utilisant internet sont prometteuses pour l'enseignement à la décision dans un contexte opératoire, comme la plateforme « think like a surgeon ».⁹⁴ À la suite de cette formation cognitive, la partie de formation technique psychomotrice intervient au moyen d'un programme d'entraînement comme celui présenté dans cette thèse. Les autres compétences chirurgicales nécessaires au cours de l'intervention sont intitulées « compétences non techniques » et correspondent aux capacités à travailler en équipe, à prendre des décisions, à détecter les situations à risque et à réagir de manière appropriée face au stress.⁹⁷ Ces compétences peuvent également être enseignées en dehors

du bloc dans des conditions simulées. Un curriculum d'entraînement qui associe les formations cognitive, technique et non technique dans un domaine chirurgical est alors dénommé « comprehensive curriculum » selon le modèle de Zevin.²⁰ Par ailleurs, les compétences chirurgicales ne sont pas limitées au BO (réel ou virtuel). La gestion du parcours d'une patiente, incluant les suites opératoires, est actuellement enseignée par compagnonnage. Cette formation dispensée « au lit du malade » est très hétérogène, comme le montre la grande variabilité du nombre de complications évitables d'un service de chirurgie à l'autre (98). Beyer et al. ont montré l'intérêt d'une formation en parcours de soin, à partir de mises en situation virtuelles pour la prise en charge de l'appendicite aiguë.⁹⁹

La formation sur simulateur de RV reste indissociable de celle apportée par le compagnonnage au BO, surtout pour des interventions complexes comme l'HC. Au-delà des acquis techniques qu'il est susceptible de générer, le curriculum de RV est un atout potentiel pour optimiser et structurer la formation au BO. Un schéma théorique d'apprentissage au bloc a été proposé, appelé BID model (Briefing - Intraoperative teaching - Debriefing). Il s'agit pour l'interne de définir lui-même un objectif avant le bloc et d'en faire part au sénior, de manière à focaliser l'enseignement sur le temps opératoire choisi, puis de faire un débriefing sur ce point précis (Fig G).¹⁰⁰ Cette proposition est intéressante et facile à intégrer dans la pratique clinique quotidienne, même si elle a été peu évaluée. Le choix par l'interne du temps opératoire à travailler est censé faciliter les processus de la mémorisation à long terme. Cependant ce choix devrait être dicté par les besoins réels de l'interne au regard de sa courbe d'apprentissage, et ceux-ci ne correspondent pas forcément à ce qu'il ressent. Dans la théorie de la pratique délibérée, c'est le coach qui décide avec l'élève des éléments à travailler. En préparant et en évaluant l'interne en amont sur un simulateur de RV, l'enseignant devrait avoir une meilleure connaissance des temps opératoires que l'interne peut réaliser au BO et

des objectifs qui peuvent être fixés, tout en gardant la structure du BID model.⁴

Figure G : « Briefing Intra-operative teaching Debriefing » model, d’après Roberts et al.¹⁰⁰

| Table 2. Briefing, Intraoperative Teaching, Debriefing Model Card | | |
|--|--|--|
| Stage | Step | Script |
| Briefing | Set learning objectives for encounter. | “What would you like to focus on?” OR “Today I want you to focus on” |
| Intraoperative teaching | Teaching during the encounter | Focused on stated objectives |
| Debriefing | Reflection | “How do you think you did? Why?” |
| | Rules | “What did you learn for next time?” |
| | Reinforcement | “You did well at” |
| | Correction | “Next time, do this” |

7.3 Intérêt de l’échelle spécifique pour l’HC

L’échelle spécifique d’une intervention est intégrée à la proposition de curriculum de Zevin. L’évaluation séquentielle et descriptive est en effet particulièrement utile pour des procédures comme l’HC, dont certains temps opératoires sont plus difficiles que d’autres et requièrent un niveau technique différent.⁴ Cet outil a un rôle important à jouer pour la formation, car il permet de structurer les interventions pédagogiques. Plusieurs utilisations sont possibles : le feedback par un observateur qualifié ou une auto-évaluation par l’interne à partir d’une vidéo de sa performance.¹⁸ Cette échelle spécifique devrait pouvoir être utilisée sur un simulateur de RV comme au BO, et constituer un lien entre les deux théâtres principaux de l’apprentissage.

L’usage des échelles spécifiques dans le cadre de la certification est peu évalué. Le caractère très détaillé de l’échelle H-OSATS sera probablement limitant si elle est utilisée dans ce but. À l’heure actuelle, l’évaluation sommative classique reste celle du compagnonnage. Parmi les autres moyens disponibles, l’échelle globale OSATS est l’outil de référence pour l’évaluation des compétences techniques.⁵⁰

Le degré d'assistance donné à l'interne pendant la procédure n'a volontairement pas été intégré à l'échelle H-OSATS, afin que celle-ci soit aussi objective et descriptive que possible. Elle pourra ainsi être utilisée à partir d'un support vidéo d'une intervention simulée ou enregistrée au BO. Le degré d'instruction et d'aide fournie par le sénior pendant l'intervention est bien sûr un élément influant beaucoup sur les performances opératoires. Il peut être évalué par un outil complémentaire comme la Zwisch scale.⁵⁴ Cette échelle qui comprend 4 niveaux (démonstration, aide active, aide passive et capacité opératoire sans aide) est un outil proposé pour la formation et pour la certification. Cependant les limites de ce type d'outil pour juger les internes rejoignent celles du compagnonnage classique par leur importante subjectivité et leur évaluation indirecte de la performance technique, ce qui est source de biais.⁵⁰

La difficulté du cas réalisée n'est pas abordée dans l'échelle H-OSATS. Le degré de difficulté d'une intervention est lié à de multiples critères : anatomie, antécédents chirurgicaux de la patiente, efficacité du matériel utilisé et qualité des aides. Ces éléments sont difficiles à intégrer à un outil d'évaluation pédagogique. L'évaluation de l'échelle H-OSATS pour les cas d'HC « simple », dans une optique de formation (et éventuellement d'évaluation sommative) est une première étape indispensable avant d'explorer le champ des situations plus complexes. Toutefois la description (de 1 à 5) correspondant aux items retenus par le groupe Delphi ne fait volontairement pas référence à des conditions opératoires particulières (à l'exception d'un item concernant la visualisation de l'uretère en cas de distorsions anatomiques). Ainsi la validité de H-OSATS pour évaluer des cas complexes d'HC pourra être recherchée à l'avenir.

7.4 Perspectives

7.4.1 Étude de la validité de l'échelle H-OSATS pour évaluer les compétences au BO

L'échelle H-OSATS développée à l'occasion de cette thèse a montré des preuves de validité lors de l'évaluation de 3 groupes d'opérateurs d'expérience différente sur le LapMentor*, ce qui permet d'envisager son utilisation dans le cadre de la formation sur simulateur de RV. Cependant, H-OSATS a été construite pour une utilisation au BO et une phase de tests dans des conditions réelles est actuellement en cours. Cette partie de la recherche avait obtenu l'accord du comité d'éthique CEROG 2015-GYN-0801. Elle comprendra l'analyse des performances de 3 groupes d'opérateurs : un groupe d'internes en formation dans 2 services de gynécologie (CHU de Marseille et CHU de Montpellier) ayant validé le curriculum sur le pelvi-trainer FLS, un groupe de jeunes chirurgiens responsables (assistants ayant réalisé entre 2 et 20 procédures comme séniors) et un groupe d'opérateurs expérimentés ayant réalisé plus de 100 procédures. Les internes seront assistés par un senior instructeur qui leur procurera une aide passive et pourra intervenir pour la sécurité de la patiente comme dans le cadre du compagnonnage classique. Les performances des différents temps opératoires intra-péritonéaux seront filmés. Les vidéos seront anonymées et évaluées par deux opérateurs qualifiés.

7.4.2 Validation de l'impact du curriculum de RV sur l'habileté technique chirurgicale (HTC) pour l'HC

Une prochaine étude vise à évaluer l'impact du curriculum de VR pour l'HC qui est présenté dans le travail n°1. Il s'agira d'une étude randomisée contrôlée multicentrique. Les objectifs de cette étude seront les suivants : premièrement, déterminer si ce curriculum de VR a un impact positif sur l'HTC des internes lors d'une HC faite au BO ; deuxièmement déterminer si

cette formation modifie le coût de prise en charge des malades. La durée opératoire, la morbidité postopératoire selon la classification de Dindo (jusqu'à 30 jours postopératoires) et la durée d'hospitalisation seront également mesurées.¹⁰¹ L'objectif secondaire fera l'objet d'une analyse médico-économique. Les participants seront des internes réalisant un semestre en chirurgie gynécologique et les patientes seront celles programmées pour subir une HC simple, avec ou sans conservation annexielle. Les critères d'inclusion seront pour les internes d'avoir réalisé plus de 5 procédures cœlioscopiques « basiques » et moins de 5 HCs en tant qu'opérateur principal ; pour les patientes les critères d'inclusion seront d'être opérées d'une HC jugée simple, quelle que soit l'indication (adénomyose, cancer ...). Trente internes seront recrutés et recevront une formation cognitive au sujet de l'HC puis une formation sur pelvi-trainer classique. Ils seront ensuite randomisés au moyen d'enveloppes scellées après évaluation de leurs compétences techniques de base lors d'une HC virtuelle sur simulateur LapMentor™. Le premier groupe réalisera la formation classique seule par compagnonnage (groupe contrôle). Le second groupe réalisera le curriculum de RV pour l'HC sur le LapMentor, en plus de la formation classique. Le troisième groupe réalisera le curriculum en plus de la formation classique, mais aura une intervention entre les sessions de simulateur de RV selon le principe de la pratique délibérée en utilisant l'échelle H-OSATS. À l'issue de ce programme, tous les internes seront évalués lors de deux HC au BO (Fig H). Un enregistrement vidéo des interventions sera réalisé. 4 temps opératoires seront imposés et évalués en aveugle par deux observateurs indépendants : la dissection-section du ligament lombo-ovarien ou du ligament utéro-ovarien, le décollement vésico-utérin, la dissection-section de l'artère utérine gauche et la suture vaginale. Lors de ces temps, l'interne évalué sera opérateur principal et le chirurgien sénior aide opératoire. Le chirurgien sénior pourra fournir des conseils à l'interne évalué et reprendre sa place d'opérateur à tout moment s'il le juge nécessaire pour la sécurité

de la patiente. Cette partie de procédure sera évaluée au moyen d'une échelle globale (OSATS). Le reste de l'opération sera pratiqué par l'interne ou le chirurgien sénior selon le choix de ce dernier. La durée opératoire et le pourcentage de temps durant lequel l'interne est opérateur seront également quantifiés. Les patientes seront suivies de la veille de l'intervention à 30 jours postopératoires, ou à la fin de leur hospitalisation si cette dernière dépasse 30 jours. Les données démographiques, préopératoires, peropératoires et postopératoires précoces seront collectées à partir des dossiers des patientes après leur accord. Les données postopératoires tardives seront obtenues auprès de la patiente soit lors de la consultation postopératoire, soit par téléphone (Fig I).

Figure H : Protocole de l'étude randomisée.

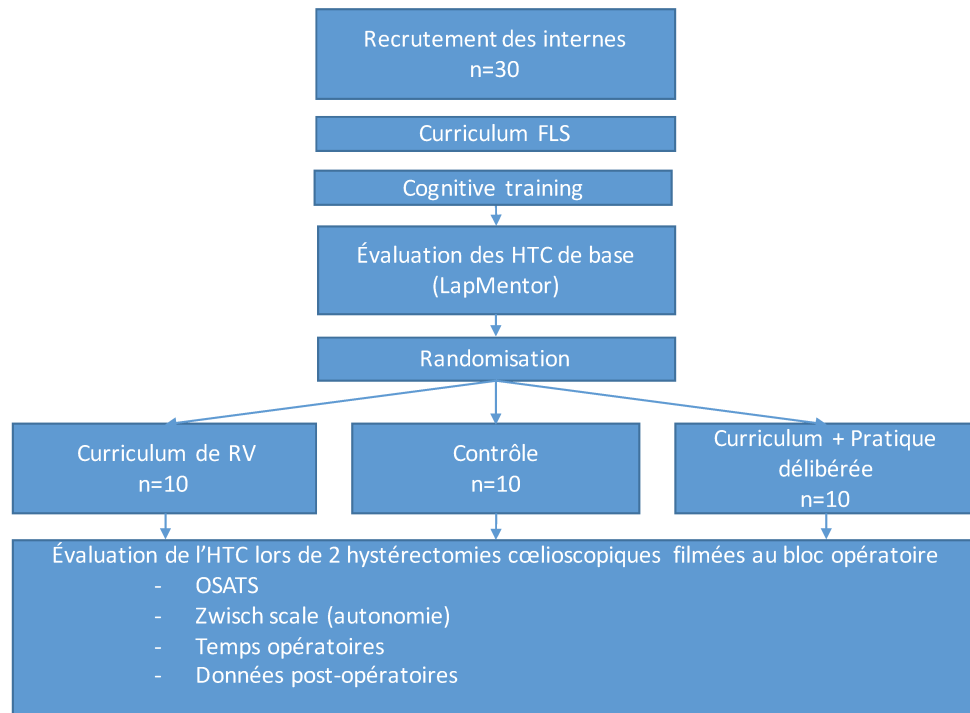
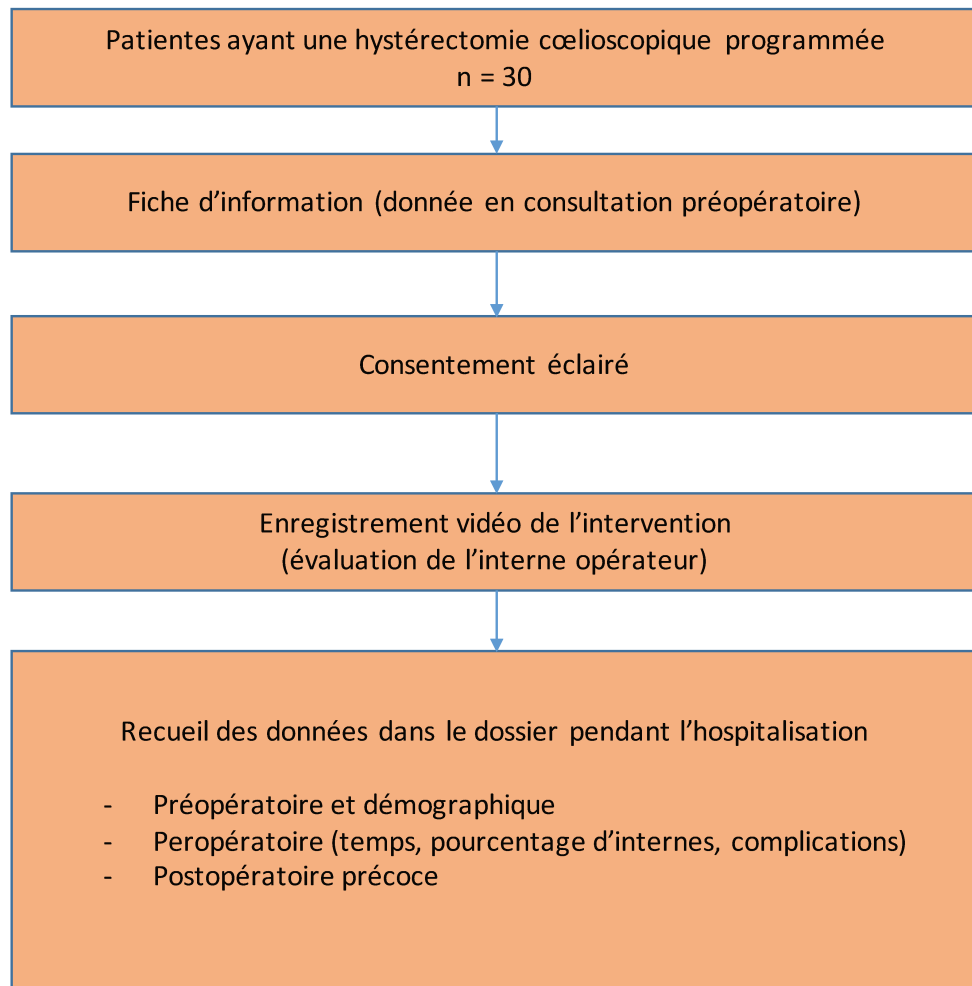


Figure I: Protocole de suivi des patientes de l'étude randomisée



9. RÉFÉRENCES BIBLIOGRAPHIQUES

1. Reznick RK, MacRae H. Teaching surgical skills--changes in the wind. *N Engl J Med*. 2006 Dec; 355(25):2664–9.
2. Nagendran M, Gurusamy KS, Aggarwal R, Loizidou M, Davidson BR. Virtual reality training for surgical trainees in laparoscopic surgery. *Cochrane Database Syst Rev*. 2013 Aug;(8): CD006575.
3. Osborne MP. William Stewart Halsted: his life and contributions to surgery. *Lancet Oncol*. 2007 Mar; 8(3):256–65.
4. Timberlake MD, Mayo HG, Scott L, Weis J, Gardner AK. What Do We Know About Intraoperative Teaching?: A Systematic Review. *Ann Surg*. 2017; doi: 10.1097/SLA.0000000000002131
5. Temple J. Resident duty hours around the globe: where are we now? *BMC Med Educ*. 2014; 14 Suppl 1:S8.
6. Jensen AR, Wright AS, Kim S, Horvath KD, Calhoun KE. Educational feedback in the operating room: a gap between resident and faculty perceptions. *Am J Surg*. 2012; 204(2):248–55.
7. Crochet P, Aggarwal R, Berdah S, Yaribakht S, Boubli L, Gamera M, et al. Current and future use of surgical skills simulation in gynecologic resident education: a French national survey. *J Gynecol Obstet Biol Reprod (Paris)*. 2014; 43(5):379–86.
8. Van Sickle KR, Ritter EM, Baghai M, Goldenberg AE, Huang I-P, Gallagher AG, et al. Prospective, randomized, double-blind trial of curriculum-based training for intracorporeal suturing and knot tying. *J Am Coll Surg*. 2008 Oct; 207(4):560–8.
9. Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Acad Med J Assoc Am Med Coll*. 2004; 79(10 Suppl):S70-81.
10. Crochet P, Aggarwal R, Dubb SS, Ziprin P, Rajaretnam N, Grantcharov T, et al. Deliberate practice on a virtual reality laparoscopic simulator enhances the quality of surgical technical skills. *Ann Surg*. 2011; 253(6):1216–22.
11. Hashimoto DA, Sirimanna P, Gomez ED, Beyer-Berjot L, Ericsson KA, Williams NN, et al. Deliberate practice enhances quality of laparoscopic surgical performance in a

- randomized controlled trial: from arrested development to expert performance. *Surg Endosc.* 2015; 29(11):3154–62.
12. Palter VN, Grantcharov TP. Individualized deliberate practice on a virtual reality simulator improves technical performance of surgical novices in the operating room: a randomized controlled trial. *Ann Surg.* 2014; 259(3):443–8.
 13. Ahmed K, Khan R, Mottrie A, Lovegrove C, Abaza R, Ahlawat R, et al. Development of a standardised training curriculum for robotic surgery: a consensus statement from an international multidisciplinary group of experts. *BJU Int.* 2015 Jul; 116(1):93–101.
 14. Palter VN. Comprehensive training curricula for minimally invasive surgery. *J Grad Med Educ.* 2011; 3(3):293–8.
 15. Moulton C-AE, Dubrowski A, Macrae H, Graham B, Grober E, Reznick R. Teaching surgical skills: what kind of practice makes perfect?: a randomized, controlled trial. *Ann Surg [Internet].* 2006; 244(3):400–9.
 16. Stefanidis D, Heniford BT. The formula for a successful laparoscopic skills curriculum. *Arch Surg (Chicago, Ill 1960).* 2009; 144(1):77–82.
 17. Strandbygaard J, Bjerrum F, Maagaard M, Winkel P, Larsen CR, Ringsted C, et al. Instructor feedback versus no instructor feedback on performance in a laparoscopic virtual reality simulator: a randomized trial. *Ann Surg.* 2013;257(5):839–44.
 18. Singh P, Aggarwal R, Tahir M, Pucher PH, Darzi A. A randomized controlled study to evaluate the role of video-based coaching in training laparoscopic skills. *Ann Surg.* 2015; 261(5):862–9.
 19. Aggarwal R, Grantcharov TP, Darzi A. Framework for systematic training and assessment of technical skills. *J Am Coll Surg.* 2007; 204(4):697–705.
 20. Zevin B, Levy JS, Satava RM, Grantcharov TP. A consensus-based framework for design, validation, and implementation of simulation-based training curricula in surgery. *J Am Coll Surg.* 2012; 215(4):580–586.e3.
 21. Fraser SA, Klassen DR, Feldman LS, Ghitulescu GA, Stanbridge D, Fried GM. Evaluating laparoscopic skills: setting the pass/fail score for the MISTELS system. *Surg Endosc.* 2003; 17(6):964–7.
 22. Zendejas B, Ruparel RK, Cook DA. Validity evidence for the Fundamentals of Laparoscopic Surgery (FLS) program as an assessment tool: a systematic review. *Surg Endosc.* 2016; 30(2):512–20.

23. Simpson JS, Scheer AS. A Review of the Effectiveness of Breast Surgical Oncology Fellowship Programs Utilizing Kirkpatrick's Evaluation Model. *J Cancer Educ Off J Am Assoc Cancer Educ.* 2016;31(3):466–71.
24. Cook DA, Hatala R, Brydges R, Zendejas B, Szostek JH, Wang AT, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. *JAMA.* 2011; 306(9):978–88.
25. Zendejas B, Brydges R, Hamstra SJ, Cook DA. State of the evidence on simulation-based training for laparoscopic surgery: a systematic review. *Ann Surg.* 2013; 257(4):586–93.
26. Harrysson IJ, Cook J, Sirimanna P, Feldman LS, Darzi A, Aggarwal R. Systematic review of learning curves for minimally invasive abdominal surgery: a review of the methodology of data collection, depiction of outcomes, and statistical analysis. *Ann Surg.* 2014; 260(1):37–45.
27. Aggarwal R, Grantcharov T, Moorthy K, Milland T, Papasavas P, Dosis A, et al. An evaluation of the feasibility, validity, and reliability of laparoscopic skills assessment in the operating room. *Ann Surg.* 2007; 245(6):992–9.
28. Aggarwal R, Grantcharov TP, Eriksen JR, Blirup D, Kristiansen VB, Funch-Jensen P, et al. An evidence-based virtual reality training program for novice laparoscopic surgeons. *Ann Surg.* 2006; 244(2):310–4.
29. Igwe E, Hernandez E, Rose S, Uppal S. Resident participation in laparoscopic hysterectomy: impact of trainee involvement on operative times and surgical outcomes. *Am J Obstet Gynecol.* 2014; 211(5):484.e1-7.
30. Feldman LS, Hagarty SE, Ghitulescu G, Stanbridge D, Fried GM. Relationship between objective assessment of technical skills and subjective in-training evaluations in surgical residents. *J Am Coll Surg.* 2004 ;198(1):105–10.
31. Goff B, Mandel L, Lentz G, Vanblaricom A, Oelschlager A-MA, Lee D, et al. Assessment of resident surgical skills: is testing feasible? *Am J Obstet Gynecol.* 2005; 192(4):1331-1338-1340.
32. Moorthy K, Munz Y, Dosis A, Bello F, Chang A, Darzi A. Bimodal assessment of laparoscopic suturing skills: construct and concurrent validity. *Surg Endosc.* 2004;18(11):1608–12.
33. Eubanks TR, Clements RH, Pohl D, Williams N, Schaad DC, Horgan S, et al. An objective scoring system for laparoscopic cholecystectomy. *J Am Coll Surg.* 1999; 189(6):566–74.

34. Weizman NF, Manoucheri E, Vitonis AF, Hicks GJ, Einarsson JI, Cohen SL. Design and validation of a novel assessment tool for laparoscopic suturing of the vaginal cuff during hysterectomy. *J Surg Educ.* 2015; 72(2):212–9.
35. Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg.* 1997; 84(2):273–8.
36. Szasz P, Louridas M, Harris KA, Aggarwal R, Grantcharov TP. Assessing Technical Competence in Surgical Trainees: A Systematic Review. *Ann Surg.* 2015; 261(6):1046–55.
37. Aggarwal R, Grantcharov T, Moorthy K, Milland T, Darzi A. Toward feasible, valid, and reliable video-based assessments of technical surgical skills in the operating room. *Ann Surg [Internet].* 2008; 247(2):372–9.
38. Vassiliou MC, Feldman LS, Andrew CG, Bergman S, Leffondré K, Stanbridge D, et al. A global assessment tool for evaluation of intraoperative laparoscopic skills. *Am J Surg.* 2005; 190(1):107–13.
39. Regehr G, MacRae H, Reznick RK, Szalay D. Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination. *Acad Med J Assoc Am Med Coll.* 1998;73(9):993–7.
40. Ahmed K, Miskovic D, Darzi A, Athanasiou T, Hanna GB. Observational tools for assessment of procedural skills: a systematic review. *Am J Surg.* 2011; 202(4):469–480.e6.
41. Beard JD, Choksy S, Khan S, Vascular Society of Great Britain and Ireland. Assessment of operative competence during carotid endarterectomy. *Br J Surg.* 2007; 94(6):726–30.
42. Birkmeyer JD, Finks JF, O’Reilly A, Oerline M, Carlin AM, Nunn AR, et al. Surgical skill and complication rates after bariatric surgery. *N Engl J Med.* 2013; 369(15):1434–42.
43. Larsen CR, Grantcharov T, Schouenborg L, Ottosen C, Soerensen JL, Ottesen B. Objective assessment of surgical competence in gynaecological laparoscopy: development and validation of a procedure-specific rating scale. *BJOG.* 2008; 115(7):908–16.
44. Zevin B, Bonrath EM, Aggarwal R, Dedy NJ, Ahmed N, Grantcharov TP, et al. Development, feasibility, validity, and reliability of a scale for objective assessment of operative performance in laparoscopic gastric bypass surgery. *J Am Coll Surg.* 2013;

- 216(5):955-965-1031, 1033.
45. Tremblay C, Grantcharov T, Urquia ML, Satkunarathnam A. Assessment tool for total laparoscopic hysterectomy: a Delphi consensus survey among international experts. *J Obstet Gynaecol Canada JOGC = J d'obstétrique gynécologie du Canada JOGC*. 2014; 36(11):1014–23.
 46. Frederick PJ, Szender JB, Hussein AA, Kesterson JP, Shelton JA, Anderson TL, et al. Surgical Competency for Robot-Assisted Hysterectomy: Development and Validation of a Robotic Hysterectomy Assessment Score (RHAS). *J Minim Invasive Gynecol*. 2017; 1;24(1):55-61.
 47. Bonrath EM, Zevin B, Dedy NJ, Grantcharov TP. Error rating tool to identify and analyse technical errors and events in laparoscopic surgery. *Br J Surg*. 2013; 100(8):1080–8.
 48. Husslein H, Shirreff L, Shore EM, Lefebvre GG, Grantcharov TP. The Generic Error Rating Tool: A Novel Approach to Assessment of Performance and Surgical Education in Gynecologic Laparoscopy. *J Surg Educ*. 2015; 72(6):1259–65.
 49. Joice P, Hanna GB, Cuschieri A. Errors enacted during endoscopic surgery--a human reliability analysis. *Appl Ergon*. 1998; 29(6):409–14.
 50. Fecso AB, Szasz P, Kerezov G, Grantcharov TP. The Effect of Technical Performance on Patient Outcomes in Surgery: A Systematic Review. *Ann Surg*. 2017; 265(3):492-501.
 51. de Montbrun S, Satterthwaite L, Grantcharov TP. Setting pass scores for assessment of technical performance by surgical trainees. *Br J Surg*. 2016; 103(3):300–6.
 52. Jelovsek JE, Walters MD, Korn A, Klingele C, Zite N, Ridgeway B, et al. Establishing cutoff scores on assessments of surgical skills to determine surgical competence. *Am J Obstet Gynecol*. 2010; 203(1):81.e1-6.
 53. Miskovic D, Wyles SM, Carter F, Coleman MG, Hanna GB. Development, validation and implementation of a monitoring tool for training in laparoscopic colorectal surgery in the English National Training Program. *Surg Endosc*. 2011; 25(4):1136–42.
 54. George BC, Teitelbaum EN, Meyerson SL, Schuller MC, DaRosa DA, Petrusa ER, et al. Reliability, validity, and feasibility of the Zwisch scale for the assessment of intraoperative performance. *J Surg Educ*. 2014; 71(6):e90-96.
 55. Joint Committee on Standards for Educational and Psychological Testing APA, American Educational Research Association and National Council on Measurement in Education. Standards for educational and psychological testing. Washington, DC; 1999.

56. Stunkard AJ, Messick S. The three-factor eating questionnaire to measure dietary restraint, disinhibition and hunger. *J Psychosom Res.* 1985; 29(1):71–83.
57. Downing S. Validity: on the meaning ful interpretation of assessment data. *Med Educ.* 2003;37:830–7.
58. Korndorffer JR, Kasten SJ, Downing SM. A call for the utilization of consensus standards in the surgical education literature. *Am J Surg.* 2010; 199(1):99–104.
59. Ghaderi I, Manji F, Park YS, Juul D, Ott M, Harris I, et al. Technical skills assessment toolbox: a review using the unitary framework of validity. *Ann Surg.* 2015; 261(2):251–62.
60. Cook DA. Much ado about differences: why expert-novice comparisons add little to the validity argument. *Adv Heal Sci Educ Theory Pract.* 2015;20(3):829–34.
61. Bruhat MA, Manhes H, Mage G, Pouly JL. Treatment of ectopic pregnancy by means of laparoscopy. *Fertil Steril.* 1980; 33(4):411–4.
62. Way LW, Stewart L, Gantert W, Liu K, Lee CM, Whang K, et al. Causes and prevention of laparoscopic bile duct injuries: analysis of 252 cases from a human factors and cognitive psychology perspective. *Ann Surg.* 2003; 237(4):460–9.
63. Bharathan R, Setchell T, Miskry T, Darzi A, Aggarwal R. Gynecologic endoscopy skills training and assessment: review. *J Minim Invasive Gynecol.* 2014; 21(1):28–43.
64. Zheng B, Hur H-C, Johnson S, Swanström LL. Validity of using Fundamentals of Laparoscopic Surgery (FLS) program to assess laparoscopic competence for gynecologists. *Surg Endosc.* 2010; 24(1):152–60.
65. Molinas CR, Campo R. Retention of laparoscopic psychomotor skills after a structured training program depends on the quality of the training and on the complexity of the task. *Gynecol Surg.* 2016; 13(4):395–402.
66. Botchorishvili R, Rabischong B, Larraín D, Khoo CK, Gaia G, Jardon K, et al. Educational value of an intensive and structured interval practice laparoscopic training course for residents in obstetrics and gynecology: a four-year prospective, multi-institutional recruitment study. *J Surg Educ.* 2012; 69(2):173–9.
67. Schreuder HWR, van Dongen KW, Roeleveld SJ, Schijven MP, Broeders IAMJ. Face and construct validity of virtual reality simulation of laparoscopic gynecologic surgery. *Am J Obstet Gynecol.* 2009; 200(5):540.e1-8.
68. Schreuder HWR, van Hove PD, Janse JA, Verheijen RRM, Stassen LPS, Dankelman J. An

- “intermediate curriculum” for advanced laparoscopic skills training with virtual reality simulation. *J Minim Invasive Gynecol.* 2011;18(5):597–606.
69. Tunitsky-Bitton E, Propst K, Muffly T. Development and validation of a laparoscopic hysterectomy cuff closure simulation model for surgical training. *Am J Obstet Gynecol.* 2016; 214(3):392.e1-6.
 70. King CR, Donnellan N, Guido R, Ecker A, Althouse AD, Mansuria S. Development and Validation of a Laparoscopic Simulation Model for Suturing the Vaginal Cuff. *Obstet Gynecol.* 2015; 126 Suppl 4:27S–35S.
 71. Tunitsky-Bitton E, King CR, Ridgeway B, Barber MD, Lee T, Muffly T, et al. Development and validation of a laparoscopic sacrocolpopexy simulation model for surgical training. *J Minim Invasive Gynecol.* 2014; 21(4):612–8.
 72. Levine RL, Kives S, Cathey G, Blinchevsky A, Acland R, Thompson C, et al. The use of lightly embalmed (fresh tissue) cadavers for resident laparoscopic training. *J Minim Invasive Gynecol.* 2006; 13(5):451–6.
 73. Hart R, Doherty DA, Karthigasu K, Garry R. The value of virtual reality-simulator training in the development of laparoscopic surgical skills. *J Minim Invasive Gynecol [Internet].* 2006; 13(2):126–33.
 74. Tang B, Tait I, Ross G, Chien P. Development and use of a restructured animal tissue model for training in laparoscopic salpingostomy and salpingectomy. *J Minim Invasive Gynecol.* 2011; 18(6):785–91.
 75. Hoffman MS. Simulation of robotic hysterectomy utilizing the porcine model. *Am J Obstet Gynecol.* 2012;206(6):523.e1-2.
 76. Hoffman MS, Ondrovic LE, Wenham RM, Apte SM, Shames ML, Zervos EE, et al. Evaluation of the porcine model to teach various ancillary procedures to gynecologic oncology fellows. *Am J Obstet Gynecol.* 2009; 201(1):116.e1-3.
 77. Bharathan R, Vali S, Setchell T, Miskry T, Darzi A, Aggarwal R. Psychomotor skills and cognitive load training on a virtual reality laparoscopic simulator for tubal surgery is effective. *Eur J Obstet Gynecol Reprod Biol.* 2013;169(2):347–52.
 78. Aggarwal R, Tully A, Grantcharov T, Larsen CR, Miskry T, Farthing A, et al. Virtual reality simulation training can improve technical skills during laparoscopic salpingectomy for ectopic pregnancy. *BJOG an Int J Obstet Gynaecol.* 2006; 113(12):1382–7.
 79. Larsen CR, Grantcharov T, Aggarwal R, Tully A, Sørensen JL, Dalsgaard T, et al. Objective

- assessment of gynecologic laparoscopic skills using the LapSimGyn virtual reality simulator. *Surg Endosc.* 2006; 20(9):1460–6.
80. Banks EH, Chudnoff S, Karmin I, Wang C, Pardanani S. Does a surgical simulator improve resident operative performance of laparoscopic tubal ligation? *Am J Obstet Gynecol.* 2007;197(5):541.e1-5.
 81. Gala R, Orejuela F, Gerten K, Lockrow E, Kilpatrick C, Chohan L, et al. Effect of validated skills simulation on operating room performance in obstetrics and gynecology residents: a randomized controlled trial. *Obstet Gynecol.* 2013; 121(3):578–84.
 82. Shore EM, Lefebvre GG, Husslein H, Bjerrum F, Sorensen JL, Grantcharov TP. Designing a Standardized Laparoscopy Curriculum for Gynecology Residents: A Delphi Approach. *J Grad Med Educ.* 2015; 7(2):197–202.
 83. Shore EM, Grantcharov TP, Husslein H, Shirreff L, Dedy NJ, McDermott CD, et al. Validating a standardized laparoscopy curriculum for gynecology residents: a randomized controlled trial. *Am J Obstet Gynecol.* 2016;215(2):204.e1-204.e11.
 84. Larsen CR, Soerensen JL, Grantcharov TP, Dalsgaard T, Schouenborg L, Ottosen C, et al. Effect of virtual reality training on laparoscopic surgery: randomised controlled trial. *BMJ.* 2009; 338:b1802.
 85. Mäkinen J, Johansson J, Tomás C, Tomás E, Heinonen PK, Laatikainen T, et al. Morbidity of 10 110 hysterectomies by type of approach. *Hum Reprod.* 2001;16(7):1473–8.
 86. Terzi H, Biler A, Demirtas O, Guler OT, Peker N, Kale A. Total laparoscopic hysterectomy: Analysis of the surgical learning curve in benign conditions. *Int J Surg.* 2016; 35:51–7.
 87. Palter VN, Orzech N, Reznick RK, Grantcharov TP. Validation of a structured training and assessment curriculum for technical skill acquisition in minimally invasive surgery: a randomized controlled trial. *Ann Surg.* 2013; 257(2):224–30.
 88. Shetty S, Zevin B, Grantcharov TP, Roberts KE, Duffy AJ. Perceptions, training experiences, and preferences of surgical residents toward laparoscopic simulation training: a resident survey. *J Surg Educ.* 2014; 71(5):727–33.
 89. Våpenstad C, Hofstad EF, Langø T, Mårvik R, Chmarra MK. Perceiving haptic feedback in virtual reality simulators. *Surg Endosc.* 2013; 27(7):2391–7.
 90. Chan B, Martel G, Poulin EC, Mamazza J, Boushey RP. Resident training in minimally invasive surgery: a survey of Canadian department and division chairs. *Surg Endosc.* 2010; 24(3):499–503.

91. Bridges M, Diamond DL. The financial impact of teaching surgical residents in the operating room. *Am J Surg.* 1999; 177(1):28–32.
92. Koperna T. How long do we need teaching in the operating room? The true costs of achieving surgical routine. *Langenbeck's Arch Surg.* 2004; 389(3):204–8.
93. Kovacs G. Procedural skills in medicine: linking theory to practice. *J Emerg Med.* 1997; 15(3):387–91.
94. Madani A, Watanabe Y, Bilgic E, Pucher PH, Vassiliou MC, Aggarwal R, et al. Measuring intra-operative decision-making during laparoscopic cholecystectomy: validity evidence for a novel interactive Web-based assessment tool. *Surg Endosc.* 2017; 31(3):1203-1212
95. Back DA, Haberstroh N, Antolic A, Sostmann K, Schmidmaier G, Hoff E. Blended learning approach improves teaching in a problem-based learning environment in orthopedics - a pilot study. *BMC Med Educ.* 2014; 14:17.
96. Rowe M, Frantz J, Bozalek V. The role of blended learning in the clinical education of healthcare students: a systematic review. *Med Teach.* 2012; 34(4):e216-221.
97. Arora S, Sevdalis N, Nestel D, Woloshynowych M, Darzi A, Kneebone R. The impact of stress on surgical performance: a systematic review of the literature. *Surgery.* 2010;147(3):318–30, 330-6.
98. Pucher PH, Aggarwal R, Darzi A. Surgical ward round quality and impact on variable patient outcomes. *Ann Surg.* 2014; 259(2):222–6.
99. Beyer-Berjot L, Patel V, Acharya A, Taylor D, Bonrath E, Grantcharov T, et al. Surgical training: design of a virtual care pathway approach. *Surgery.* 2014; 156(3):689–97.
100. Roberts NK, Williams RG, Kim MJ, Dunnington GL. The briefing, intraoperative teaching, debriefing model for teaching in the operating room. *J Am Coll Surg.* 2009;208(2):299–303.
101. Dindo D, Demartines N, Clavien P-A. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann Surg.* 2004; 240(2):205–13.

10. ANNEXES

10.1 Structure pour développer un curriculum basé sur la simulation, d'après Zevin et al.

10.2 Utilisation des simulateurs pour former les internes de chirurgie gynécologique en France : un état des lieux en 2013

SUPPLEMENTARY APPENDIX

An evidence-based framework for the design of a simulation-based surgical training curriculum *Predevelopment analysis*

1. Perform a needs assessment/define the problem
 - a. Determine the goals and objectives for the curriculum
 - b. Determine the desired learning outcomes, measures and proficiencies to be achieved
 - i. Include governing authorities—Boards, Societies, ACGME, etc, when possible
 - c. Determine the common errors that need to be taught
 - d. Define the population of interest (novice, intermediate, advanced trainee, expert surgeon)
 - e. Include program director, faculty, educator, user (eg, resident, medical student, etc) and boards
2. Determine the available resources at your institution, society, etc (optional)
 - a. Knowledge content
 - b. Physical space
 - c. Equipment (simulators? types? high or low fidelity?)
 - d. Personnel including faculty with formal educator expertise (ie, fellowship, masters, or PhD)
 - e. Financial resources
 - f. Support of faculty members, chair, chief, and/or program-director

Curriculum development

Cognitive component (focused on specific technical issues)

1. Pretest of procedure-specific knowledge
 - a. Written test format
 - i. Multiple choice
 - b. Interactive, graphic-based image software that delivers procedural training
2. Preprocedure patient assessment and preparation
 - a. Review of indications/contraindications for surgery
 - b. Teamwork and environment considerations (functioning equipment, instruments, suture/stapling devices, etc.)
3. Procedure-specific operative knowledge
 - a. Key steps of the procedure
 - b. Anatomy: descriptions of important structures and tissue planes

- c. Instruments: identify types, indications, safety and limitations of instruments
- d. Psychomotor skills: demonstration of fundamental technical skills needed to complete the procedure
- e. Errors (most common and most critical though infrequent):
 - i. Anticipation/avoidance
 - ii. Recognition
 - iii. Management
- f. Postoperative complications (potential and actual):
 - i. Knowledge of
 - ii. Anticipation/avoidance
 - iii. Recognition and management
- g. Inclusion of video examples of failures/disasters
4. Faculty interaction: time for questions to the expert; answers and reflections by trainees
5. Post-test of learned procedure-specific knowledge
 - a. Written test format
 - i. Multiple choice questions
 - ii. Interactive, graphic-based image software that delivers procedural training
 - b. Mock operative note dictation
6. Delivery of the cognitive component
 - a. Procedure-specific video-based tutorials with pauses in the video at critical steps of the procedure
 - b. Interactive software that delivers procedural training
 - i. Interactive graphic-based image of procedure has to be identical to psychomotor skill simulator (this is critical)
 - c. Progress to the psychomotor skills component of the training is not permitted until cognitive part is passed, preferably with 100%.

Psychomotor skills (technical component)

1. Deconstruction of the procedure of interest into its component tasks/steps
 - a. Acquire videorecordings of different approaches of experts performing the same procedure of interest
 - i. Include only the tasks in common from multiple different approaches
 - b. Perform a hierarchical tasks analysis (HTA) of the procedure (ideally should include experts in surgical education, behavioral psychologists, expert clinicians)
 - i. Two or more independent reviewers to watch the video-recordings and write down all the tasks/steps of the procedure

- ii. Establish consensus between reviewers on the tasks/steps
 - iii. Each tasks/steps should have a clear and unambiguous definition of the task/subtask, start point and end point, and quantitative measure (if possible) or clearly defined Likert Scale
 - iv. By reviewing the video-recordings of novices, identify which tasks/steps are the most difficult for a novice to perform and what are the most common errors
2. Create a procedure-specific assessment tool using the identified tasks/steps/errors
 - a. Use Delphi methodology to identify the tasks/steps/errors from the HTA to be included in the technical skills assessment tool
 - b. Establish reliability of the tool
 - c. Establish validity of the tool
 - d. Demonstrate construct validity by demonstrating a difference in scores of novice surgeons (<10 procedures), intermediates (10 to 100 procedures), and experts (>100 procedures). The number of procedures at each level can vary significantly depending on the type of procedure.
 - e. Establish the benchmark criteria for competency (Experts' mean \pm 1 SD)
 - i. Analyze experts' learning curves (stop when 2 consecutive trials show no improvement, ie, they have reached a plateau)
 - ii. Establish the mean of all the experts' plateaus
 - iii. Establish the mean \pm 1 SD of the mean of all the experts
 - iv. The benchmark of competence is experts' mean minus 1 SD
 - f. Define a cutoff value to be used in formative assessment (from the errors list)
 - g. Define a cutoff value to be used in summative assessment to define proficiency (from the benchmark criteria)
3. Conduct initial assessment of trainees technical skill using the procedure-specific assessment tool
4. Identify existing simulation models for training in the key tasks/steps of the procedure
 - a. If no procedure-specific simulator exists, determine how to develop the desired simulator
 - b. Basic skills training
 - i. Synthetic models
 - ii. Low fidelity computer-based models
 - iii. Aim to select models with embedded objective metrics
 - c. Advanced (procedure-specific) skills training
 - i. Synthetic models
 - ii. Computer-based models
 - iii. Cadaveric/tissue models
 - iv. Live animal models
 - d. Confirm or define validity for the chosen models
 - i. Face
 - ii. Content
 - iii. Concurrent
 - iv. Construct
 - v. Predictive
5. Define different levels of difficulty for each simulation model
6. Establish expert proficiency benchmarks for each simulation model (see item 2e)
 - a. Using measurement parameters imbedded in the model
 - b. Using the previously developed procedure-specific assessment tool
7. Set acceptable level of proficiency as the "pass" score for each model in the curriculum (see item 2e)
 - a. Forced proficiency at each level before progression to the next level (learn basic skills before move on to advanced levels)
8. Define a specific practice schedule for trainees
 - a. Distributed practice
 - b. Deliberate practice
 - c. Variability of practice (variable levels of difficulty)
 - d. Practice until acceptable level of proficiency is achieved
 - e. Define amount of overtraining (if needed)
9. Provide extrinsic feedback (during practice sessions) to trainees on their performance
 - a. Formative
 - i. Immediate computer generated feedback
 - ii. Immediate, face-to-face feedback by expert instructors with alternative approaches demonstrated and practiced in the presence of the instructor
 - b. Summative
 - i. Review areas of weakness based on the procedure-specific assessment tool

- ii. Review of performance on video

Team-based component (supports technical aspect of procedure)

1. Develop a module for teaching nontechnical skills (teamwork, communication, situation awareness, etc) as related to the procedure of interest
 - a. Identify the roles and responsibilities of allied health care team (individual and team training)
 - b. Develop crisis-based scenarios for learning nontechnical skills in a simulated environment (appropriate for level of learner and of incremental difficulty)
 - c. Use available assessment tools for formative evaluation of nontechnical skills (evaluate the reliability, validity and applicability of instruments; if not available develop new tools)
 - a. Use an existing team training model (eg, Team-STEPPS) if available
 - b. Incorporate train-the-trainer programs for teaching faculty
 - i. Provide specific training for debriefing skills

Curriculum validation, evaluation, and improvement Validation

1. Conduct a randomized controlled trial to demonstrate transfer of learned skills to the real environment
 - a. Compare simulator trained group with traditionally trained group

- a. Demonstrate difference in cognitive knowledge and technical skill
- b. Evaluate the impact of simulation-based training on the learning curves in the operating room
- c. Calculate the transfer-effectiveness ratio for the training curriculum

Evaluation and improvement

1. Prospectively collect data on curriculum validity
 - a. Pre-post test learner performance data comparison
 - b. Learner retention data
 - c. Learner feedback
 - d. Impact of curriculum on clinical outcomes
2. Perform periodic evaluations of curriculum and adjust curriculum based on feedback/experience
 - a. Identify strengths and weaknesses and areas for improvement
 - b. Develop methods for continuous program improvement

Maintenance of training

1. Define post-training intervals at which cognitive knowledge and technical skills tests will be administered to assure ongoing proficiency of the learner
2. Develop a specific "retraining" curriculum, based on principles listed. Generally, this will be less extensive and more focused than the full curriculum.
3. Include retraining to proficiency during testing sessions (maintenance practice) with formative feedback.
4. Study skill degradation for particular tasks.

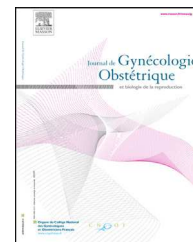


Disponible en ligne sur

ScienceDirect
www.sciencedirect.com

Elsevier Masson France

EM|consulte
www.em-consulte.com



TRAVAIL ORIGINAL

Utilisation des simulateurs pour former les internes de chirurgie gynécologique en France : un état des lieux en 2013



Current and future use of surgical skills simulation in gynecologic resident education: A French national survey

P. Crochet^{a,*}, R. Aggarwal^b, S. Berdah^c, S. Yaribakht^d,
L. Boubli^e, M. Gamberre^a, A. Agostini^a

^a Service de gynécologie obstétrique, hôpital de La Conception, AP–HM, 147, boulevard Baille, 13385 Marseille cedex 05, France

^b Department of surgery, Perelman school of medicine, university of Pennsylvania, 295, John Morgan Building, 3620 Hamilton Walk, Philadelphia, PA 19104, États-Unis

^c Service de chirurgie digestive, CERC (Centre d'enseignement et de recherche chirurgical), faculté de médecine nord, université Aix-Marseille, boulevard Pierre-Dramard, 13916 Marseille, France

^d AP–HP, 75000 Paris, France

^e Service de gynécologie obstétrique, hôpital Nord, AP–HM, chemin des Bourrely, 13015 Marseille, France

Reçu le 8 août 2013 ; avis du comité de lecture le 22 septembre 2013 ; définitivement accepté le 1^{er} octobre 2013

Disponible sur Internet le 12 décembre 2013

MOTS CLÉS

Simulation ;
Gynécologie ;
Éducation ;
Questionnaire

Résumé

But. – La simulation est une méthode prometteuse pour améliorer l'enseignement en chirurgie gynécologique. Cette enquête réalisée auprès des internes et des enseignants français fait un état des lieux de l'utilisation des simulateurs en 2013.

Matériels et méthodes. – Deux questionnaires spécifiques ont été envoyés aux internes et aux enseignants. Les questions concernaient les types de simulateurs utilisés, les modes d'utilisation et l'opinion des intéressés à propos de ces outils pédagogiques.

Résultats. – Vingt-six pour cent des internes interrogés ont répondu (258/998) et 23 % des enseignants (29/122). Soixante-cinq pour cent (167/258) des internes ont déjà utilisé un simulateur. Les simulateurs les plus utilisés sont des Pelvi-trainers de coelioscopie (84 %) et les séances sur modèle porcine vivante (63 %). Le mode d'utilisation le plus fréquent est la séance d'initiation (51 %) et les journées de formation universitaire (38 %). Les internes jugent les simulateurs très utiles pour la formation. Les enseignants sont favorables à une utilisation obligatoire à l'avenir pour la formation mais sont moins favorables concernant la certification.

* Auteur correspondant.

Adresses e-mail : patricecrochet.marseille@gmail.com, patrice.crochet@ap-hm.fr (P. Crochet).

KEYWORDS

Simulation;
Gynecology;
Education;
Questionnaire

Conclusion. – Une expérience sur simulateur est déjà une réalité pour une majorité des internes ayant répondu au questionnaire. Cependant, les types de simulateurs utilisés varient beaucoup selon les centres et l'accès aux simulateurs est majoritairement ponctuel. L'opinion des interrogés est très favorable à leur utilisation pour la formation chirurgicale.

© 2013 Elsevier Masson SAS. Tous droits réservés.

Summary

Objectives. – Simulation is a promising method to enhance surgical education in gynecology. The purpose of this study was to provide baseline information on the current use of simulators across French academic schools.

Materials and methods. – Two questionnaires were created, one specifically for residents and one for professors. Main issues included the type of simulators used and the kind of use made for training purposes. Opinions and agreement about the use of simulators were also asked.

Results. – Twenty-six percent of residents (258/998) and 24% of professors (29/122) answered the questionnaire. Sixty-five percent of residents (167/258) had experienced simulators. Laparoscopic pelvic-trainers (84%) and sessions on alive pigs (63%) were most commonly used. Residents reported access to simulators most commonly during introductory sessions (51%) and days of academic workshops (38%). Residents believed simulators very useful for training. Professors agreed that simulators should become a required part of residency training, but were less enthusiastic regarding simulation becoming a part of certification for practice.

Conclusion. – Surgical skills simulators are already experienced by a majority of French gynecologic residents. However, the use of these educational tools varies among surgical schools and remains occasional for the majority of residents. There was a strong agreement that simulation technology should be a component of training.

© 2013 Elsevier Masson SAS. All rights reserved.

Introduction

Les pratiques de simulation se sont considérablement développées au cours des 20 dernières années dans le domaine de la santé [1]. Elles présentent de nombreux avantages pour améliorer la qualité et la sécurité des soins, comme le rappelle un récent rapport de l'HAS [2]. L'entraînement en laboratoire s'applique tout particulièrement au domaine de la chirurgie, où la réduction des opportunités opératoires et les considérations éthiques ne permettent plus de considérer le compagnonnage au bloc comme seule méthode de formation [3]. La simulation peut concerner l'aspect technique ou bien non technique des compétences chirurgicales [3]. La chirurgie gynécologique fait partie des spécialités où les techniques de simulation se sont le plus développées, dans le sillage de la révolution coelioscopique. L'impact de la simulation pour améliorer les compétences chirurgicales au bloc opératoire a déjà été démontré, notamment pour la chirurgie coelioscopique annexielle [4] même si de nombreuses techniques sont encore en cours d'évaluation.

La plupart des techniques de simulation sont connues et utilisées depuis plusieurs années par certains centres français [5]. Cependant il y a de grandes disparités sur le territoire concernant la diffusion de ces outils, et la mise en place de programmes de formation en laboratoire reste le fruit d'initiatives locales. En l'absence de politique organisée dans ce domaine, il n'y a pas d'informations disponibles sur l'utilisation réelle qui est faite des simulateurs pour la chirurgie gynécologique à l'heure actuelle et sur l'accès qui est réservé aux internes.

Le but de cette étude était de faire un état des lieux de l'utilisation des simulateurs pour la formation aux compétences techniques chirurgicales dans les CHU français du

point de vue des internes et du corps enseignant. L'opinion des intéressés a également été recueillie concernant la place que devrait occuper la simulation dans la formation à l'avenir.

Matériels et méthodes

Deux questionnaires spécifiques ont été créés pour cette étude, destinés aux internes en formation et aux enseignants français – professeurs des universités (PU) et maître de conférences universitaires (MCU).

Les questionnaires ont été diffusés par courriel en février 2013 (suivi d'un rappel). Le questionnaire destiné aux enseignants a été envoyé à tous les PU et MCU de gynécologie obstétrique des 29 CHU français. Le questionnaire destiné aux internes a été diffusé en collaboration avec l'association des internes de gynécologie obstétrique en formation (AGOF) qui nous a permis d'utiliser son répertoire de 988 internes, représentant donc la quasi-totalité des internes actuellement en cours de formation (1020-30).

Ces questionnaires étaient anonymes et comportaient des questions communes et des questions spécifiques. Les questions communes étaient : l'université d'origine, la notion d'une expérience sur simulateur pendant la formation, l'opinion personnelle concernant l'utilité des simulateurs pour l'apprentissage des gestes de base et des gestes plus complexes comme l'hystérectomie (selon une échelle de 1 « inutile » à 5 « très utile »).

Le questionnaire pour les internes comportait quatre questions spécifiques : l'année d'internat, l'expérience dans des conditions réelles en tant que premier opérateur sur 4 interventions spécifiques (grossesse extra utérine, hystérectomie vaginale, hystérectomie par laparotomie,

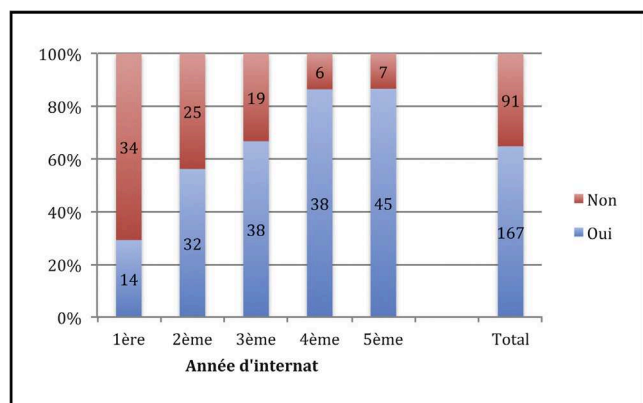


Figure 1 Réponse des internes à la question : «Au cours de votre formation en chirurgie gynécologique, avez vous eu un enseignement sur simulateur?».

Residents answers to the question : "During your surgical training, have you ever use a surgical simulator?"

hystérectomie coelioscopique), le type de simulateur utilisé et les conditions d'utilisation.

Le questionnaire des PU et MCU comportait des questions spécifiques : l'accès ou non à un outil de simulation pour les internes, le(s) type(s) de simulateur disponible(s), le lieu d'utilisation, les conditions d'utilisation, le mode de financement, l'opinion personnelle concernant l'utilisation obligatoire des simulateurs à l'avenir pour la formation et la certification des internes. Le détail des questionnaires est présenté en [Annexes 1 et 2](#).

Les résultats présentés sous forme de figure ont été réalisés à l'aide du logiciel tableur Excel 2011 (Microsoft) et du logiciel de statistique SPSS version 20,0 (SPSS Inc., Chicago, IL). Les résultats concernant les échelles d'opinion sont calculées sous forme de médiane et écart-type.

Résultats

Données démographiques

Deux cent cinquante-huit internes ont répondu sur 998 sollicités (26%). Le taux de réponse des internes en fonction de l'année de cursus était également réparti (19% en 1^{ère} année, 22% en 2^e année, 22% en 3^e année, 17% en 4^e année et 20% en 5^e année). Des réponses provenant de l'ensemble des 29 CHU français ont été recueillies.

Vingt-neuf praticiens universitaires ont répondu sur les 122 sollicités (23%). Les réponses provenaient de 18 des 29 CHU français.

Expérience sur simulateur

Concernant les internes, 65% (167/258) déclarent avoir déjà utilisé un simulateur au cours de leur formation. Le détail des réponses en fonction des années est présenté sur la [Fig. 1](#).

Concernant les enseignants, 17% (5/29) ont eu une expérience sur simulateur au cours de leur formation.

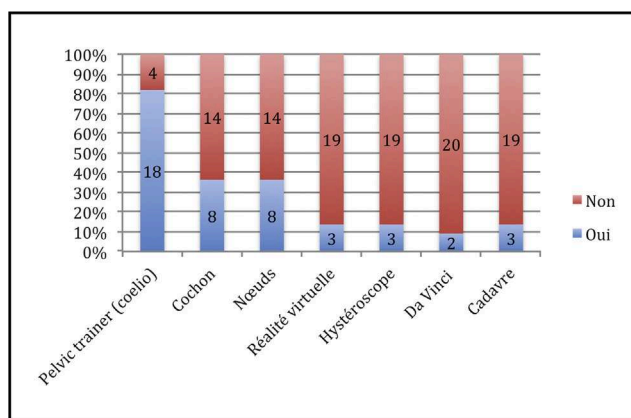


Figure 2 Réponse des enseignants à la question : «De quel type de simulateur disposez-vous?».

Consultants answers to the question : "What kind(s) of simulator(s) do you currently use in your laboratory?"

Structures disponibles (d'après réponses des enseignants)

Soixante-seize pour cent des enseignants déclarent disposer de matériel de simulation au sein de leur université. D'après ces réponses, il y a 6 des 18 CHU ayant répondu qui disposent de plus de deux types de simulateurs distincts. Le(s) lieu(x) où les simulateurs sont utilisés se situe(nt) dans 45% des cas au sein du service de chirurgie, dans 32% des cas dans un local de la faculté et dans 50% des cas dans un centre spécifique dédié. Les types de simulateurs utilisés sont détaillés sur la [Fig. 2](#). Les modes de simulation les plus fréquemment utilisés sont les *pelvi-trainer* coelioscopiques (82%), suivis des séances sur modèle porcin vivant anesthésié (36%) et les simulateurs pour réaliser des nœuds en laparotomie (36%). Les autres types de simulation sont mentionnés dans moins de 15% des cas (simulateurs de réalité virtuelle, simulateur robotique Da Vinci, séances sur cadavres).

Le mode d'utilisation des simulateurs est rapporté par les enseignants de la manière suivante : accès libre dans 73% des cas, organisation de séances d'initiations dans 64% des cas, programme développé pendant un semestre d'internat dans 32% des cas, organisation de journées de formation universitaires dans 18% des cas, utilisation à des fins de recherche dans 18% des cas et utilisation de curriculum dans 9% des cas. Le mode de financement du matériel de simulation déclaré par les enseignants est un mécénat ou financement par des laboratoires dans 50% des cas et un financement sur des crédits universitaires dans 32% des cas.

Expérience des internes

Le nombre d'interventions réalisées dans des conditions réelles en tant que premier opérateur a été demandé pour quatre interventions fréquentes en chirurgie gynécologique. Les résultats, en fonction de l'année d'internat, sont présentés sur la [Fig. 3](#).

Concernant les internes ayant eu accès aux techniques de simulation, les types de simulateur utilisés étaient : un *pelvi-trainer* coelioscopique dans 84% des cas, des séances

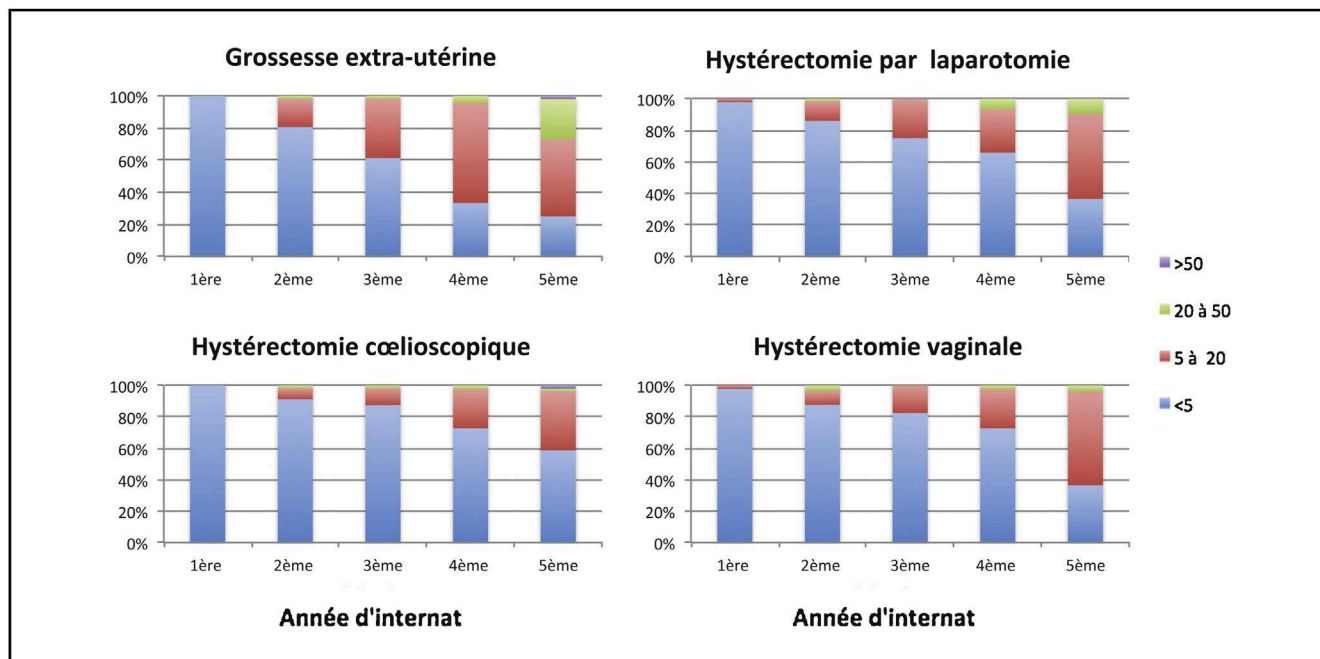


Figure 3 Expérience déclarée des internes en tant que premier opérateur pour 4 interventions (nombre de cas).
Number of surgical cases declared as first operator by the residents (regarding 4 gynecologic procedures).

sur modèle porcin vivant anesthésié dans 63% des cas, un simple simulateur pour faire des nœuds dans 53% des cas, un simulateur cœlioscopique de réalité virtuelle dans 25% des cas, un simulateur d'hystérocopie dans 21% des cas, un simulateur robotique Da Vinci dans 12% des cas, des séances sur cadavre humain dans 7% des cas.

Les deux modes d'utilisation des simulateurs les plus couramment rapporté par les internes sont les séances d'initiation (51%, 85/167) et les journées de formation universitaires (38%, 64/167). L'accès libre aux simulateurs est possible dans 34% des cas. La validation de curriculum est mentionnée dans 20% des cas, et la réalisation de programme tout au long d'un semestre d'internat dans 11% des cas.

Opinions

L'opinion des internes et des enseignants concernant l'utilité de la simulation pour l'apprentissage des gestes de bases et des gestes complexes (comme l'hystérectomie) est rapportée sur la Fig. 4. Pour les internes, la réponse médiane sur une échelle de 1 (inutile) à 5 (très utile) était de 5 ($\pm 0,8$) pour les gestes de base et de 5 ($\pm 0,856$) pour les gestes complexes. Pour les enseignants, la médiane des réponses était de 5 ($\pm 0,979$) pour les gestes de base et de 5 ($\pm 0,891$) pour les gestes complexes.

L'opinion des enseignants concernant une utilisation obligatoire à l'avenir de la simulation pour la formation des internes ainsi que la certification des internes (selon une échelle de 1 – pas du tout d'accord – à 5 – tout à fait d'accord) est présentée sur la Fig. 5. La médiane des réponses était de 5 ($\pm 0,942$) pour la formation et de 2 ($\pm 1,526$) pour la certification.

Discussion

Les techniques de simulation dans le domaine de la chirurgie gynécologique peuvent prendre différentes formes. Il y a de nombreuses déclinaisons possibles du « simulateur » en fonction de la voie d'abord (laparoscopique, vaginale, hystérocopie, chirurgie ouverte) et des procédés utilisés (modèle synthétique, modèle animal, réalité virtuelle, cadavres...). La diversité des outils disponibles ainsi que le caractère récent des recommandations dans ce domaine [2] expliquent qu'il n'y a pas eu jusqu'à présent d'évaluation globale de la situation concernant l'utilisation des simulateurs au cours de la formation des internes de gynécologie français.

Le premier enseignement de cette étude est la diffusion large des techniques de simulation puisque 76% des enseignants ayant répondu à cette enquête déclarent disposer d'au moins un simulateur. Ceci fait écho aux réponses des internes puisqu'une majorité d'entre eux déclarent avoir déjà eu accès à un simulateur au cours de leur cursus. Le taux de réponse des internes (26%) est élevé pour un questionnaire de ce type, ce qui souligne leur intérêt pour ce sujet [6]. À titre de comparaison, une enquête réalisée auprès des internes français au sujet de l'enseignement du siège en 2006 avait recueilli 19% de réponse [7]. Le taux de participation (23%) des enseignants est moins élevé. Ce plus faible taux de réponse peut être en partie expliqué par le fait que la chirurgie gynécologique n'est pas l'activité principale de certains enseignants qui ne se sont donc pas sentis concernés par cette enquête.

L'expérience des internes en tant que premier opérateur dans des conditions réelles a été évaluée sur quatre opérations majeures. C'est un témoignage du compagnonnage réalisé au bloc et des interventions réellement réalisées

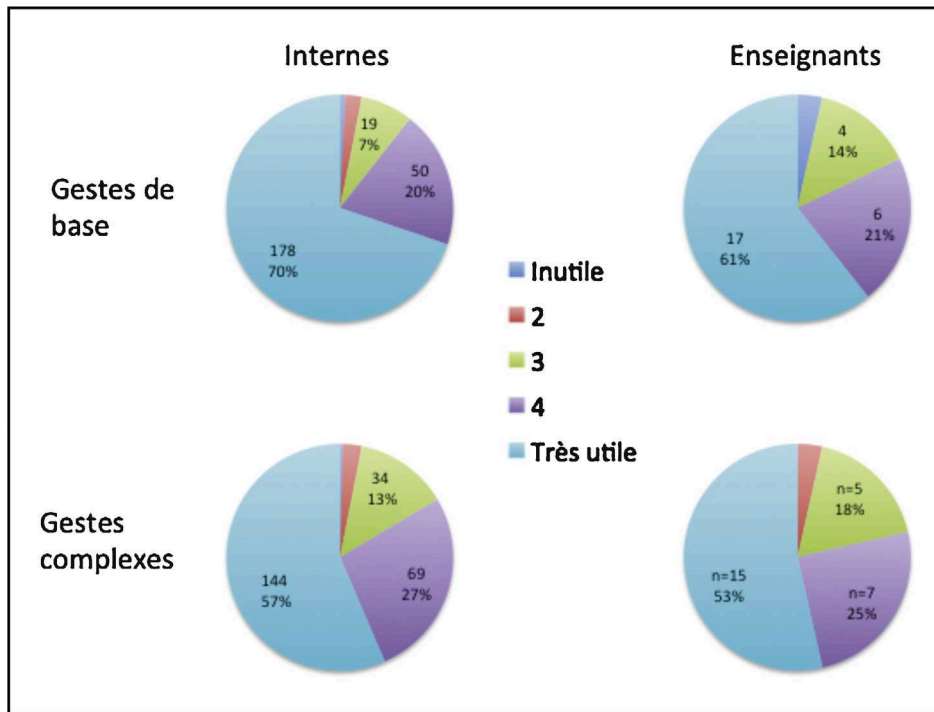


Figure 4 Opinion des enseignants et des internes concernant l'utilité de la simulation pour l'entraînement à des gestes simples et des gestes complexes (selon une échelle de 1 « non utile » à 5 « très utile »).

Consultants and residents opinion regarding the usefulness of simulators for basic skills and complexe procedures (on a scale between 1 "not useful" to 5 "very useful").

par les internes durant leur formation. On peut noter que le nombre cumulé d'interventions réalisées varie grandement à la fin du cursus. De plus, il semble y avoir moins d'opportunités opératoires concernant l'hystérectomie par

voie coelioscopique. Ceci est concordant avec les pratiques françaises puisque 20% environ des hystérectomies sont réalisées en utilisant la voie coelioscopique [8].

Si 76% des enseignants déclarent disposer d'au moins un simulateur, il y a une grande hétérogénéité des plateaux techniques disponibles. Les centres disposant d'un panel large restent une minorité de cas. Les simulateurs les plus utilisés (d'après les internes et les enseignants) sont les *pelvi-trainer* de coelioscopie ce qui est logique étant donné leur faible coût, leur facilité d'utilisation et leur efficacité [9]. L'entraînement sur tissus animal vivant dont le prix de revient est plus important est relativement développé en France car 8 centres déclarent y avoir accès. Les simulateurs utilisant le principe de la réalité virtuelle (coelio- ou hystéroscopique) ont l'avantage de permettre la répétition des même gestes et la réalisation de procédures complètes. Il sont cependant peu utilisés, probablement en raison de la nécessité d'un investissement de départ important. De plus, malgré les progrès récents de la modélisation informatique, ces simulateurs reproduisent imparfaitement les conditions opératoires sur tissu vivant.

La manière dont les simulateurs sont utilisés est une des clefs de l'efficacité de ces outils. Les enseignants ont déclaré laisser souvent les simulateurs en accès libre aux internes (73%), ce qui semble une réalité moins fréquente d'après les réponses des intéressés (34%). Pour une majorité d'internes, l'usage des moyens de simulation reste ponctuel dans le cadre de séances d'initiations. Dans un tiers des cas, cet accès est plus structuré dans le cadre de journées de formation. Il convient de noter que ces formations n'ont pas forcément été réalisées dans leur ville d'origine

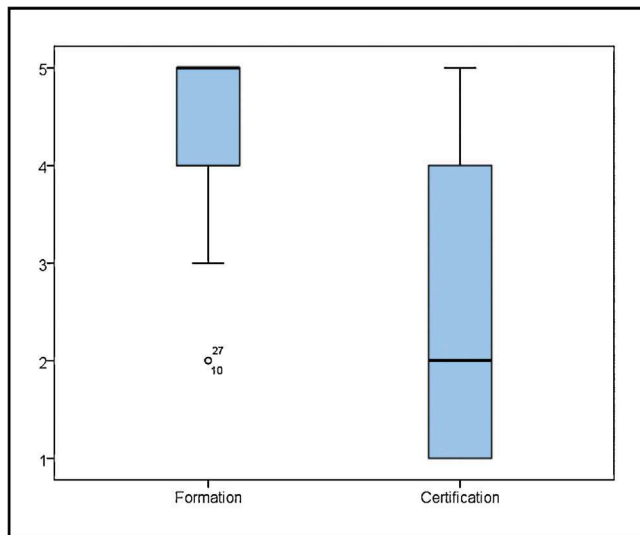


Figure 5 Opinion des enseignants concernant l'utilisation obligatoire de la simulation pour la formation et la certification à l'avenir (selon une échelle de 1 « non favorable » à 5 « très favorable »).

Consultants opinion regarding the future mandatory use of simulators for training and certification (on a scale between 1 "do not agree" to 5 "completely agree").

mais peuvent avoir eu lieu dans d'autres centres, comme par exemple dans le cadre de l'enseignement spécifique à la coelioscopie organisée pour les internes au CICE à Clermont-Ferrand depuis 2006 [5,10]. Ces formations structurées qui durent quelques jours associent cours théoriques et séances sur *pelvi-trainer* et modèle porcin vivant. Elles ont montré leur efficacité pour améliorer les performances techniques des internes [5].

L'opinion des internes et des enseignants est très favorable à l'utilisation des simulateurs pour la formation en chirurgie. La majorité des personnes interrogées ont déjà une expérience pour l'entraînement sur des gestes simples, mais elles croient utile de pouvoir s'entraîner à l'avenir sur des modèles permettant de réaliser des interventions plus complexes comme l'hystérectomie, même si ces outils sont encore peu disponibles et peu évalués à l'heure actuelle. L'opinion des enseignants est clairement en faveur d'une intégration obligatoire de la simulation dans la future formation des internes, mais ils sont plus nuancés concernant son utilisation pour évaluer les compétences des internes. Les données supportant l'utilisation de la simulation pour la certification sont prometteuses mais doivent encore être évaluées [11].

Des enquêtes similaires ont été réalisées dans d'autres spécialités chirurgicales. Toutes rapportent une opinion très positive des internes comme des enseignants en faveur de l'utilisation de laboratoires de simulation pour la formation [12,13]. Dans une enquête américaine réalisée auprès de chirurgiens orthopédiques, Karam et al. ont identifié que les principaux obstacles à l'utilisation de la simulation sont le manque de financement et de programmes d'entraînement disponibles [13]. Dans une enquête canadienne, les enseignants soulignent les difficultés liées aux moyens humains nécessaires pour faire fonctionner les laboratoires dans de bonnes conditions [14]. Cette question des coûts matériels mais surtout humains se pose également dans notre pays et constitue actuellement la principale limite à la création de centres de simulation, dont le nombre reste faible en France. À l'occasion de cette enquête la majeure partie des financements déclarés par les enseignants vient d'un partenariat avec l'industrie, ce qui souligne l'absence et la nécessité d'une politique pédagogique.

À ce jour, la seule politique d'utilisation de la simulation institutionnalisée concerne le programme FLS (*Fundamental of Laparoscopic Surgery*) dont la validation sur *pelvi-trainer* est obligatoire depuis 2008 aux États-Unis pour tous les internes de chirurgie générale [15]. Un groupe de travail dans ce domaine (*Gynaecological Endoscopic Surgical Education and Assessment Program*) a été institué au sein

de l'European Society of Gynecologic Endoscopy (ESGE) et devrait rendre ses conclusions à la fin l'année 2013.

Certains éléments limitent les enseignements de cette enquête. Afin d'obtenir un taux de participation maximum, un nombre restreint de questions a été retenu et par là même les informations recueillies sont partielles. En particulier, les conditions d'utilisation des simulateurs restent à préciser. Il serait par exemple intéressant de connaître le taux d'internes dont l'accès à des formations structurées sur simulateurs a été possible grâce à un déplacement dans un autre centre que leur ville d'origine. Par ailleurs, la notion de curriculum n'était pas formellement définie dans le questionnaire, et il est possible qu'il y ait eu une confusion sur ce point dans les réponses des enseignants entre un programme d'entraînement et un curriculum validé étape par étape par rapport à des performances d'expert [16]. Enfin, les enseignants de plusieurs grands centres français n'ont pas répondu et ceci peut biaiser les résultats sur les structures disponibles sur le territoire.

Conclusion

Cette étude montre la diversité des techniques de simulation déjà utilisées dans les écoles de chirurgie gynécologique françaises. Elle permet également de souligner les grandes disparités qui existent dans l'utilisation de ces techniques. Le défi des années à venir sera de mettre en place sur le territoire des structures permettant une formation selon des programmes d'entraînement validés et un accès égal pour tous les internes.

Déclaration d'intérêts

Les auteurs déclarent ne pas avoir de conflits d'intérêt en relation avec cet article.

Remerciements

Les membres du bureau de l'Association des gynécologues obstétriciens en formation (AGOF) qui ont accepté de collaborer à ce travail en permettant l'utilisation la liste d'adresses email des internes gynécologues français.

Mme Capelle, secrétaire de l'AGOF, pour son efficacité et sa disponibilité lors de la diffusion du questionnaire aux internes.

Tous les internes et enseignants qui ont consacré un peu de leur temps à remplir le questionnaire.

Annexe 1. Questionnaire adressé aux internes.

1. Dans quel CHU êtes vous en formation? 2. En quelle année êtes vous?
2. Combien d'interventions avez vous réalisé jusqu'à présent (opérateur principal)?
- Coéloscopie pour GEU : <5 5-20 20-50 >50
 - Hystérectomie vaginale : <5 5-20 20-50 >50
 - Hystérectomie voie haute : <5 5-20 20-50 >50
 - Hystérectomie coéloscopique : <5 5-20 20-50 >50
3. Au cours de votre formation en chirurgie gynécologique, avez vous eu un enseignement sur simulateur?
- oui non
- A. Si oui, sous quel forme?
- Simulateur pour faire des nœuds
 - Simulateur coéloscopique de réalité virtuelle (ex : Lapsim*, Lapmentor*,...)
 - Simulateur d'hystéroscopie Simulateur robotique Da Vinci
 - Entraînement sur tissus réels : Cadavre Animal (cochons)
- B. Si oui, dans quelles conditions avez vous utilisé ces méthodes de formation?
- Utilisation « libre » par mise à disposition à proximité du service.
 - Utilisation « encadrée » au cours de :
 - Séance(s) d'initiation.
 - Journées de formation universitaire.
 - Programmes intégrés lors d'un semestre d'internat.
 - Programmes de formation structurés, validés scientifiquement (selon le terme anglo-saxon de « curriculum »).

Commentaires :

-
-
4. Plus généralement, pensez vous qu'il y a une utilité des outils de simulation dans notre formation chirurgicale?
- Pour l'acquisition des gestes de base :

| | | | | | | |
|---------|---|---|---|---|---|------------|
| Inutile | 1 | 2 | 3 | 4 | 5 | Très utile |
|---------|---|---|---|---|---|------------|
 - Pour l'acquisition des interventions plus complexes (ex : hystérectomie)

| | | | | | | |
|---------|---|---|---|---|---|------------|
| Inutile | 1 | 2 | 3 | 4 | 5 | Très utile |
|---------|---|---|---|---|---|------------|

Annexe 2. Questionnaire adressé aux PU et MCU.

1. Dans quel centre universitaire êtes-vous formateur en chirurgie gynécologique?
2. Au cours de votre formation en chirurgie gynécologique, avez-vous eu un enseignement sur simulateur?
- oui non
3. Au sein de votre centre universitaire, existe-t-il un accès à un outil de simulation chirurgicale pour les internes en formations :
- oui non
- A. Si oui, quels sont les lieux où sont situés le(s) simulateur(s)?
- Le service de chirurgie Un local de la faculté
 - Un centre spécifique dédié à cette activité
- B. Si oui, sous quelle(s) forme(s)?
- Simulateur pour faire des nœuds Pelvic trainer coéloscopique
 - Simulateur coéloscopique de réalité virtuelle (ex : Lapsim*, Lapmentor*...)
 - Simulateur d'hystéroscopie Simulateur robotique Da Vinci
 - Entraînement sur tissus réels : Cadavre Animal (cochons)
- C. Si oui, dans quelles conditions utilisez-vous cet outil de formation?
- Utilisation « libre » par mise à disposition à proximité du service.
 - Utilisation « encadrée » au cours de :
 - Séances d'initiation.
 - Journées de formation universitaire.
 - Programmes intégrés lors d'un semestre d'internat.
 - Programmes de formation structurés, validés scientifiquement (selon le terme anglo-saxon de « curriculum »)

Annexe 2. (Suite)

Dans le cadre de programmes de recherche.

D. Si oui, quels sont les modes de financement qui ont permis l'achat et l'entretien de ce matériel?

.....

.....

4. Plus généralement, pensez-vous qu'il y ait une utilité des outils de simulation dans notre formation chirurgicale?

• Pour l'acquisition des gestes de base :

Inutile 1 2 3 4 5 Très utile

• Pour l'acquisition des interventions plus complexes (ex : hystérectomie)

Inutile 1 2 3 4 5 Très utile

5. Pensez-vous que l'entraînement sur simulateur devrait obligatoirement faire partie de la formation chirurgicale des internes à l'avenir?

Pas du tout d'accord 1 2 3 4 5 Tout à fait d'accord

6. Pensez-vous que l'entraînement sur simulateur devrait obligatoirement faire partie de la certification des internes à l'avenir?

Pas du tout d'accord 1 2 3 4 5 Tout à fait d'accord

Commentaires :

.....

.....

Références

- [1] Zendejas B, Brydges R, Hamstra SJ, Cook DA. State of the evidence on simulation-based training for laparoscopic surgery: a systematic review. *Ann Surg* 2013;257:586–93.
- [2] Guide de bonnes pratiques en matière de simulation en santé. HAS; 2012.
- [3] Selzer DJ, Dunnington GL. Surgical skills simulation: a shift in the conversation. *Ann Surg* 2013;257:594–5.
- [4] Larsen CR, Soerensen JL, Grantcharov TP, Dalsgaard T, Schouenborg L, Ottosen C, et al. Effect of virtual reality training on laparoscopic surgery: randomised controlled trial. *BMJ* 2009;14:338 [b1802].
- [5] Botchorishvili R, Rabischong B, Larrain D, Khoo CK, Gaia G, Jardon K, et al. Educational value of an intensive and structured interval practice laparoscopic training course for residents in obstetrics and gynecology: a four-year prospective, multi-institutional recruitment study. *J Surg Educ* 2012;69:173–9.
- [6] Hudry D. Training. ... *Gynecol Obstet Fertil* 2012;40:390.
- [7] Carcopino X, Shojai R, D'Ercole C, Boubli L. French trainees in obstetrics and gynaecology theoretical training and practice of vaginal breech delivery: a national survey. *Eur J Obstet Gynecol Reprod Biol* 2007;135:17–20.
- [8] Huchon C. État des lieux des hystérectomies en France. *Réalités en gynécologie obstétrique* 2013;167:25–8.
- [9] Sroka G, Feldman LS, Vassiliou MC, Kaneva PA, Fayez R, Fried GM. Fundamentals of laparoscopic surgery simulator training to proficiency improves laparoscopic performance in the operating room – a randomized controlled trial. *Am J Surg* 2010;199:115–20.
- [10] Philippe AC, Botchorishvili R, Pereira B, Canis M, Bourdel N, Mage G, Pouly JL, Houille C, Jardon K, Rabischong B. Interest of a structured laparoscopy training in a simulation center: survey of resident's point of view. *J Gynecol Obstet Biol Reprod* 2013;42:238–45.
- [11] Vassiliou MC, Feldman LS. Objective assessment, selection, and certification in surgery. *Surg Oncol* 2011;20:140–5.
- [12] Aggarwal R, Balasundaram I, Darzi A. Training opportunities and the role of virtual reality simulation in acquisition of basic laparoscopic skills. *J Surg Res* 2008;145:80–6.
- [13] Karam MD, Pedowitz RA, Natividad H, Murray J, Marsh JL. Current and future use of surgical skills training laboratories in orthopaedic resident education: a national survey. *J Bone Joint Surg Am* 2013;295:e4 [DOI: 10.2106].
- [14] Kapadia M, DaRosa DA, MacRae HM, Dunnington GL. Current assessment and future directions of surgical skills laboratories. *J Surg Educ* 2007;64:260–5.
- [15] Accreditation Council for Graduate Medical Education. Available from: http://www.acgme.org/acWebsite/downloads/RRC_progReq/440_general_surgery_01012008_u08102008.pdf
- [16] Zevin B, Levy JS, Satava RM, Grantcharov TP. A consensus-based framework for design, validation, and implementation of simulation-based training curricula in surgery. *J Am Coll Surg* 2012;215:580–6 [e3].