

Faculté des sciences de la motricité

**La plus-value du GNRB® dans le
diagnostic de rupture partielle ou
complète du ligament croisé
antérieur : une revue narrative
systématisée**

Auteur : Rodolphe Willocx

Promoteur(s) : Edith Van Bunnan – Philippe Mahaudens

Année académique 2023-2024

Master en kinésithérapie et réadaptation [60.0] – KINE2M

*JE TIENS À EXPRIMER MES REMERCIEMENTS LES PLUS
SINCÈRES À MES PROMOTEURS DE MÉMOIRE,
EDITH VAN BUNNEN ET PHILIPPE MAHAUDENS,
QUI M'ONT ENCADRÉ, SOUTENU, ORIENTÉ ET AIDÉ DANS LA
RÉALISATION DE CE MÉMOIRE.*

*JE SOUHAITERAIS ÉGALEMENT REMERCIER DE TOUT CŒUR
MES PARENTS AINSI QUE MA COPINE POUR LEUR SOUTIEN ET
L'INTÉRÊT QU'ILS ONT PORTÉ À MON TRAVAIL, AINSI QUE
POUR LEUR PRÉSENCE ET LEURS ENCOURAGEMENTS DURANT
MES ANNÉES D'ÉTUDES.*

*MERCI À TOUT LE CORPS PROFESSORAL DE LA FSM DE
L'UCLouvain POUR CES BELLES ANNÉES D'ÉTUDES AINSI
QUE POUR LA QUALITÉ DE LEUR ENSEIGNEMENT.*

ABRÉVIATIONS

GNRB[®]	GeNouRoB
IRM	Imagerie par Résonance Magnétique
LCA	Ligament Croisé Antérieur
LCP	Ligament Croisé Postérieur
LDA	Laximétrie Dynamique Automatisée
NIH	National Institutes of Health
NOS	Newcastle-Ottawa Scale
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analysis
RC	Rupture complète
RDP	Radiographie Dynamique Passive
RP	Rupture partielle
Se	Sensibilité
Sp	Spécificité

TABLE DES MATIÈRES

INTRODUCTION	6
I. MÉTHODOLOGIE.....	9
I.1 Stratégie de recherche.....	9
I.2 Sélection des études et analyse des données.....	10
II. RÉSULTATS	12
II.1 Sélection des études.....	12
II.2 Évaluation de la qualité méthodologique des articles	14
II.3 Caractéristiques des études sélectionnées.....	16
III. DISCUSSION.....	26
CONCLUSION	36
BIBLIOGRAPHIE	37
ANNEXES.....	42

INTRODUCTION

Le ligament croisé antérieur (LCA) est une structure ligamentaire importante du genou, reliant directement le fémur au tibia. Il est constitué de deux faisceaux, le faisceau antéro-médial et le faisceau postéro-latéral (Stone *et al.*, 2021). Il s'insère en proximal sur la face postéro-médiale du condyle fémoral latéral et se termine en distal sur l'aire inter-condylaire antérieure du tibia (Masik *et al.*, 2022; Stone *et al.*, 2021; Zantop *et al.*, 2005). Son orientation médiale, vers l'avant et le bas lui confère une fonction de stabilisation antérieure du genou, il limite ainsi le tiroir antérieur tibia-fémur et joue également un rôle important dans le contrôle de la rotation médiale du tibia (Masik *et al.*, 2022; Pappas *et al.*, 2013).

La rupture du LCA est une des pathologies les plus fréquentes du genou (Masik *et al.*, 2022). Bien qu'il soit compliqué d'en évaluer le nombre annuel exact, il est estimé qu'environ 1 personne sur 3500 ($\approx 0,03\%$) est touchée chaque année par une rupture du LCA (Moses *et al.*, 2012; Noyes & Barber-Westin, 2016). Différentes études ont cependant soutenu l'idée que ce chiffre pourrait en réalité être encore plus élevé (Sanders *et al.*, 2016). La rupture du LCA est particulièrement commune chez les sportifs de tout niveau, plus encore chez les femmes que les hommes et est surtout retrouvée dans les sports de pivots (Larwa *et al.*, 2021; Masik *et al.*, 2022). Différents mécanismes lésionnels sont observés ; il peut s'agir d'une hyper-extension du genou ou, plus fréquemment, d'un genou fléchi, en valgus et en rotation interne du tibia par rapport au fémur, souvent due à un mouvement de pivot lorsque le pied est fixe (De Carli *et al.*, 2022; Koga *et al.*, 2010; Masik *et al.*, 2022; Vavken & Murray, 2013). Il est important de noter que la majorité des blessures du LCA surviennent dans des situations sans contact ; lors de l'atterrissage d'un saut, un changement de direction, une accélération ou un freinage (Larwa *et al.*, 2021; Sokal *et al.*, 2022; Vavken & Murray, 2013).

Le diagnostic de rupture du LCA est clinique. Il est réalisé sur base de l'histoire du patient, associée à des tests cliniques et quasiment systématiquement complété par une imagerie par résonance magnétique (IRM) (Brady & Weiss, 2018; Cojean *et al.*, 2023a; Sokal *et al.*, 2022). Le test de Lachman est souvent considéré comme étant le test clinique le plus fiable pour diagnostiquer une rupture du LCA

(Ferretti *et al.*, 2022; Masik *et al.*, 2022), mais cela est sujet à pas mal de désaccord au sein de la littérature scientifique récente (Sokal *et al.*, 2022). D'autres tests, tels que le test de ressaut rotatoire, le test du tiroir antérieur et le signe du levier, sont également utilisés dans le diagnostic clinique de lésion du LCA et pourraient même avoir une meilleure exactitude diagnostique d'après une étude récente de Sokal *et al.*. Le problème de ces tests cliniques, c'est qu'ils sont opérateur-dépendants, qu'ils demandent souvent une certaine expérience et qu'ils ne permettent pas de quantifier objectivement le taux de laxité mesurée (Khan *et al.*, 2017; Robert *et al.*, 2009), rendant la différenciation entre rupture partielle (RP) et complète (RC) du LCA parfois compliquée. Or, cette différenciation permet entre autres d'aiguiller la décision thérapeutique, lorsqu'il faut choisir entre le traitement conservateur et l'acte chirurgical (Colombet *et al.*, 2010; Stone *et al.*, 2021).

À ce jour, l'outil non-invasif le plus fréquemment utilisé pour détecter la rupture du LCA, est l'IRM. Cette technique comporte bon nombre d'avantages, dont le fait que le contraste entre les tissus mous est clair et qu'elle permet d'évaluer simultanément l'état des structures adjacentes (Cojean *et al.*, 2023a). Il existe cependant également quelques désavantages à l'IRM. La spécificité et la sensibilité de l'IRM dépendent de la technique de scan utilisée et sont également fortement influencées par l'expérience et les capacités du radiologue. De plus, l'IRM est onéreux et le temps d'attente pour le patient est souvent long (Brady & Weiss, 2018; Cojean *et al.*, 2023a; Masik *et al.*, 2022).

Une technique, développée depuis les années '70 et qui se développe et évolue toujours, a été mise en avant afin d'améliorer le diagnostic du LCA, la laximétrie ou l'arthrométrie. Comme son nom l'indique, la laximétrie consiste à mesurer la laxité d'une articulation à l'aide d'un arthromètre (Rousseau, 2018). Dans le cas du LCA, cela permet de quantifier la translation du tibia par rapport au fémur. Il existe plusieurs machines développées dans ce but, dont les plus connues et étudiées sont le KT-1000[®], le Rolimeter[®], Telos[®] et le GNRB[®] (Genourob, Laval, France) (Cojean *et al.*, 2023a; Jenny *et al.*, 2017; Lefevre *et al.*, 2014).

Le KT-1000[®] et le Rolimeter[®] sont assez simples à utiliser mais sont opérateur-dépendants, rendant les résultats peu reproductibles et inexacts (Cojean *et al.*, 2023a; Lefevre *et al.*, 2014). Telos[®] est une machine qui met le LCA sous

contrainte lors d'une radiographie, elle semblerait obtenir des résultats plus précis et reproductibles que le KT-1000[®] et le Rolimeter[®], mais l'appareil est très coûteux et expose le patient à des radiations (Lefevre *et al.*, 2014).

Le GNRB[®] a été créé dans le but de réaliser des mesures plus reproductibles et précises dans les ruptures du LCA que les machines déjà existantes (Lefevre *et al.*, 2014). L'appareil est formé par un système entièrement automatisé qui réalise des tests dynamiques pour mesurer la translation antérieure du tibia, appelée Laximétrie Dynamique Automatisée (LDA). Nous nous sommes intéressés au GNRB[®] spécifiquement car c'est un appareil récent, qui a tout de même déjà pas mal été étudié, permettant une comparaison d'articles récents assez large. De plus, les études du GNRB[®] ont fréquemment obtenu des résultats très encourageants, suggérant qu'il puisse avoir une vraie valeur ajoutée dans le diagnostic de lésion du LCA (Bouguennec *et al.*, 2015; Collette *et al.*, 2012).

L'objectif de ce mémoire est d'évaluer, sur base de la littérature scientifique récente, la valeur ajoutée et l'intérêt du GNRB[®] dans le diagnostic de RP et RC du LCA, de voir quel rôle le GNRB[®] peut avoir lors de l'évaluation de blessures du LCA. À terme, le GNRB[®] pourrait-il remplacer l'IRM ou, à défaut, de diminuer le recours à l'IRM en première intention, diminuant ainsi les coûts et les délais de prise en charge ?

I. MÉTHODOLOGIE

I.1 Stratégie de recherche

La présente revue de la littérature a été élaborée en suivant de près les directives de la déclaration du *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA).

Une question de recherche a été construite selon la méthode PICO : la Population (P), l'Intervention (I), la Comparaison (C) et les Résultats (O : Outcome). Ceci a donné la question de recherche suivante : « Quelle est la plus-value du GNRB® dans le diagnostic de rupture partielle et complète du ligament croisé antérieur » ? L'élaboration d'une question de recherche précise à l'aide de cette méthode permet de bien cerner le sujet de notre travail afin de réaliser une recherche ciblée ainsi que d'être méthodique et le plus exhaustif possible dans la sélection des articles qui traitent le sujet.

La recherche bibliographique a été réalisée sur base de cette question de recherche. Quatre bases de données différentes ont été épluchées dans le but d'obtenir un aperçu le plus exhaustif possible de la littérature traitant du GNRB® dans le diagnostic de rupture du LCA. Ces bases de données sont Pubmed, Embase, ScienceDirect et Scopus, elles ont été choisies compte tenu du très large nombre d'articles référencés sur chacune d'elles.

Afin d'obtenir tous les articles mentionnant le GNRB®, les recherches dans les trois premières bases de données ont été réalisées en utilisant uniquement les mots-clés 'gnrb OR genourob' en examinant tous les champs. Cette décision a été prise suite au relativement faible nombre d'articles obtenus lors de ces recherches, ne demandant pas à ce que celles-ci soient particulièrement affinées. La recherche dans Scopus a, elle, été construite à l'aide d'une équation de recherche car le nombre de résultats obtenus était plus important en ne

recherchant que 'gnrb OR genourob'. L'objectif était d'affiner la recherche dès le départ dans le but d'avoir un nombre de résultats moins conséquent tout en parvenant à cibler tous les articles potentiellement pertinents pour cette revue.

L'équation de recherche utilisée dans Scopus est la suivante: « (gnr OR genourob) AND ('anterior AND cruciate AND ligament' OR acl OR 'anterior AND cruciate AND ligaments') AND (diagnosis OR diagnostic OR diagnose) ».

Dans chacune des bases de données analysées, les recherches ont été limitées aux publications parues à partir de 2007, année de création du GNRB® par la société Genourob à Laval, France.

I.2 Sélection des études et analyse des données

Les recherches dans ces quatre bases de données ont mené à un total de 525 références. Toutes ces références ont été intégrées à un outil de gestion et de tri de références scientifiques appelé RAYYAN. Les doublons ont été détectés et supprimés dès le départ, puis les titres, abstraits et mots-clés de chacune des références restantes ont été analysés en nous tenant à des critères d'inclusion et d'exclusion prédéfinis.

Ont été incluses : (1) toutes les études quantitatives, expérimentales, comparatives et/ou diagnostiques (2) publiées en français ou en anglais, qui (3) comparent le GNRB® à tout autre moyen de mesure dans (4) le diagnostic de RP ou RC du LCA (5) réalisées sur des sujets avec un LCA pathologique.

Un article est considéré comme non pertinent lorsqu'il présente un ou plusieurs critères d'exclusion suivants : (1) étude qualitative (revue ou méta-analyse), (2) étude réalisée sur des cadavres, (3) étude post-opératoire, (4) étude sur genoux sains. (**Tableau 1**).

Tableau 1. Critères d'inclusion et d'exclusion de la revue

CRITÈRES D'INCLUSION	CRITÈRES D'EXCLUSION
• Étude quantitative, expérimentale, comparative et/ou diagnostique • Langue française ou anglaise • Comparaison GNRB® à autre moyen de mesure • Diagnostic de RP ou RC du LCA • LCA pathologique	• Étude qualitative • Étude cadavérique • Étude post-opératoire • Genoux sains

II. RÉSULTATS

II.1 Sélection des études

Nos recherches à l'aide des mots-clés 'gnrb or genourob' ainsi que de notre équation de recherche ont mené à l'élaboration d'une base de 525 références dans RAYYAN, un logiciel de gestion de références bibliographiques. De ces 525 références, 66 ont été obtenues sur Pubmed, 94 sur Embase, 194 sur ScienceDirect et 171 sur Scopus.

Lors de l'analyse de ces références à l'aide du logiciel RAYYAN, 198 doublons ont été détectés de prime abord. Nous les avons vérifiés individuellement afin de limiter au maximum le risque d'erreurs puis ils ont été supprimés immédiatement afin de ne laisser que 327 références à analyser.

L'analyse systématique des titres, abstraits et mots-clés de ces 327 références a finalement mené à l'exclusion de 314 articles ne respectant pas les critères d'inclusion et d'exclusion de notre recherche. Tout cela a permis la sélection de 13 articles pour une première lecture. Suite à celle-ci, trois articles ont été écartés car deux d'entre eux sont des études post-opératoires et que le dernier ne compare pas le GNRB® à un autre moyen de diagnostic de RP ou RC du LCA. Finalement, un total de 10 articles a été inclus dans les résultats de notre revue et compilé dans le **Tableau 3**. La **Figure 1** illustre clairement le processus de sélection, d'inclusion et d'exclusion des articles retenus dans cette revue de la littérature.

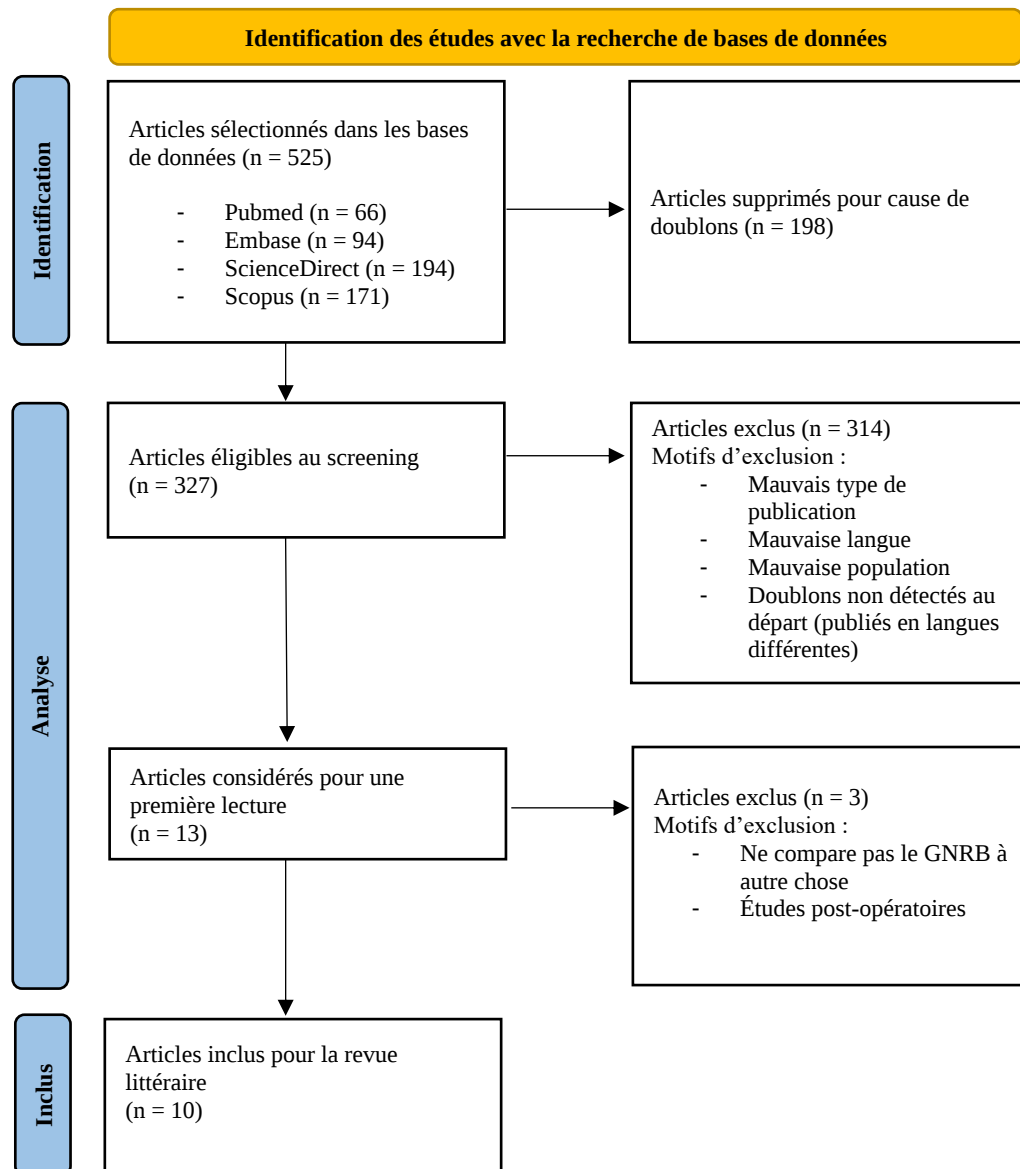


Figure 1. Diagramme de recherche selon les directives PRISMA.

II.2 Évaluation de la qualité méthodologique des articles

Une fois les études sélectionnées dans cette revue de la littérature, elles ont été analysées, à commencer par leur qualité méthodologique, appréciée à l'aide d'un outil d'évaluation de la qualité des études scientifiques. Pour cela, nous avons fait le choix d'utiliser l'outil développé spécialement pour les études de cohorte et les études transversales par les National Institutes of Health (NIH), le « *NIH Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies* » (en annexe). Le choix s'est porté sur cet outil en particulier car les études incluses dans cette revue sont majoritairement des études de cohorte ou des études transversales. Il présente l'avantage de pouvoir évaluer les deux types littéraires à la fois, à l'inverse d'autres outils tels que la plus largement utilisée « *Newcastle-Ottawa Scale (NOS) for cohort studies* », qui se limitent en général aux études de cohorte. Par ailleurs, le « *NIH Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies* » semble relativement simple d'utilisation et de compréhension. (Ma *et al.*, 2020).

Le **Tableau 2** reprend les résultats pour chaque étude obtenus lors de l'analyse de la méthodologie à l'aide de l'échelle des NIH. Cette échelle est composée de 14 points, sous forme de questions auxquelles il faut répondre par « *Oui* » ou « *Non* » en fonction de si elle s'applique à l'étude analysée. Il est également possible de choisir une autre réponse lorsqu'il n'est pas possible de répondre par oui ou par non ; « *CD* » pour « *Cannot determine* », « *NA* » pour « *Not applicable* » ou « *NR* » pour « *Not reported* ». Littéralement cela signifie ; « indéterminable », « ne s'applique pas » et « non communiqué ». Dans le **Tableau 2**, « *Oui* » est représenté par un 'V' blanc dans une boule verte, « *Non* » par une croix blanche dans une boule rouge et les trois autres options sont regroupées en un seul et même signe, le point d'exclamation dans une boule jaune. Chaque « *Oui* » rapporte un point, toute autre réponse n'en rapporte aucun. Plus une étude obtient de points, plus la méthodologie de celle-ci est considérée qualitative.

De manière générale, les dix études incluses dans cette revue ont obtenu des résultats satisfaisants lors de l'évaluation à l'aide de l'échelle des NIH. Les scores obtenus s'étalent de 7 à 10 sur 14. En moyenne, chaque étude obtient le score de 8,3. La médiane se situe à 8,5 et le mode est de 9. Ces résultats satisfaisants et équilibrés sont à prendre avec un peu de recul, car les études incluses ne sont pas toutes des études transversales ou de cohorte. Il y a également un certain nombre d'études diagnostiques et comparatives, pour lesquelles l'échelle utilisée n'est pas forcément la plus adaptée. De plus, certaines études contiennent des incohérences, des erreurs et/ou sont plus évasives quant au développement de leurs calculs (e.a. (Cojean *et al.*, 2023a; Lefevre *et al.*, 2014; Saravia *et al.*, 2020; Sidibé *et al.*, 2020). Ces erreurs ou incohérences ne sont pas directement prises en compte par l'échelle des NIH et pourraient jouer un rôle plus ou moins important dans l'acceptabilité des conclusions auxquelles sont arrivés leurs auteurs. Parallèlement, la question de l'objectivité peut se poser pour l'étude de Robert *et al.*, dont les auteurs sont les inventeurs du GNRB[®], ainsi que celle de Cojean *et al.*, dont le Dr. Robert est co-auteur.

Tableau 2. Risque de biais pour les études incluses. NIH Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies.

	Question NIH														Total de "oui"
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
Beldame <i>et al.</i> , 2012	✓	✓	⚠	✓	✗	✓	✓	✓	✓	✗	✓	⚠	✓	⚠	9/14
Cojean <i>et al.</i> , 2023	✓	✓	⚠	✓	✗	✓	✓	✗	✓	✗	✓	⚠	✓	⚠	8/14
Di Iorio <i>et al.</i> , 2014	✓	✓	⚠	✓	✗	✓	✓	✗	✓	✗	✓	⚠	✓	⚠	8/14
Jenny <i>et al.</i> , 2013	✓	✓	⚠	✓	✗	✓	✓	✗	✓	✗	✓	⚠	⚠	⚠	7/14
Lefevre <i>et al.</i> , 2014	✓	✓	⚠	✓	✗	✓	✓	✓	✓	✗	✓	⚠	✓	⚠	9/14
Murgier <i>et al.</i> , 2018	✓	✓	⚠	✓	✓	✓	✓	✗	✓	✗	✓	⚠	✓	⚠	9/14
Robert <i>et al.</i> , 2009	✓	✓	⚠	✓	✗	✓	✓	✓	✓	✗	✓	⚠	✓	⚠	9/14
Ryu <i>et al.</i> , 2018	✓	✓	⚠	✓	✗	✓	✓	✓	✓	✗	✓	⚠	✓	✓	10/14
Saravia <i>et al.</i> , 2020	✓	✓	✗	✗	✗	✓	✓	✗	✓	✗	✓	⚠	✓	⚠	7/14
Sidibé <i>et al.</i> , 2020	✓	✓	⚠	✓	✗	✓	✓	✗	✓	✗	✓	⚠	⚠	⚠	7/14

Question 8 a été comptabilisée à partir du moment où différents paramètres ont été testés. La 13 a été comptabilisée à partir du moment où il y a au moins autant de mesures pour les tests que de personnes incluses (ou < de 20% de différence).

II.3 Caractéristiques des études sélectionnées

Les caractéristiques majeures de chacune des 10 études sélectionnées dans cette revue sont reprises en détails dans le **Tableau 3**. Ce tableau reprend le nom du premier auteur, l'année de publication, le design, la taille de l'échantillon, l'âge moyen des participants et l'état des LCA étudiés pour chacune des études. Il s'attarde également sur d'autres informations particulièrement intéressantes provenant de chaque article telles que le ou les outil(s) comparé(s) au GNRB®, le protocole mis en place, les mesures effectuées, les résultats obtenus et ce qu'en ont conclu les examinateurs.

Les 10 études incluses dans ce tableau sont toutes des études observationnelles, majoritairement de cohorte ou transversale, prospective ou rétrospective. Elles comparent chacune le GNRB® à un autre outil utilisé dans le diagnostic de RP ou RC du LCA ; les radiographies dynamiques passives (RDP) : Telos® et la méthode « Lerat » (5 études ; Beldame *et al.*, 2012; Jenny & Arndt, 2013; Lefevre *et al.*, 2014; Murgier *et al.*, 2018; Ryu *et al.*, 2018), l'IRM (4 études ; (Cojean *et al.*, 2023a; Di Iorio *et al.*, 2014; Saravia *et al.*, 2020; Sidibé *et al.*, 2020), l'examen clinique (2 études ; Di Iorio *et al.*, 2014; Ryu *et al.*, 2018), le KT-1000® (2 études ; Murgier *et al.*, 2018; Robert *et al.*, 2009) et/ou le Rolimeter® (1 étude ; Murgier *et al.*, 2018).

Beldame *et al.*, Cojean *et al.*, Di Iorio *et al.*, Robert *et al.* ainsi que Sidibé *et al.*, ont utilisé l'évaluation arthroscopique comme référence, considérant que cet outil permettait l'obtention de l'information la plus fiable sur l'état de santé des structures ligamentaires observées. Jenny & Arndt, quant à eux, ont pris l'Orthopilot®, un système de navigation peropératoire, pour référence lors de la mesure de laxité du genou tandis que Lefèvre *et al.*, Ryu *et al.*, et Saravia *et al.*, ont considéré, de leur côté, que les images obtenues à l'IRM, éventuellement en combinaison avec l'examen clinique, feraient office de référence.

Tableau 3. Synthèse des articles inclus dans l'étude.

Auteurs et année	Design	Sujets	État LCA	Comparaison GNRB®	Protocole	Mesure des résultats	Résultats																													
1. Beldame et al., 2012	Étude prospective de cohorte	N = 157 cas (H : 121, F : 36) Âge moyen = 40,0 ans [18–69]	- Sain - RP - RC	RDP : Telos® – méthode « Lerat » Arthroscopie (référence)	1) Telos® 2 genoux 2) Lerat 2 genoux 3) GNRB® 2 genoux 4) Arthroscopie	Se et Sp prédite pour chaque appareil (RP exclues)	<table><thead><tr><th colspan="5">Appareil (seuil différentiel)</th></tr><tr><th></th><th>Telos® (3 mm)</th><th>Lerat (3 mm)</th><th colspan="2">GNRB® (1,5 mm)</th></tr><tr><th></th><th></th><th></th><th>à 250N</th><th>à 134N</th></tr></thead><tbody><tr><td>Se</td><td>64,8%</td><td>43,2%</td><td>62,2%</td><td>59,4%</td></tr><tr><td>Sp</td><td>75,8%</td><td>82,7%</td><td>75,9%</td><td>93,1%</td></tr></tbody></table> ➔ Valeur diagnostique similaire entre GNRB® à 250N et RDP (Telos® et Lerat), mais GNRB® n’est pas irradiant, donc à favoriser.	Appareil (seuil différentiel)						Telos® (3 mm)	Lerat (3 mm)	GNRB® (1,5 mm)					à 250N	à 134N	Se	64,8%	43,2%	62,2%	59,4%	Sp	75,8%	82,7%	75,9%	93,1%				
Appareil (seuil différentiel)																																				
	Telos® (3 mm)	Lerat (3 mm)	GNRB® (1,5 mm)																																	
			à 250N	à 134N																																
Se	64,8%	43,2%	62,2%	59,4%																																
Sp	75,8%	82,7%	75,9%	93,1%																																
2. Cojean et al., 2023	Étude prospective de cohorte	N = 214 cas (H : 143, F : 63)** Âge moyen = 30 ± 12 ans [18-71]	- Sain - RP - RC	IRM Arthroscopie (référence)	1) IRM 2) GNRB® 2 genoux 3) Arthroscopie	Se, Sp, VPP et VPN pour chaque appareil	<table><thead><tr><th rowspan="2"></th><th colspan="2">IRM</th><th colspan="2">GNRB® à 134N</th></tr><tr><th>RP</th><th>RC</th><th>RP (1,3 mm)</th><th>RC (3 mm)</th></tr></thead><tbody><tr><td>Se</td><td>29,51%</td><td>80,81%</td><td>73,77%</td><td>77,78%</td></tr><tr><td>Sp</td><td>88,97%</td><td>64,49%</td><td>85,52%</td><td>85,98%</td></tr><tr><td>VPP</td><td>52,94%</td><td>67,8%</td><td>68,18%</td><td>83,7%</td></tr><tr><td>VPN</td><td>75%</td><td>78,41%</td><td>88,57%</td><td>80,7%</td></tr></tbody></table> ➔ GNRB® et IRM sont aussi performants pour détecter les RC et GNRB® est plus performant pour les RP que IRM. Combinaison GNRB® et IRM : capacité diagnostique augmentée.		IRM		GNRB® à 134N		RP	RC	RP (1,3 mm)	RC (3 mm)	Se	29,51%	80,81%	73,77%	77,78%	Sp	88,97%	64,49%	85,52%	85,98%	VPP	52,94%	67,8%	68,18%	83,7%	VPN	75%	78,41%	88,57%	80,7%
	IRM		GNRB® à 134N																																	
	RP	RC	RP (1,3 mm)	RC (3 mm)																																
Se	29,51%	80,81%	73,77%	77,78%																																
Sp	88,97%	64,49%	85,52%	85,98%																																
VPP	52,94%	67,8%	68,18%	83,7%																																
VPN	75%	78,41%	88,57%	80,7%																																
3. Di Iorio et al., 2014	Étude prospective de cohorte	N = 49 cas (H : 35, F : 14) Âge moyen = 31 ± 11 ans	- RP - RC	Examen clinique IRM Arthroscopie (référence)	1) Examen clinique bilatéral (test de LT et ressaut) (anesthésié) 2) IRM 3) GNRB® 2 genoux 4) Arthroscopie	Capacité à différencier une RP d’une RC	<u>Examen clinique</u> : PAS capable de différencier significativement <u>IRM</u> : Différence significativement les RP des RC (Se = 84% et Sp = 92% pour détection RP) <u>GNRB®^a</u> : Différence significativement les RP des RC (Se = 72% pour détection RP) ^a seuil de 1,5 mm à 134N																													

H : Hommes ; F : Femmes ; RDP : Radiographie Dynamique Passive ; GNRB® : GeNouRoB ; D : Droit ; G : Gauche ; Se : sensibilité ; Sp : spécificité ; IRM : Imagerie à Résonance Magnétique ; VPP : Valeur Prédictive Positive ; VPN : Valeur Prédictive Négative ; LT : Lachman-Trillat ; CCCs : coefficients de corrélation de concordance ; A2 genoux : différence entre les 2 genoux ; n.c. : non connu ; IY : Index J de Youden ; ICC : Intra-class correlation

Tableau 3. Synthèse des articles inclus dans l'étude.

4. Jenny & Arndt, 2013	Étude transversale	N = 21 cas (H : 14, F : 7) Âge moyen = 28 ans	Laxité chronique^b ^b diagnostic par examen clinique et (parfois) IRM	RDP : méthode « Lerat » Navigation peropératoire OrthoPilot® (référence)	1) GNRB® genou pathologique (éveillé) 2) RDP genou pathologique : protocole proche de Lerat (anesthésié) 3) OrthoPilot® durant l'opération (anesthésié)	Précision de la mesure de laxité antérieure du genou	<u>RDP vs OrthoPilot®</u> : différence significative <u>GNRB® vs OrthoPilot®</u> : différence significative <u>RDP vs GNRB®</u> : PAS de différence significative ➔ RDP et GNRB® : même fiabilité donc privilégier GNRB® (non-irradiant) MAIS sous-estime la laxité réelle. ^c à 250N	
5. Lefevre et al., 2014	Étude diagnostique prospective	N = 139 cas (H : 86, F : 53) Âge moyen = 30,7 ± 9,3 ans	- RP^d - RC^d ^d diagnostic par IRM et examen clinique	RDP : Telos®	1) Telos® 2 genoux (pathologique en premier) 2) GNRB® 2 genoux (sain en premier)	Corrélation entre mesures du Telos® et du GNRB®. Se et Sp aux seuils optimaux pour les 2 machines Valeur diagnostique de chaque test	Corrélation raisonnable, meilleure entre Telos® 250N et GNRB® 250N. <u>Telos® (RP)</u> à 150N, seuil optimal : 4 mm (Se = 62,3% et Sp = 73,7%) à 250N, seuil optimal : 3,6 mm (Se = 81,5% et Sp = 59,5%) <u>GNRB® (RP)</u> à 134N, seuil optimal : 2 mm (Se = 83,2% et Sp = 64,3%) à 250N, seuil optimal : 2,5 mm (Se = 84% et Sp = 81%) ➔ Valeur diagnostique du GNRB® est meilleure que Telos® pour les RP.	
6. Murgier et al., 2018	Étude comparative prospective	N = 52 cas (H : 40, F : 12) Âge moyen = 27,5 ± 8,2 ans	RC^d ^d diagnostic par IRM et examen clinique	RDP : Telos® KT-1000® Rolimeter®	1) KT-1000® 2) Rolimeter® 3) GNRB® à 200N 4) Telos® à 25KPa ^e 3 mesures par genou puis moyenne	Calcul de CCCs ; comparabilité d'un outil par rapport à un autre	<u>Mauvaise corrélation globale</u> entre les outils donc mesures peu comparables. Comparabilité entre le KT-1000® et le GNRB® est moyenne (CCC = 0,51) et est faible pour les autres outils.	

H : Hommes ; *F* : Femmes ; *RDP* : Radiographie Dynamique Passive ; *GNRB®* : GeNouRoB ; *D* : Droit ; *G* : Gauche ; *Se* : sensibilité ; *Sp* : spécificité ; *IRM* : Imagerie à Résonance Magnétique ; *VPP* : Valeur Prédictive Positive ; *VPN* : Valeur Prédictive Négative ; *LT* : Lachman-Trillat ; *CCCs* : coefficients de corrélation de concordance ; *Δ2* genoux : différence entre les 2 genoux ; *n.c.* : non connu ; *IX* : Index J de Youden ; *ICC* : Intra-class correlation

Tableau 3. Synthèse des articles inclus dans l'étude.

7. Robert et al., 2009	Étude comparative prospective	(série 1) N = 20 cas (H : 17, F : 3) Âge moyen = n.c. [19-22ans] (série 2a) N = 21 cas (H : 15, F : 6) Âge moyen = 26 ans [16-36] (série 2b) N = 24 cas (H : 19, F : 5) Âge moyen = 31 ans [15-59]	(série 1) • Sain (série 2) • a) RC • b) RP	KT-1000®	1) Multiples mesures des 2 genoux avec GNRB® et KT-1000® (ordre aléatoire, total de 308 paires de mesures, 2 examinateurs différents) 2) 3 mesures par genou avec minimum GNRB® (total de 241 mesures) 3) Arthroscopie (série 2 uniquement)	(série 1) Dispersion, précision des mesures (variances) (séries 2) Se et Sp maximales aux seuils optimaux pour le GNRB à 134N	<u>GNRB®f</u> : faible dispersion des mesures <u>KT-1000®</u> : forte dispersion ➔ Meilleure reproductibilité du GNRB®, indépendamment de l'expérience de l'examineur à l'inverse du KT-1000® <u>GNRB®f</u> : seuils de laxité différentielle optimaux : RP : 1,5 mm (Se = 80% et Sp = 87%) RC : 3 mm (Se = 70% et Sp = 99%) f à 134N
8. Ryu et al., 2018	Étude comparative rétrospective	N = 40 cas (H : 26, F : 14) Âge moyen = 29,6 ans [16-57]	• RC^g	Test de LT sous radiographie RDP : Telos®	1) Test de LT des 2 genoux (80N, flexion 30°, 2 examinateurs) 2) Telos® des 2 genoux (80N, flexion 30°, 2 examinateurs) 3) GNRB® des 2 genoux (0 à 134N, 2 examinateurs) En parallèle : 3 groupes en fonction du délais entre la blessure et le diagnostic : 1) [0-10]] (n=10) 2) [11-20]] (n=12) 3) [20-30]] (n=18)	Valeur diagnostique de chaque outil Influence du délai, de la présence de pathologies associées légères, du sexe et de l'âge	Les 3 outils sont 'raisonnablement informatifs'. Le <u>GNRB®</u>, obtenant la meilleure valeur diagnostique, est considéré comme le <u>plus fiable</u>. <u>GNRB®</u> : • Seul à obtenir des valeurs de laxité différentielle ($\Delta 2$ genoux) significatives dès les 10 premiers jours. • Seul à montrer une corrélation significative entre l'âge et la détection des ruptures du LCA. • Fiabilité inter-observateur légèrement supérieure. Les 3 outils ont une valeur diagnostique relativement plus importante lorsqu'il y a une pathologie associée, ainsi que chez les femmes.

H : Hommes ; F : Femmes ; RDP : Radiographie Dynamique Passive ; GNRB® : GeNouRoB ; D : Droit ; G : Gauche ; Se : sensibilité ; Sp : spécificité ; IRM : Imagerie à Résonance Magnétique ; VPP : Valeur Prédictive Positive ; VPN : Valeur Prédictive Négative ; LT : Lachman-Trillat ; CCCs : coefficients de corrélation de concordance ; $\Delta 2$ genoux : différence entre les 2 genoux ; n.c. : non connu ; IY : Index J de Youden ; ICC : Intra-class correlation

Tableau 3. Synthèse des articles inclus dans l'étude.

9. Saravia et al., 2020	Étude Transversale diagnostique	N = 79 cas (157 genoux testés) Âge moyen = n.c. [17-60 ans], médiane = 17,5 ans	• Sain • RC	Combinaison examen clinique et IRM (référence)	1) Examen clinique et IRM 2) GNRB® : 3 mesures par genou, par 3 examinateurs ≠	Concordance entre les examinateurs, Se, Sp et IY du GNRB® au seuil de 3 mm et 6,8 mm Se = 90,7%, Sp = 40,3%, IY = 0,31 <u>6,8 mm</u> Se = 74,4%, Sp = 93,8%, IY = 0,67 ➔ GNRB® est un outil valide et fiable pour la mesure de laxité chez les patients avec une lésion du LCA.	Bonne concordance entre les examinateurs (ICC = 0,990). <u>3 mm</u> Se = 90,7%, Sp = 40,3%, IY = 0,31 <u>6,8 mm</u> Se = 74,4%, Sp = 93,8%, IY = 0,67 ➔ GNRB® est un outil valide et fiable pour la mesure de laxité chez les patients avec une lésion du LCA.
10. Sidibé et al., 2020	Étude transversale	N = 30 cas (H : 25, F : 5) Âge moyen = 27,2 ans [17-41]	• RP • RC	IRM Arthroscopie (référence)	Pas d'ordre spécifié, comparaison des diagnostics obtenus avec l'IRM et le GNRB® par rapport à l'arthroscopie per opératoire.	Se et Sp de l'IRM et du GNRB®	<u>IRM</u> RP : Se = 85,7% et Sp = 95,7% RC : Se = 95,7% et Sp = 85,7% <u>GNRB®</u> RP (seuil = 1,5 mm à 250N) : Se = 42,9% et Sp = 87% RC (seuil = 3,2 mm à 250N) : Se = 87% et Sp = 42,9% ➔ IRM est plus performante que GNRB® et GNRB® ne permet PAS d'améliorer les résultats de l'IRM.

H : Hommes ; F : Femmes ; RDP : Radiographie Dynamique Passive ; GNRB® : GeNouRoB ; D : Droit ; G : Gauche ; Se : sensibilité ; Sp : spécificité ; IRM : Imagerie à Résonance Magnétique ; VPP : Valeur Prédictive Positive ; VPN : Valeur Prédictive Négative ; LT : Lachman-Trillat ; CCCs : coefficients de corrélation de concordance ; A2 genoux : différence entre les 2 genoux ; n.c. : non connu ; IY : Index J de Youden ; ICC : Intra-class correlation

Au total, 846 sujets ont été étudiés, ce qui donne un échantillon moyen d'un peu moins de 85 sujets par étude, dont l'écrasante majorité est constituée d'hommes [62%-83%]. Le plus petit échantillon est celui de l'étude de Jenny et Arndt et le plus large celui de Cojean *et al.*. Ils ont respectivement inclus 21 et 214 sujets dans leur recherche. (**Tableau 4**). La plupart des études s'est concentrée sur les adultes. Elles ont d'ailleurs toutes une moyenne d'âge située entre 26 et 40 ans, sauf Saravia *et al.* qui ne citent pas de moyenne d'âge mais un âge médian, 17,5 ans. Seules quatre études ont inclus des patients plus jeunes, à la limite entre l'adolescence et l'âge adulte, ayant 15, 16 et 17 ans (Robert *et al.*, 2009; Ryu *et al.*, 2018; Saravia *et al.*, 2020; Sidibé *et al.*, 2020).

De plus, chaque étude a inclus des patients pour lesquels il y avait des suspicions de lésion, ou des lésions confirmées, du LCA. Il n'y a que quatre études dans lesquelles des LCA sains ont été retrouvés au sein de l'échantillon analysé (Beldame *et al.*, 2012; Cojean *et al.*, 2023a; Robert *et al.*, 2009; Saravia *et al.*, 2020). Lorsque nous nous intéressons uniquement aux LCA pathologiques, nous pouvons remarquer que les proportions de RC du LCA sont plus importantes dans 87,5 % des études (7 études sur 8). Il y a d'ailleurs trois études qui se sont intéressées exclusivement aux RC. La seule étude dans laquelle la proportion de RP est plus importante est celle de Robert *et al.* Dans cette étude, les proportions de RC, de RP et de LCA sains sont en réalité assez similaires, chaque état étant représenté dans environ un tiers des cas. Cela est dû au design de leur étude qui diffère légèrement des autres. Les auteurs ont réalisé deux recherches associées, en composant des échantillons différents pour chaque recherche. Une première expérimentation a été faite sur des genoux sains uniquement. Ensuite, ils ont effectué une deuxième expérimentation, cette fois sur des genoux pathologiques (RC et RP). Les sujets inclus dans la deuxième partie de cette étude ont été choisis spécifiquement sur base de leur pathologie. Cela explique pourquoi les proportions de LCA sains, RC et RP sont pratiquement équivalents.

Deux études, Jenny et Arndt, 2012 et Sidibé *et al.*, 2020, n'ont pas mentionné clairement les proportions de RC et RP dans leurs échantillons. Dans la première, une 'laxité chronique du LCA' était mentionnée, sans plus de précision. Dans la seconde, les auteurs évoquent des proportions de RC et RP obtenues lors

des mesures expérimentales à l'aide de l'IRM et du GNRB[®] mais ne donnent pas d'information sur les proportions obtenues lors de la mesure de référence, l'arthroscopie.

Tableau 4. Détails des échantillons de chaque étude.

Étude	N	% hommes	% LCA sains	% RP	% RC
Beldame <i>et al.</i>	157	77%	50%	17%	33%
Cojean <i>et al.</i>	214	67%	21,5%	28,5%	50%
Di Iorio <i>et al.</i>	49	71%	0%	28,6%	71,4%
Jenny & Arndt	21	67%	0%	-	-
Lefèvre <i>et al.</i>	139	62%	0%	30%	70%
Murgier <i>et al.</i>	52	77%	0%	0%	100%
Robert <i>et al.</i>	65	78%	31%	37%	32%
Ryu <i>et al.</i>	40	65%	0%	0%	100%
Saravia <i>et al.</i>	79	-	46%	0%	54%
Sidibé <i>et al.</i>	30	83%	0%	-	-

N : nombre de cas ; *RP* : ruptures partielles ; *RC* : ruptures complètes

Tous les articles ont évalué la fiabilité des mesures obtenues avec le GNRB[®] en les comparant à celles obtenues avec d'autres outils diagnostiques cités précédemment, dans le cadre du diagnostic de rupture du LCA. Cependant, toutes n'ont pas toujours utilisé exactement les mêmes paramètres de mesure ni les mêmes seuils de laxité différentielle dans la détermination de la présence et du type de rupture du LCA.

La majorité des auteurs a réalisé les mesures avec le GNRB[®] en positionnant les genoux analysés en position neutre de rotation (0°) et fléchis avec un angle de 20°, comme décrit par les créateurs de la machine (Robert *et al.*, 2009). Deux études parlent cependant d'un angle de flexion du genou légèrement différent, 25° et 30° respectivement (Jenny & Arndt, 2013; Ryu *et al.*, 2018). Une dernière étude ne précise pas l'amplitude de la flexion du genou lors du test avec le GNRB[®] (Sidibé *et al.*, 2020).

Concernant les forces appliquées par le GNRB[®] lors des mesures expérimentales, les auteurs se sont à nouveau intéressés à différentes valeurs. Plus de la moitié des articles inclus s'est intéressée à la laxité mesurée lors de l'application d'une force de 134N ou d'une force progressivement croissante de 0 à 134N (Cojean *et al.*, 2023a; Di Iorio *et al.*, 2014; Lefevre *et al.*, 2014; Robert *et al.*, 2009; Ryu *et al.*, 2018; Saravia *et al.*, 2020). Murgier *et al.* sont les seuls à s'être

intéressés à l'application d'une force de 200N. Robert *et al.* et Lefèvre *et al.* ont, eux, également poussé les mesures jusqu'à 250N tandis que Beldame *et al.*, Jenny et Arndt et Sidibé *et al.* se sont, pour leur part, intéressés uniquement à cette force de 250N sans porter d'intérêt particuliers aux paliers plus faibles.

Le **Tableau 5** reprend toutes les valeurs de sensibilité (Se) et spécificité (Sp) calculées pour le GNRB® par les différents auteurs repris dans cette revue, en fonction de la force exercée et du seuil choisi pour la détection d'une RP ou RC. L'indice de Youden calculé sur base des Se et Sp mesurées permet de donner une idée de la précision d'un moyen de mesure, dans ce cas-ci, le GNRB®. Cet indice se calcule en faisant la somme des Se et Sp puis en y soustrayant 1 ($Se + Sp - 1$). Si l'indice est égal à 1, l'outil évalué est parfait. A contrario, plus cet indice se rapproche de 0, moins l'outil est efficace (Delacour *et al.*, 2005).

Concernant les RC, Beldame *et al.* ont, lors de leur comparaison du GNRB® avec les RDP, obtenu des valeurs de Se de 64,8 % et Sp de 75,8 % pour le Telos® avec un seuil de 3 mm et des Se et Sp de 62,2 % et 75,9 % pour le GNRB® avec une valeur seuil de laxité différentielle de 1,5 mm à 250N. La méthode de Lerat *et al.* a, de son côté, obtenu une valeur diagnostique faible lors de son analyse par Beldame *et al.*. Malgré une spécificité élevée dépassant les 80 %, la sensibilité ne tournait, elle, qu'autour de 40 % (Beldame *et al.*, 2012).

Cojean *et al.* ont publié, en 2023, une étude comparant le GNRB® à l'IRM avec validation arthroscopique sur un large échantillon de patients. Lors de cette comparaison, ils ont obtenu différentes valeurs de Se et Sp pour le GNRB® et l'IRM en fonction des situations. Ils ont démontré une $Se = 100 \%$ et une $Sp = 95 \%$ à l'IRM contre une $Se = 95,65 \%$ et une $Sp = 97,5 \%$ avec le GNRB®, pour les genoux sains. Pour la détection de RC, le GNRB® à 134 N (seuil 3 mm) a perdu environ 15% en sensibilité (77,78 %) et en spécificité (85,98 %). Le même phénomène est survenu pour l'IRM, avec une perte plus importante, de l'ordre des 20 à 30% ($Se = 80,81 \%$ et $Sp = 64,49 \%$) (Cojean *et al.*, 2023a).

L'étude de Robert *et al.*, datant de 2009, écrite par les inventeurs du GNRB®, a, quant à elle, obtenu une $Se = 70 \%$ et une $Sp = 99 \%$ dans le diagnostic de lésions complètes en utilisant le GNRB® à 134N avec un seuil de 3 mm.

D'après Saravia *et al.*, 2018, le GNRB[®] posséderait une Se de 90,7 % et une Sp de 40,3 % à un seuil de 3 mm de laxité ainsi qu'une Se de 74,4 % et Sp de 93,8 % lorsque ce seuil de laxité était de 6,8 mm en considérant l'IRM comme la référence de l'état ligamentaire.

Sidibé *et al.* ont calculé une Se de 87 % et une Sp de 42,9 % pour le GNRB[®] à 250N avec un seuil de 3 mm contre 95,7 % et 85,7 % respectivement pour l'IRM. (Sidibé *et al.*, 2020)

Malgré l'inclusion de RC dans les études de Di Iorio *et al.*, 2014; Lefevre *et al.*, 2014; Murgier *et al.*, 2018; Ryu *et al.*, 2018, elles n'ont pas calculé de valeurs métrologiques pour le GNRB[®].

Au sujet des RP, l'étude de Di Iorio *et al.* n'a pas présenté directement de valeurs métrologiques pour le GNRB[®], mais les auteurs ont mesuré une détection de 72 % des RP au seuil de 1,5 mm en utilisant une force de 134N. Di Iorio *et al.* comparent le GNRB[®] à l'IRM dans leur étude. La Se qu'ils ont calculé pour l'IRM est de 84 % et la Sp de 92 % et ces valeurs de Se et Sp sont toutes les deux montées jusqu'à 100 % lors de l'utilisation de l'IRM 3 Tesla. (Di Iorio *et al.*, 2014)

Pour Cojean *et al.*, les qualités métrologiques GNRB[®] restent élevées lors du diagnostic de RP (seuil 1,3 mm), à l'inverse du l'IRM. Les valeurs du GNRB[®] à 134N n'ont quasiment pas chuté par rapport au diagnostic de RC, obtenant une Se = 73,77 % et une Sp = 85,52 %. L'IRM, par contre, a énormément diminué en Se, passant à 29,51 %, tout en maintenant une Sp assez élevée, égale à 88,97 %. (Cojean *et al.*, 2023a)

Lefevre *et al.* ont obtenu une Se = 83,2 % et une Sp = 64,3% pour le GNRB[®] à 134N avec un seuil de laxité différentielle de 2 mm. Pour le GNRB[®] à 250N avec un seuil de 2,5 mm, Lefevre *et al.* ont calculé une Se = 84 % et Sp = 81 %. À l'inverse, ils ont obtenu une Se = 81,5 % et Sp = 59,5 % pour le Telos[®] à 250N, avec un seuil de 3,6 mm.

Les valeurs de Se et Sp présentées par Robert *et al.* pour la mesure diagnostique de lésions partielles du LCA à l'aide du GNRB[®] (134N, seuil 1,5 mm) étaient respectivement égales à 80 % (Se) et 87 % (Sp).

Sidibé *et al.* ont également calculé des valeurs de Se et Sp pour les RP, résultant en une Se = 42,9 % et une Sp = 87 % pour le GNRB[®] (250N, seuil différentiel de 1,5 mm) contre une Se = 85,7 % et une Sp = 95,7 % pour l'IRM.

À l'instar des études de Lefevre *et al.*, Murgier *et al.* et Ryu *et al.* pour les RC, Beldame *et al.* n'ont pas publié de valeurs métrologiques pour le GNRB® dans le diagnostic de RP du LCA.

Tableau 5. Sensibilités (Se), spécificités (Sp) et Indice de Youden du GNRB® dans la détection de ruptures partielles et complètes du ligament croisé antérieur (LCA).

	Seuil de laxité	GNRB® à 134N		Indice de Youden	GNRB® à 250N		Indice de Youden
		Se	Sp		Se	Sp	
RUPTURES COMPLÈTES							
Beldame <i>et al.</i> , (2012)	1,5 mm	59,4 %	93,1 %	0,525	62,2 %	75,9 %	0,381
Cojean <i>et al.</i> , (2023)	3 mm	77,8 %	86 %	0,638	-	-	-
Di Iorio <i>et al.</i> , (2014)	3 mm	-	-	-	-	-	-
Lefevre <i>et al.</i> , (2014)	-	-	-	-	-	-	-
Murgier <i>et al.</i> , (2018)	-	-	-	-	-	-	-
Robert <i>et al.</i> , (2009)	3 mm	70 %	99 %	0,690	-	-	-
Ryu <i>et al.</i> , (2018)	-	-	-	-	-	-	-
Saravia <i>et al.</i> , (2020)	3 mm	90,7 %	40,3 %	0,310	-	-	-
	6,8 mm	74,4 %	93,8 %	0,682	-	-	-
Sidibé <i>et al.</i> , (2020)	3,2 mm	-	-	-	87 %	42,9 %	0,299
RUPTURES PARTIELLES							
Beldame <i>et al.</i> , (2012)	-	-	-	-	-	-	-
Cojean <i>et al.</i> , (2023)	1,3 mm	73,8 %	85,5 %	0,593	-	-	-
Di Iorio <i>et al.</i> , (2014)	1,5 mm	72 %	-	-	-	-	-
Lefevre <i>et al.</i> , (2014)	2 mm	83,2 %	64,3 %	0,475	-	-	-
	2,5 mm	-	-	-	84 %	81 %	0,650
Robert <i>et al.</i> , (2009)	1,5 mm	80 %	87 %	0,670	-	-	-
Sidibé <i>et al.</i> , (2020)	1,5 mm	-	-	-	42,9 %	87 %	0,299

III. DISCUSSION

Jusqu'à présent, le diagnostic de lésion du LCA est essentiellement basé sur l'examen clinique, associé à l'IRM et éventuellement à une mesure quantitative de la laxité antéro-postérieure du genou (Brady & Weiss, 2018; Cojean *et al.*, 2023a). Pour ce faire, il existe différents outils et techniques, dont les plus communément utilisés et étudiés sont le KT-1000[®], les radiographies dynamiques passives (RDP) à l'aide du Telos[®] et/ou de la méthode décrite par Lerat *et al.* au début des années 2000, le Rolimeter[®] et, forcément, le GNRB[®] (Murgier *et al.*, 2018).

L'objectif de ce mémoire était de réaliser une revue de la littérature analysant la plus-value du GNRB[®] dans le diagnostic de ruptures du LCA. Pour ce faire, dix articles scientifiques comparant le GNRB[®] à tout autre outil d'évaluation ont été considérés.

GNRB[®] comparé aux RDP

Dans un premier temps, portons notre attention sur la comparaison entre le GNRB[®] et les RDP, réalisées avec le Téllos[®] et selon la méthode « Lerat », dans le diagnostic de lésion du LCA. Les RDP sont l'outil qui a reçu le plus d'attention de la part des auteurs que nous avons analysés. La moitié des études incluses dans cette revue s'y sont d'ailleurs intéressées.

A notre connaissance, Beldame *et al.*, 2012 sont les premiers à avoir entrepris la comparaison de ces deux machines sur des patients dont le LCA est lésé. Ils ont confirmé que le Telos[®] et le GNRB[®] avaient tous les deux bel et bien une valeur diagnostique lors du bilan des RC du LCA et qu'elles étaient à peu près équivalentes. Cependant, ces valeurs se sont avérées relativement limitées. À performances diagnostiques égales, Beldame *et al.*, recommandent le GNRB[®] au dépend du Telos[®] car le GNRB[®] est plus facile d'utilisation et qu'il a l'avantage majeur de ne pas être irradiant, donnant la possibilité d'être utilisé de manière répétée lors du suivi du patient. Concernant les RP, les auteurs de cette étude n'ont pas retrouvé de valeur diagnostique ni du GNRB[®], ni du Telos[®], à l'inverse des

inventeurs du GNRB® (Robert *et al.*, 2009) et de l'étude de Lefevre *et al.*, publiée en 2014.

Cette dernière a conclu que tant le Telos® que le GNRB® possédaient assurément une valeur diagnostique lors du dépistage de RP. De plus, les résultats obtenus par Lefevre *et al.* penchaient en faveur du GNRB® à 250N avec un seuil de laxité de 2,5 mm pour les RP. Ces différences importantes entre l'étude de Beldame *et al.* et de Lefevre *et al.* pourraient être attribuées à la grosse divergence des seuils utilisés par les chercheurs. La force principale de ces deux études est leur échantillon. Ils sont relativement importants ($N_{\text{Beldame et al.}} = 157$ et $N_{\text{Lefevre et al.}} = 139$) et sont comparables en taille et en proportions, bien que l'échantillon de l'étude de Lefevre *et al.* avait une proportion d'hommes 15 % plus faible que l'échantillon de Beldame *et al.*. Cela pourrait jouer sur les valeurs de laxité obtenues, car celles-ci ont tendance à être plus importantes chez les femmes que chez les hommes (Ryu *et al.*, 2018).

À l'inverse des trouvailles de Beldame *et al.* selon lesquels le GNRB® et le Telos® possèdent une valeur diagnostique équivalente mais faible, l'étude de Ryu *et al.*, datant de 2018, soutient que tant le Telos® que le GNRB® sont « assez informatifs » dans le diagnostic aigu de RC du LCA. De plus, toutes les valeurs diagnostiques (aire sous la courbe, différences significatives de laxité dès 10j post-blessure, fiabilité inter-observateur, fiabilité dans une population âgée ou féminine) obtenues pour le GNRB® étaient meilleures que celles du Telos®. Cependant, les auteurs de cette étude n'ont pas calculé de Se et de Sp, rendant une comparaison purement objective impossible. L'échantillon analysé par Ryu *et al.* était fort limité, d'une part par le nombre de sujets inclus, d'autre part par l'inclusion exclusive de RC, ne leur permettant pas de s'intéresser aux RP.

La conclusion que Jenny & Arndt, 2013 ont tiré de leur étude était similaire à celle de Beldame *et al.* : « *Le système GNRB® offre la même fiabilité de mesure que les clichés dynamiques (RDP) en évitant toute irradiation, et doit donc être privilégié* » (Jenny & Arndt, 2013). Cependant, il a également été conclu dans cette étude que les deux outils avaient tendance à minorer la laxité antérieure réelle telle qu'évaluée par un système de navigation peropératoire validé, pris pour référence. En revanche, cette étude comporte quelques limitations dont les plus gênantes sont une faible taille de l'échantillon et des circonstances de prise de mesures très inégales. La mesure avec le GNRB® a été réalisée sur des patients bien éveillés, à

l'inverse des RDP, faites sous anesthésie. Cela pourrait jouer un rôle sur l'état de relâchement musculaire et tissulaire du patient et par conséquent sur la laxité mesurée lors des différents tests.

Dans les expériences menées par Murgier *et al.*, les mesures du Telos® étaient largement dispersées et la corrélation entre les deux outils analysés était mauvaise. D'après cette étude, il n'est par conséquent pas possible de comparer des valeurs de laxité différentielle d'un appareil à l'autre, ce qui démontre qu'il est important de rester très vigilant lorsqu'elles sont comparées dans la littérature. Quoiqu'il en soit, les auteurs n'ont malheureusement pas calculé les valeurs métrologiques de chacun des quatre outils qu'ils ont étudiés, les empêchant de déterminer quel outil était le plus fiable.

Comme le démontre le **Tableau 6**, deux des cinq études qui se sont penchées sur la comparaison du GNRB® aux RDP ont obtenu des valeurs diagnostiques supérieures pour le GNRB®. Deux autres ont tiré la conclusion que ces deux outils étaient équivalents en terme de capacité diagnostique, mais ont néanmoins tous les deux préféré le GNRB® aux RDP de par son caractère non-irradiant. La dernière étude qui s'est intéressée à ces deux outils n'a pas su désigner un outil supérieur à l'autre.

Ces conclusions globalement positives à l'égard du GNRB® en comparaison aux RDP, et en particulier au Telos®, sont comparables à ce que nous pouvons retrouver dans des études effectuées dans d'autres conditions que celles incluses dans la présente revue (Bouguennec *et al.*, 2015). L'étude publiée par Bouguennec et son équipe porte sur la reproductibilité des mesures du GNRB® par rapport au Telos®, sur des genoux sains. Elle a, elle aussi, conclu que les mesures du GNRB® étaient plus reproductibles que celles du Telos®, et que le GNRB® avait l'avantage de ne pas être irradiant. Les auteurs ont cependant limité la force exercée par le GNRB® à 89N, ce qui est considérablement moins important que les forces appliquées dans les études incluses dans cette revue.

Tableau 6. Quel outil du GNRB® ou des RDP (Telos® - Lerat) possède la meilleure valeur diagnostique ?

GNRB® ou RDP ?		
Beldame <i>et al.</i> , 2012		
Jenny et Arndt, 2013		 : Ne sait pas dire
Lefevre <i>et al.</i> , 2014		 : GNRB® = RDP
Murgier <i>et al.</i> , 2018		 : GNRB® > RDP
Ryu <i>et al.</i> , 2018		

Flèche rouge descendante : l'étude n'a pas su se prononcer ;

Flèche jaune horizontale : les deux outils sont équivalents, mais GNRB® est préféré car non-irradiant ;

Flèche verte ascendante : le GNRB® a une meilleure valeur diagnostique que les RDP.

GNRB® comparé à l'IRM

Intéressons-nous maintenant aux articles qui ont effectué la comparaison du GNRB® à l'IRM. Dans cette revue, l'IRM est le second outil le plus comparé au GNRB® dans le cadre du diagnostic de rupture du LCA. Quatre études sur dix s'y sont notamment intéressées.

Une des premières études à s'être attaquée à la comparaison, sur genoux pathologiques, du GNRB® à l'IRM dans le diagnostic de rupture de LCA, date d'il y a tout juste dix ans (Di Iorio *et al.*, 2014). Dans cette étude, il a été conclu que tant le GNRB® que l'IRM étaient capables de différencier significativement les RP des RC. Concernant les RP, la Se calculée pour l'IRM dans cette étude est de 84 % et la Sp de 92 %. Ces valeurs de Se et Sp grimpent toutes les deux même jusqu'à 100 % lors de l'utilisation de l'IRM 3 Tesla, mais cette prédiction de qualité parfaite ne repose que sur la mesure de cinq cas et est donc à nuancer. Les auteurs n'ont pas calculé directement de valeurs métrologiques pour le GNRB®, mais présentent une détection de 72 % des RP. Cette valeur est inférieure à la sensibilité de 80 % calculée par les inventeurs de la machine pour le diagnostic de RP au même seuil pour une force appliquée identique (Robert *et al.*, 2009). Le nombre de sujets inclus par Di Iorio *et al.* limite la puissance de leur étude et des conclusions tirées.

À première vue, l'étude de Saravia *et al.*, 2020 a produit de bons résultats lors de la comparaison des résultats obtenus à l'aide du GNRB® et ceux de l'IRM, combinée à l'examen clinique. Plusieurs choses posent malgré tout problème dans cet article. Différentes incohérences ont été retrouvées ; entre autres leur échantillon, sa distribution, les patients inclus, ... et les raisons sont souvent floues, menant à des contradictions. De plus, les auteurs ont utilisé des seuils de laxité de

3 mm et 6,8 mm pour chaque genou individuellement, à l'inverse de la laxité différentielle qui s'intéresse à la différence entre le genou sain et le genou pathologique, utilisée dans la majorité des études et par les inventeurs du GNRB® (Robert *et al.*, 2009). Il existe pourtant une laxité physiologique dans les genoux sains (Alqahtani *et al.*, 2018; Smith *et al.*, 2022). Un tel seuil de 3 mm pour diagnostiquer une RC du LCA n'a donc pas vraiment lieu d'être. Saravia et son équipe concluent de leur étude que le GNRB® est un instrument fiable et valide dans le diagnostic de RC du LCA. Cette étude semble plutôt nous démontrer que le GNRB®, malgré une précision et une reproductibilité importantes, n'est pas aussi puissant que l'IRM à l'égard du diagnostic de lésion du LCA, cette dernière ayant été la référence pour les calculs de Se et Sp du GNRB®.

Sidibé *et al.*, 2020 se sont également penchés sur la comparaison du GNRB® à l'IRM dans leur étude transversale avec validation arthroscopique. La conclusion de leur étude était que le GNRB® est moins performant que l'IRM, qu'il n'a pas d'apport diagnostique et qu'il ne permet pas d'améliorer les mesures de l'imagerie. D'après leur étude, la laxité du genou évaluée par le GNRB® ne dépend pas uniquement du LCA (Sidibé *et al.*, 2020). En revanche, ils n'ont inclus que 30 patients dans leur étude, ce qui limite fortement la puissance de leur étude. Par ailleurs, ils n'ont jamais présenté, dans leur article, les valeurs obtenues lors de l'arthroscopie de référence, ne permettant pas de comparer les proportions de RC et RP retrouvées avec le GNRB® (80%/20%) ou l'IRM (76,7%/23,3%). Les auteurs n'ont donné que très peu d'informations au sujet des calculs effectués et des statistiques employées pour déterminer que l'IRM surpassait si franchement le GNRB®. Les valeurs de Se et Sp mises en avant par Sidibé *et al.* nous ont interpellés, car il s'avère qu'elles sont exactement inversées lorsqu'on compare les valeurs des RC et des RP, tant au sujet du GNRB® que de l'IRM.

Saravia *et al.*, 2020 ne sont pas les seuls à avoir rapporté une comparaison positive entre les valeurs du GNRB® et celles de l'IRM. L'étude de Cojean *et al.* (2023a) est également très positive quant à la capacité diagnostique du GNRB® par rapport à celle de l'IRM. Cojean et son équipe ont conclu trois choses. Premièrement, le GNRB® et l'IRM sont au moins aussi performants pour détecter les LCA sains et les RC du LCA. Deuxièmement, le GNRB® est capable de mieux détecter les RP du LCA que l'IRM. Et dernièrement, la combinaison de la mesure du GNRB® et de l'IRM permet d'augmenter l'exactitude du diagnostic de l'état du

LCA, qu'il soit sain, partiellement lésé ou complètement rompu. Il faut cependant souligner que l'IRM utilisée était une IRM 1,5 Tesla, sans utilisation de plans obliques. Or, différentes études dont celle de Di Iorio *et al.*, 2014 ont obtenu de meilleurs résultats avec des IRM 3 Tesla. De plus, un des co-auteurs de l'étude de Cojean *et al.* n'est autre que le Dr. Henri Robert, co-inventeur du GNRB[®], ce qui soulève quelques doutes sur l'entière objectivité de cette étude.

Deux des quatre études ayant comparé le GNRB[®] à l'IRM ont trouvé que l'IRM était clairement supérieure au GNRB[®], tandis que les deux autres études ont déclaré que le GNRB[®] et l'IRM étaient au moins aussi performants pour détecter les RC, une des deux allant jusqu'à soutenir l'idée que le GNRB[®] était plus performant que l'IRM dans les RP (**Tableau 7**).

Tableau 7. Quel outil du GNRB[®] ou de l'IRM possède la meilleure valeur diagnostique ?

	GNRB [®] ou IRM ?		
	RP	RC	
Di Iorio <i>et al.</i> , 2014	↓	↓	↓ : GNRB [®] < IRM
Saravia <i>et al.</i> , 2020	/	→	→ : GNRB [®] = IRM
Sidibé <i>et al.</i> , 2020	↓	↓	↑ : GNRB [®] > IRM
Cojean <i>et al.</i> , 2023	↑	→	

Flèche rouge descendante : valeur diagnostique du GNRB[®] est moindre que celle de l'IRM ;

Flèche jaune horizontale : valeur diagnostique du GNRB[®] comparable à celle de l'IRM ;

Flèche verte ascendante : le GNRB[®] a une meilleure valeur diagnostique que l'IRM.

GNRB[®] comparé à l'examen clinique

La comparaison du GNRB[®] et de l'examen clinique, le test de Lachman-Trillat et le test de ressaut, a été abordée par deux études incluses dans cette revue. Les deux études ont obtenu des résultats plus probants pour le GNRB[®] que l'examen clinique, mais n'ont pas réalisé l'examen clinique dans exactement les mêmes conditions. La première, celle de Di Iorio *et al.*, a réalisé l'examen clinique sous anesthésie et n'est pas parvenue à démontrer une différence significative entre une RP et une RC, à l'inverse du GNRB[®] (Di Iorio *et al.*, 2014). Dans la seconde étude, celle publiée en 2018 par Ryu *et al.*, le test de Lachman-Trillat était effectué avec une force de 80N mesurée par dynamomètre et la laxité évaluée par une radiographie. Cette étude a tout de même démontré une valeur diagnostique 'relativement informative' de l'examen clinique, mais toutes les valeurs diagnostiques obtenues pour le GNRB[®] étaient meilleures, en particulier dans les lésions aiguës. Chacune des deux études avait des petits échantillons, rendant leur

puissance limitée. Par ailleurs, aucune des deux études n'a calculé de qualités métrologiques pour le test de Lachman. Cela complique la comparaison objective des deux outils ainsi que la comparaison avec les données de littérature.

La littérature existante semble majoritairement s'accorder sur le fait que le test de Lachman est le meilleur test clinique pour le diagnostic de lésions du LCA (Benjaminse *et al.*, 2006; Huang *et al.*, 2016; van Eck *et al.*, 2013) bien que cet avis n'est pas unanime et a récemment été quelque peu nuancé (Sokal *et al.*, 2022).

Tableau 8. Quel outil du GNRB® ou de l'examen clinique possède la meilleure valeur diagnostique ?

GNRB® ou Examen clinique ?		
Di Iorio <i>et al.</i> , 2014		 : GNRB® < Examen clinique
Ryu <i>et al.</i> , 2018		 : GNRB® = Examen clinique
		 : GNRB® > Examen clinique

Flèche rouge descendante : valeur diagnostique du GNRB® est moindre que celle de l'examen clinique ;
 Flèche jaune horizontale : valeur diagnostique du GNRB® comparable à celle de l'examen clinique ;
 Flèche verte ascendante : le GNRB® a une meilleure valeur diagnostique que l'examen clinique.

GNRB® comparé au KT-1000®

Selon Robert *et al.*, 2009, les mesures effectuées à l'aide du GNRB® sont plus reproductibles que celles du KT-1000®, car il y avait un effet opérateur significatif avec le KT-1000® non-présent avec le GNRB® ainsi qu'un effet d'apprentissage également propre au KT-1000®. Ils ont conclu que le GNRB® avait une bonne Se et Sp lors de la détection de RP et RC. (> 80 % pour toutes, sauf la Se pour les RP ; 70 %). Il faut cependant nuancer les conclusions de cette étude, car elle a été publiée par l'inventeur du GNRB®.

De son côté, l'étude de Murgier *et al.* a conclu que la relation entre les mesures (en mm) du GNRB® et du KT-1000® était moyenne. C'est une meilleure relation qu'entre le GNRB® et les deux autres outils les plus utilisés ; le Telos®, le Rolimeter®.

Ces conclusions sont comparables à celles retrouvées dans la littérature lors de la comparaison du GNRB® au KT-1000® sur des genoux sains (Collette *et al.*, 2012).

Tableau 9. Quel outil du GNRB® ou du KT-1000® possède la meilleure valeur diagnostique ?

GNRB® ou KT-1000® ?		
Murgier <i>et al.</i> , 2018		 : GNRB® < KT-1000®
Robert <i>et al.</i> , 2009		 : GNRB® = KT-1000®  : GNRB® > KT-1000®

Flèche rouge descendante : valeur diagnostique du GNRB® est moindre que celle du KT-1000® ;

Flèche jaune horizontale : valeur diagnostique du GNRB® comparable à celle du KT-1000® ;

Flèche verte ascendante : le GNRB® a une meilleure valeur diagnostique que le KT-1000®.

GNRB® comparé au Rolimeter®

Murgier *et al.*, 2018 sont les seuls à avoir également comparé le GNRB® au Rolimeter®. À nouveau, la conclusion de cette étude était que les mesures de laxité en mm n'étaient pas comparables entre le GNRB® et le Rolimeter®.

Quelle suite ?

Nous pouvons déduire plusieurs choses sur base des informations recueillies dans cette revue.

Premièrement, la majorité des auteurs a reconnu une efficacité et une utilité certaine au GNRB®. Lorsqu'il est comparé aux RDP, à l'examen clinique ou au KT-1000®, le GNRB® a obtenu, pratiquement à chaque fois, au minimum une qualité diagnostique similaire, voire une mesure plus efficace, plus reproductible et plus précise que l'outil auquel il était comparé. Cela nous démontre que le GNRB® a bel et bien sa place dans le diagnostic des lésions du LCA.

Deuxièmement, néanmoins, comparé à l'IRM, les résultats sont plus mitigés, tout en restant globalement positifs. Plus d'études comparant la qualité diagnostique du GNRB® par rapport à l'IRM seraient intéressantes pour confirmer les bons résultats obtenus par Cojean *et al.* tant pour les RC que les RP.

À ce jour, le GNRB® ne nous semble pas en mesure de remplacer l'IRM dans un futur proche, l'IRM étant trop ancrée dans les habitudes et les décisions en première intention lors d'une suspicion de lésion du LCA, comme en témoignent les articles de Cojean *et al.*, 2023a; Masik *et al.*, 2022.

Par contre, il nous semble raisonnable de penser qu'il y a bien une valeur ajoutée au GNRB® dans un parcours de soin classique lors de la suspicion d'entorse

du LCA. L'intérêt du GNRB® serait de diminuer le nombre d'IRM faites en première intention pour ne garder l'IRM qu'en cas de test GNRB® positif, afin de prendre la décision thérapeutique adéquate en combinant les informations du GNRB® aux images de l'IRM pour connaître l'étendue des dégâts et la présence de lésions associées.

Une des limites du GNRB®, soulevée par différents auteurs, est le fait qu'il évalue exclusivement la laxité antéro-postérieure, en deux dimensions (Beldame *et al.*, 2012; Di Iorio *et al.*, 2014; Jenny & Arndt, 2013). Or, le LCA a également pour rôle de limiter la rotation excessive du tibia par rapport au fémur, ainsi, une lésion du LCA devrait en théorie également causer une augmentation de la laxité rotatoire tibiale.

Une étude de Mouton *et al.*, 2015 a démontré que la combinaison de la mesure de laxité antérieure et rotatoire améliorerait le diagnostic de blessures du LCA. Une nouvelle machine créée par Genourob, société créatrice du GNRB®, a vu le jour récemment et n'a pas encore beaucoup été étudiée, le Dyneelax®. Il s'agit essentiellement d'une version évoluée du GNRB® à laquelle a été rajoutée la mesure de la laxité rotatoire du tibia par rapport au fémur (Nascimento *et al.*, 2024). Les premières études semblent encourageantes (Cojean *et al.*, 2022, 2023b; Nascimento *et al.*, 2024). De plus, le projet de développement du Dyneelax® a également été soutenu par l'Union Européenne, démontrant son sérieux et des objectifs reconnus utiles à la société.

Cette piste nous semble intéressante à explorer à l'avenir afin d'augmenter la capacité diagnostique de l'arthrométrie dans un but de diminuer la prescription d'IRM en première intention sans nécessité réelle, car elle est coûteuse pour le patient et la société et que cela pourrait permettre de désengorger les listes d'attente souvent bondées pour l'obtention d'un rendez-vous pour une IRM.

Certaines limites peuvent être observées concernant la présente revue de la littérature.

D'une part, nous avons inclus des études de différents designs qui n'ont pas toutes calculé de valeurs diagnostiques quantitatives (Se – Sp). Cela a donc rendu les comparaisons entre elles parfois laborieuses. Deux types majeurs de designs ont été inclus ; il y avait des études de cohorte ou transversales, pour lesquelles le NIH

a spécifiquement été créé, mais aussi des études diagnostiques et/ou comparatives. Le NIH n'est pas forcément le meilleur outil pour évaluer leur qualité méthodologique, et l'utilisation de l'échelle QUADAS-2 aurait pu être plus pertinente pour ces études-là. Ceci dit, il était nécessaire de faire un choix pour ne garder qu'un seul outil d'évaluation méthodologique et ainsi permettre la comparaison des études les unes avec les autres.

D'autre part, nous nous sommes intéressés uniquement à la précision diagnostique pour les lésions isolées du LCA, ne prenant pas en compte les blessures du LCA avec lésions associées (lésions méniscales, LCP, ligaments latéraux, ...) qui sont très fréquentes chez les patients avec une suspicion de rupture du LCA (Stranger *et al.*, 2023).

De plus, le processus de sélection des articles n'a été effectué que par un seul observateur (R.W.), favorisant le risque d'erreurs liées à l'inclusion ou l'exclusion des articles dans notre revue.

CONCLUSION

La plupart des études incluses dans cette étude ont rapporté une capacité diagnostique importante du GNRB[®], toujours a minima équivalente et très souvent supérieure à celles des outils comparés, mise à part lors de la comparaison avec l'IRM. Cela confirme l'hypothèse selon laquelle le GNRB[®] a une valeur-ajoutée lors du diagnostic de lésions du LCA, en particulier lorsqu'il faut déterminer si la lésion est partielle ou totale.

Le GNRB[®] n'est pas (encore) en mesure de rivaliser avec l'IRM, car ce dernier est ancré profondément dans nos habitudes face à la suspicion d'une rupture du LCA.

Le GNRB[®] pourrait cependant être utilisé à bon escient en prenant plus de place dans le parcours de soin classique, par exemple pour diminuer le nombre d'IRM prescrites en première intention en ne prescrivant une IRM qu'en cas de lésion avérée lors de l'examen avec le GNRB[®]. Le GNRB[®] pourrait également être utilisé pour discriminer une lésion partielle et complète lorsque des doutes persistent à l'IRM et ainsi orienter la décision thérapeutique.

Un nouvel outil diagnostique, le Dyneelax[®], combine l'évaluation de la laxité antéro-postérieure et de la laxité en rotation. Cette solution pourrait théoriquement améliorer encore l'exactitude des mesures de l'arthrométrie mais des recherches supplémentaires à ce sujet sont nécessaires pour pouvoir l'affirmer avec certitude.

BIBLIOGRAPHIE

1. Alqahtani, Y., Murgier, J., Beaufils, P., Boisrenoult, P., Steltzlen, C., & Pujol, N. (2018). Anterior tibial laxity using the GNRB® device in healthy knees. *The Knee*, 25(1), 34-39. <https://doi.org/10.1016/j.knee.2017.03.004>
2. Beldame, J., Mouchel, S., Bertiaux, S., Adam, J.-M., Mouilhade, F., Roussignol, X., & Dujardin, F. (2012). Anterior knee laxity measurement : Comparison of passive stress radiographs Telos(®) and « Lerat », and GNRB(®) arthrometer. *Orthopaedics & Traumatology, Surgery & Research : OTSR*, 98(7), 744-750. <https://doi.org/10.1016/j.otsr.2012.05.017>
3. Benjaminse, A., Gokeler, A., & van der Schans, C. P. (2006). Clinical diagnosis of an anterior cruciate ligament rupture : A meta-analysis. *The Journal of Orthopaedic and Sports Physical Therapy*, 36(5), 267-288. <https://doi.org/10.2519/jospt.2006.2011>
4. Bouguennec, N., Odri, G. A., Graveleau, N., & Colombet, P. (2015). Comparative reproducibility of TELOS™ and GNRB® for instrumental measurement of anterior tibial translation in normal knees. *Orthopaedics & Traumatology, Surgery & Research : OTSR*, 101(3), 301-305. <https://doi.org/10.1016/j.otsr.2015.01.007>
5. Brady, M. P., & Weiss, W. (2018). Clinical Diagnostic Tests Versus MRI Diagnosis of ACL Tears. *Journal of Sport Rehabilitation*, 27(6), 596-600. <https://doi.org/10.1123/jsr.2016-0188>
6. Cojean, T., Batailler, C., Robert, H., & Cheze, L. (2022). Reliability evaluation of the DYNEELAX, a new knee arthrometer. *SB 2022, 47eme Congrès de la Société de Biomécanique*, 25, pp S73-S75. <https://hal.science/hal-03862298>
7. Cojean, T., Batailler, C., Robert, H., & Cheze, L. (2023a). GNRB® laximeter with magnetic resonance imaging in clinical practice for complete and partial anterior cruciate ligament tears detection : A prospective diagnostic study with arthroscopic validation on 214 patients. *The Knee*, 42, 373-381. <https://doi.org/10.1016/j.knee.2023.03.017>
8. Cojean, T., Batailler, C., Robert, H., & Cheze, L. (2023b). Sensitivity, repeatability and reproducibility study with a leg prototype of a recently developed knee arthrometer : The DYNEELAX®. *Medicine in Novel Technology and Devices*, 19((Cojean T., theo.cojean@etu.univ-lyon1.fr; Batailler C., cecile.batailler@chu-lyon.fr; Cheze L., laurence.cheze@univ-lyon1.fr) Univ Lyon, Univ Gustave Eiffel, Univ Claude Bernard Lyon 1, LBMC UMR_T 9406, Lyon, France). Embase. <https://doi.org/10.1016/j.medntd.2023.100254>

9. Collette, M., Courville, J., Forton, M., & Gagnière, B. (2012). Objective evaluation of anterior knee laxity; comparison of the KT-1000 and GNRB® arthrometers. *Knee Surgery, Sports Traumatology, Arthroscopy: Official Journal of the ESSKA*, 20(11), 2233-2238. <https://doi.org/10.1007/s00167-011-1869-2>
10. Colombet, P., Dejour, D., Panisset, J.-C., & Siebold, R. (2010). Current concept of partial anterior cruciate ligament ruptures. *Orthopaedics & Traumatology: Surgery & Research*, 96(8, Supplement), S109-S118. <https://doi.org/10.1016/j.otsr.2010.09.003>
11. De Carli, A., Ferretti, A., & Maestri, B. (2022). Mechanism of Injury of ACL Tears. In A. Ferretti (Éd.), *Anterolateral Rotatory Instability in ACL Deficient Knee* (p. 71-81). Springer International Publishing. https://doi.org/10.1007/978-3-031-00115-4_7
12. Delacour, H., Servonnet, A., Perrot, A., Vigezzi, J. F., & Ramirez, J. M. (2005). La courbe ROC (receiver operating characteristic): Principes et principales applications en biologie clinique. *Ann Biol Clin*, 63.
13. Di Iorio, A., Carnesecchi, O., Philppot, R., & Farizon, F. (2014). Multiscale analysis of anterior cruciate ruptures : Prospective study of 49 cases. *Orthopaedics & Traumatology, Surgery & Research: OTSR*, 100(7), 751-754. <https://doi.org/10.1016/j.otsr.2014.02.011>
14. Ferretti, A., Maestri, B., & Iannotti, F. (2022). Diagnostics in ACL-Deficient Knee. In A. Ferretti (Éd.), *Anterolateral Rotatory Instability in ACL Deficient Knee* (p. 83-99). Springer International Publishing. https://doi.org/10.1007/978-3-031-00115-4_8
15. Huang, W., Zhang, Y., Yao, Z., & Ma, L. (2016). Clinical examination of anterior cruciate ligament rupture : A systematic review and meta-analysis. *Acta Orthopaedica Et Traumatologica Turcica*, 50(1), 22-31. <https://doi.org/10.3944/AOTT.2016.14.0283>
16. Jenny, J.-Y., & Arndt, J. (2013). Anterior knee laxity measurement using stress radiographs and the GNRB® system versus intraoperative navigation. *Orthopaedics and Traumatology: Surgery and Research*, 99(6 SUPPL), S297-S300. Embase. <https://doi.org/10.1016/j.otsr.2013.07.008>
17. Jenny, J.-Y., Puliero, B., Schockmel, G., Harnoist, S., & Clavert, P. (2017). Experimental validation of the GNRB® for measuring anterior tibial translation. *Orthopaedics & Traumatology: Surgery & Research*, 103(3), 363-366. <https://doi.org/10.1016/j.otsr.2016.12.011>
18. Khan, N., Dockter, E., Fithian, D., Navarro, R., & Luetzow, W. (2017). Development of Arthrometry. In V. Musahl, J. Karlsson, R. Kuroda, & S. Zaffagnini (Éds.), *Rotatory Knee Instability : An Evidence Based Approach* (p. 115-129). Springer International Publishing. <https://doi.org/10.1007/978-3-319->

19. Koga, H., Nakamae, A., Shima, Y., Iwasa, J., Myklebust, G., Engebretsen, L., Bahr, R., & Krosshaug, T. (2010). Mechanisms for Noncontact Anterior Cruciate Ligament Injuries : Knee Joint Kinematics in 10 Injury Situations from Female Team Handball and Basketball. *The American Journal of Sports Medicine*, 38(11), 2218-2225. <https://doi.org/10.1177/0363546510373570>
20. Larwa, J., Stoy, C., Chafetz, R. S., Boniello, M., & Franklin, C. (2021). Stiff Landings, Core Stability, and Dynamic Knee Valgus : A Systematic Review on Documented Anterior Cruciate Ligament Ruptures in Male and Female Athletes. *International Journal of Environmental Research and Public Health*, 18(7), Article 7. <https://doi.org/10.3390/ijerph18073826>
21. Lefevre, N., Bohu, Y., Naouri, J. F., Klouche, S., & Herman, S. (2014). Validity of GNRB® arthrometer compared to Telos™ in the assessment of partial anterior cruciate ligament tears. *Knee Surgery, Sports Traumatology, Arthroscopy*, 22(2), 285-290. <https://doi.org/10.1007/s00167-013-2384-4>
22. Ma, L.-L., Wang, Y.-Y., Yang, Z.-H., Huang, D., Weng, H., & Zeng, X.-T. (2020). Methodological quality (risk of bias) assessment tools for primary and secondary medical studies : What are they and which is better? *Military Medical Research*, 7(1), 7. <https://doi.org/10.1186/s40779-020-00238-8>
23. Masik, D., Robert, H., Godet, R., & Raclot, V. (2022). Intérêt d'un nouveau parcours de soins dans la prise en charge des entorses du genou. *Médecine*, 18(2), 82-86. <https://doi.org/10.1684/med.2022.740>
24. Moses, B., Orchard, J., & Orchard, J. (2012). Systematic Review : Annual Incidence of ACL Injury and Surgery in Various Populations. *Research in Sports Medicine*, 20(3-4), 157-179. <https://doi.org/10.1080/15438627.2012.680633>
25. Mouton, C., Theisen, D., Meyer, T., Agostinis, H., Nührenbörger, C., Pape, D., & Seil, R. (2015). Combined anterior and rotational knee laxity measurements improve the diagnosis of anterior cruciate ligament injuries. *Knee Surgery, Sports Traumatology, Arthroscopy*, 23(10), 2859-2867. <https://doi.org/10.1007/s00167-015-3757-7>
26. Murgier, J., Béranger, J. S., Boisrenoult, P., Steltzlen, C., & Pujol, N. (2018). Prospective comparative study of knee laxity with four different methods in anterior cruciate ligament tears. *International Orthopaedics*, 42(8), 1845-1851. <https://doi.org/10.1007/s00264-018-3791-4>
27. Nascimento, N., Kotsifaki, R., Papakostas, E., Zikria, B. A., Alkhelaifi, K., Hagert, E., Olory, B., D'Hooghe, P., & Whiteley, R. (2024). DYNEELAX Robotic Arthrometer Reliability and Feasibility on Healthy and Anterior Cruciate Ligament Injured/Reconstructed Persons. *Translational Sports Medicine*, 2024, e3413466.

28. Noyes, F., & Barber-Westin, S. (2016). *ACL Injury : Everything You Need to Know to Make the Right Treatment Decision*: ERROR:#NAME? Publish Green. <http://ebookcentral.proquest.com/lib/uclouvainbe/detail.action?docID=4517475>
29. Pappas, E., Zampeli, F., Xergia, S. A., & Georgoulis, A. D. (2013). Lessons learned from the last 20 years of ACL-related in vivo-biomechanics research of the knee joint. *Knee Surgery, Sports Traumatology, Arthroscopy*, 21(4), 755-766. <https://doi.org/10.1007/s00167-012-1955-0>
30. Robert, H., Nouveau, S., Gageot, S., & Gagnière, B. (2009). A new knee arthrometer, the GNRB®: Experience in ACL complete and partial tears. *Orthopaedics and Traumatology: Surgery and Research*, 95(3), 171-176. Embase. <https://doi.org/10.1016/j.otsr.2009.03.009>
31. Rousseau, R. (2018). De l'intérêt de la laximétrie en pratique quotidienne après une rupture du ligament croisé antérieur... Faut-il rester clinique ? *Journal de Traumatologie du Sport*, 35(4), 195-196. <https://doi.org/10.1016/j.jts.2018.10.006>
32. Ryu, S. M., Na, H. D., & Shon, O. J. (2018). Diagnostic Tools for Acute Anterior Cruciate Ligament Injury : GNRB, Lachman Test, and Telos. *Knee Surgery & Related Research*, 30(2), 121-127. <https://doi.org/10.5792/ksrr.17.014>
33. Sanders, T. L., Maradit Kremers, H., Bryan, A. J., Larson, D. R., Dahm, D. L., Levy, B. A., Stuart, M. J., & Krych, A. J. (2016). Incidence of Anterior Cruciate Ligament Tears and Reconstruction : A 21-Year Population-Based Study. *The American Journal of Sports Medicine*, 44(6), 1502-1507. <https://doi.org/10.1177/0363546516629944>
34. Saravia, A., Cabrera, S., Molina, C. R., Pacheco, L., & Muñoz, G. (2020). Validity of the Genourob arthrometer in the evaluation of total thickness tears of anterior cruciate ligament. *Journal of Orthopaedics*, 22, 203-206. <https://doi.org/10.1016/j.jor.2020.03.041>
35. Sidibé, S., Deme, H., Akpo, G., Niane, M. M., Daffe, M., Diop, A. D., Soko, T. O., Kinkpe, C., Sane, A. D., Ba, S., Niang, E. H., & Sidibé, S. (2020). Performance of magnetic resonance imaging ang genourob laximetry in anterior cruciate ligament rupture in 30 cases. *Le Mali medical*, 35(2), 20-25. Medline.
36. Smith, K., Miller, N., & Laslovich, S. (2022). The Reliability of the GNRB® Knee Arthrometer in Measuring ACL Stiffness and Laxity : Implications for Clinical Use and Clinical Trial Design. *International Journal of Sports Physical Therapy*, 17(6), 1016-1025. <https://doi.org/10.26603/001c.38252>

37. Sokal, P. A., Norris, R., Maddox, T. W., & Oldershaw, R. A. (2022). The diagnostic accuracy of clinical tests for anterior cruciate ligament tears are comparable but the Lachman test has been previously overestimated : A systematic review and meta-analysis. *Knee Surgery, Sports Traumatology, Arthroscopy*, 30(10), 3287-3303. <https://doi.org/10.1007/s00167-022-06898-4>

38. Stone, A. V., Marx, S., & Conley, C. W. (2021). Management of Partial Tears of the Anterior Cruciate Ligament : A Review of the Anatomy, Diagnosis, and Treatment. *JAAOS - Journal of the American Academy of Orthopaedic Surgeons*, 29(2), 60. <https://doi.org/10.5435/JAAOS-D-20-00242>

39. Stranger, N., Kaulfersch, C., Mattiassich, G., Mandl, J., Hausbrandt, P. A., Szolar, D., Schöllnast, H., & Tillich, M. (2023). Frequency of anterolateral ligament tears and ramp lesions in patients with anterior cruciate ligament tears and associated injuries indicative for these lesions—A retrospective MRI analysis. *European Radiology*, 33(7), 4833-4841. <https://doi.org/10.1007/s00330-023-09444-z>

40. van Eck, C. F., van den Bekerom, M. P. J., Fu, F. H., Poolman, R. W., & Kerkhoffs, G. M. M. J. (2013). Methods to diagnose acute anterior cruciate ligament rupture : A meta-analysis of physical examinations with and without anaesthesia. *Knee Surgery, Sports Traumatology, Arthroscopy: Official Journal of the ESSKA*, 21(8), 1895-1903. <https://doi.org/10.1007/s00167-012-2250-9>

41. Vavken, P., & Murray, M. M. (2013). ACL Injury Epidemiology. In M. M. Murray, P. Vavken, & B. Fleming (Éds.), *The ACL Handbook : Knee Biology, Mechanics, and Treatment* (p. 3-17). Springer New York. https://doi.org/10.1007/978-1-4614-0760-7_1

42. Zantop, T., Petersen, W., & Fu, F. H. (2005). Anatomy of the anterior cruciate ligament. *Anatomic ACL Reconstruction, Part 1*, 15(1), 20-28. <https://doi.org/10.1053/j.oto.2004.11.011>

ANNEXES

Tableau 1. Critères d'inclusion et d'exclusion de la revue

CRITÈRES D'INCLUSION	CRITÈRES D'EXCLUSION
• Étude quantitative, expérimentale, comparative et/ou diagnostique • Langue française ou anglaise • Comparaison GNRB® à autre moyen de mesure • Diagnostic de RP ou RC du LCA • LCA pathologique	• Étude qualitative • Étude cadavérique • Étude post-opératoire • Genoux sains

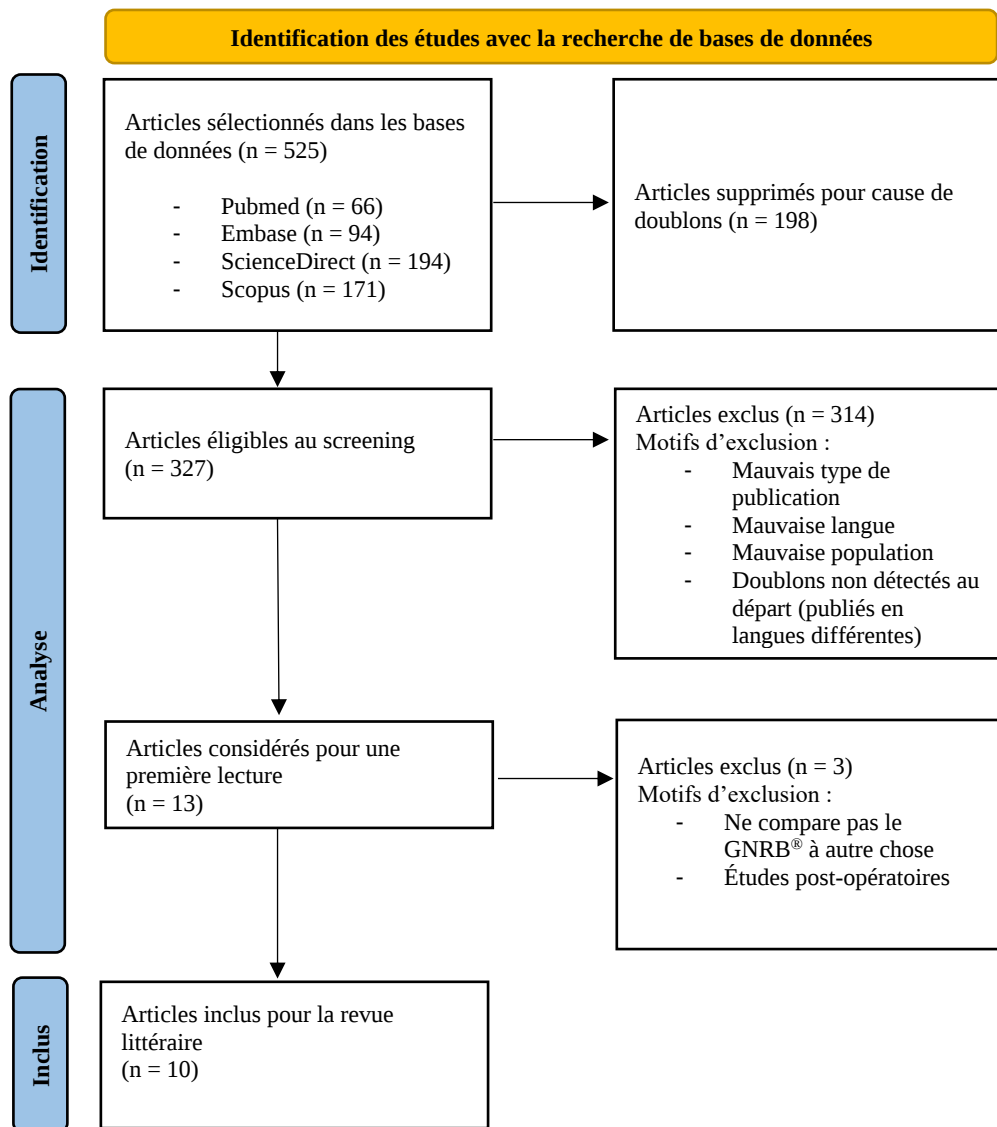


Figure 1. Diagramme de recherche selon les directives PRISMA.

Tableau 2. Risque de biais pour les études incluses. NIH Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies.

	Question NIH														Total de "oui"
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
Beldame et al., 2012	✓	✓	⚠	✓	✗	✓	✓	✓	✓	✗	✓	⚠	✓	⚠	9/14
Cojean et al., 2023	✓	✓	⚠	✓	✗	✓	✓	✗	✓	✗	✓	⚠	✓	⚠	8/14
Di Iorio et al., 2014	✓	✓	⚠	✓	✗	✓	✓	✗	✓	✗	✓	⚠	✓	⚠	8/14
Jenny et Arndt, 2013	✓	✓	⚠	✓	✗	✓	✓	✗	✓	✗	✓	⚠	⚠	⚠	7/14
Lefèvre et al., 2014	✓	✓	⚠	✓	✗	✓	✓	✓	✓	✗	✓	⚠	✓	⚠	9/14
Murgier et al., 2018	✓	✓	⚠	✓	✓	✓	✓	✗	✓	✗	✓	⚠	✓	⚠	9/14
Robert et al., 2009	✓	✓	⚠	✓	✗	✓	✓	✓	✓	✗	✓	⚠	✓	⚠	9/14
Ryu et al., 2018	✓	✓	⚠	✓	✗	✓	✓	✓	✓	✗	✓	⚠	✓	✓	10/14
Saravia et al., 2020	✓	✓	✗	✗	✗	✓	✓	✗	✓	✗	✓	⚠	✓	⚠	7/14
Sidibé et al., 2020	✓	✓	⚠	✓	✗	✓	✓	✗	✓	✗	✓	⚠	⚠	⚠	7/14

Question 8 a été comptabilisée à partir du moment où différents paramètres ont été testés. La 13 a été comptabilisée à partir du moment où il y a au moins autant de mesures pour les tests que de personnes incluses (ou < de 20% de différence).

Tableau 4. Détails des échantillons de chaque étude.

Étude	N	% hommes	% LCA sains	% RP	% RC
Beldame et al.	157	77%	50%	17%	33%
Cojean et al.	214	67%	21,50%	28,50%	50%
Di Iorio et al.	49	71%	0%	28,60%	71,40%
Jenny & Arndt	21	67%	0%	-	-
Lefèvre et al.	139	62%	0%	30%	70%
Murgier et al.	52	77%	0%	0%	100%
Robert et al.	65	78%	31%	37%	32%
Ryu et al.	40	65%	0%	0%	100%
Saravia et al.	79	-	46%	0%	54%
Sidibé et al.	30	83%	0%	-	-

N : nombre de cas ; RP : ruptures partielles ; RC : ruptures complètes

Tableau 5. Sensibilités (Se), spécificités (Sp) et Indice de Youden du GNRB® dans la détection de ruptures partielles et complètes du ligament croisé antérieur (LCA).

	Seuil de laxité	GNRB® à 134N		Indice de Youden	GNRB® à 250N		Indice de Youden
		Se	Sp		Se	Sp	
RUPTURES COMPLÈTES							
Beldame <i>et al.</i> , (2012)	1,5 mm	59,4 %	93,1 %	0,525	62,2 %	75,9 %	0,381
Cojean <i>et al.</i> , (2023)	3 mm	77,8 %	86 %	0,638	-	-	-
Di Iorio <i>et al.</i> , (2014)	3 mm	-	-	-	-	-	-
Lefèvre <i>et al.</i> , (2014)	-	-	-	-	-	-	-
Murgier <i>et al.</i> , (2018)	-	-	-	-	-	-	-
Robert <i>et al.</i> , (2009)	3 mm	70 %	99 %	0,690	-	-	-
Ryu <i>et al.</i> , (2018)	-	-	-	-	-	-	-
Saravia <i>et al.</i> , (2020)	3 mm	90,7 %	40,3 %	0,310	-	-	-
	6,8 mm	74,4 %	93,8 %	0,682	-	-	-
Sidibé <i>et al.</i> , (2020)	3 mm	-	-	-	87 %	42,9 %	0,299
RUPTURES PARTIELLES							
Beldame <i>et al.</i> , (2012)	-	-	-	-	-	-	-
Cojean <i>et al.</i> , (2023)	1,3 mm	73,8 %	85,5 %	0,593	-	-	-
Di Iorio <i>et al.</i> , (2014)	1,5 mm	72 %	-	-	-	-	-
Lefèvre <i>et al.</i> , (2014)	2 mm	83,2 %	64,3 %	0,475	-	-	-
	2,5 mm	-	-	-	84 %	81 %	0,650
Robert <i>et al.</i> , (2009)	1,5 mm	80 %	87 %	0,670	-	-	-
Sidibé <i>et al.</i> , (2020)	1,5 mm	-	-	-	42,9 %	87 %	0,299

Annexe 2. NIH Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies.

12/11/2017

Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies - NHLBI, NIH



Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies

Criteria	Yes	No	Other (CD, NR, NA)*
1. Was the research question or objective in this paper clearly stated?			
2. Was the study population clearly specified and defined?			
3. Was the participation rate of eligible persons at least 50%?			
4. Were all the subjects selected or recruited from the same or similar populations (including the same time period)? Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants?			
5. Was a sample size justification, power description, or variance and effect estimates provided?			
6. For the analyses in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured?			
7. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?			
8. For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)?			
9. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?			
10. Was the exposure(s) assessed more than once over time?			
11. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?			
12. Were the outcome assessors blinded to the exposure status of participants?			
13. Was loss to follow-up after baseline 20% or less?			
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?			

Quality Rating (Good, Fair, or Poor) (see guidance)
Rater #1 initials:
Rater #2 initials:
Additional Comments (If POOR, please state why):

*CD, cannot determine; NA, not applicable; NR, not reported

Guidance for Assessing the Quality of Observational Cohort and Cross-Sectional Studies

The guidance document below is organized by question number from the tool for quality assessment of observational cohort and cross-sectional studies.

Question 1. Research question

Did the authors describe their goal in conducting this research? Is it easy to understand what they were looking to find? This issue is important for any scientific paper of any type. Higher quality scientific research explicitly defines a research question.

Questions 2 and 3. Study population

Did the authors describe the group of people from which the study participants were selected or recruited, using demographics, location, and time period? If you were to conduct this study again, would you know who to recruit, from where, and from what time period? Is the cohort population free of the outcomes of interest at the time they were recruited?

An example would be men over 40 years old with type 2 diabetes who began seeking medical care at Phoenix Good Samaritan Hospital between January 1, 1990 and December 31, 1994. In this example, the population is clearly described as: (1) who (men over 40 years old with type 2 diabetes); (2) where (Phoenix Good Samaritan Hospital); and (3) when (between January 1, 1990 and December 31, 1994). Another example is women ages 34 to 59 years of age in 1980 who were in the nursing profession and had no known coronary disease, stroke, cancer, hypercholesterolemia, or diabetes, and were recruited from the 11 most populous States, with contact information obtained from State nursing boards.

In cohort studies, it is crucial that the population at baseline is free of the outcome of interest. For example, the nurses' population above would be an appropriate group in which to study incident coronary disease. This information is usually found either in descriptions of population recruitment, definitions of variables, or inclusion/exclusion criteria.

You may need to look at prior papers on methods in order to make the assessment for this question. Those papers are usually in the reference list.

If fewer than 50% of eligible persons participated in the study, then there is concern that the study population does not adequately represent the target population. This increases the risk of bias.

Question 4. Groups recruited from the same population and uniform eligibility criteria

Were the inclusion and exclusion criteria developed prior to recruitment or selection of the study population? Were the same underlying criteria used for all of the subjects involved? This issue is related to the description of the study population, above, and you may find the information for both of these questions in the same section of the paper.

<https://www.nhlbi.nih.gov/health-pro/guidelines/in-develop/cardiovascular-risk-reduction/tools/cohort>

1/4

Most cohort studies begin with the selection of the cohort; participants in this cohort are then measured or evaluated to determine their exposure status. However, some cohort studies may recruit or select exposed participants in a different time or place than unexposed participants, especially retrospective cohort studies—which is when data are obtained from the past (retrospectively), but the analysis examines exposures prior to outcomes. For example, one research question could be whether diabetic men with clinical depression are at higher risk for cardiovascular disease than those without clinical depression. So, diabetic men with depression might be selected from a mental health clinic, while diabetic men without depression might be selected from an internal medicine or endocrinology clinic. This study recruits groups from different clinic populations, so this example would get a "no."

However, the women nurses described in the question above were selected based on the same inclusion/exclusion criteria, so that example would get a "yes."

Question 5. Sample size justification

Did the authors present their reasons for selecting or recruiting the number of people included or analyzed? Do they note or discuss the statistical power of the study? This question is about whether or not the study had enough participants to detect an association if one truly existed.

A paragraph in the methods section of the article may explain the sample size needed to detect a hypothesized difference in outcomes. You may also find a discussion of power in the discussion section (such as the study had 85 percent power to detect a 20 percent increase in the rate of an outcome of interest, with a 2-sided alpha of 0.05). Sometimes estimates of variance and/or estimates of effect size are given, instead of sample size calculations. In any of these cases, the answer would be "yes."

However, observational cohort studies often do not report anything about power or sample sizes because the analyses are exploratory in nature. In this case, the answer would be "no." This is not a "fatal flaw." It just may indicate that attention was not paid to whether the study was sufficiently sized to answer a prespecified question—i.e., it may have been an exploratory, hypothesis-generating study.

Question 6. Exposure assessed prior to outcome measurement

This question is important because, in order to determine whether an exposure causes an outcome, the exposure must come before the outcome.

For some prospective cohort studies, the investigator enrolls the cohort and then determines the exposure status of various members of the cohort (large epidemiological studies like Framingham used this approach). However, for other cohort studies, the cohort is selected based on its exposure status, as in the example above of depressed diabetic men (the exposure being depression). Other examples include a cohort identified by its exposure to fluoridated drinking water and then compared to a cohort living in an area without fluoridated water, or a cohort of military personnel exposed to combat in the Gulf War compared to a cohort of military personnel not deployed in a combat zone.

With either of these types of cohort studies, the cohort is followed forward in time (i.e., prospectively) to assess the outcomes that occurred in the exposed members compared to nonexposed members of the cohort. Therefore, you begin the study in the present by looking at groups that were exposed (or not) to some biological or behavioral factor, intervention, etc., and then you follow them forward in time to examine outcomes. If a cohort study is conducted properly, the answer to this question should be "yes," since the exposure status of members of the cohort was determined at the beginning of the study before the outcomes occurred.

For retrospective cohort studies, the same principal applies. The difference is that, rather than identifying a cohort in the present and following them forward in time, the investigators go back in time (i.e., retrospectively) and select a cohort based on their exposure status in the past and then follow them forward to assess the outcomes that occurred in the exposed and nonexposed cohort members. Because in retrospective cohort studies the exposure and outcomes may have already occurred (it depends on how long they follow the cohort), it is important to make sure that the exposure preceded the outcome.

Sometimes cross-sectional studies are conducted (or cross-sectional analyses of cohort-study data), where the exposures and outcomes are measured during the same timeframe. As a result, cross-sectional analyses provide weaker evidence than regular cohort studies regarding a potential causal relationship between exposures and outcomes. For cross-sectional analyses, the answer to Question 6 should be "no."

Question 7. Sufficient timeframe to see an effect

Did the study allow enough time for a sufficient number of outcomes to occur or be observed, or enough time for an exposure to have a biological effect on an outcome? In the examples given above, if clinical depression has a biological effect on increasing risk for CVD, such an effect may take years. In the other example, if higher dietary sodium increases BP, a short timeframe may be sufficient to assess its association with BP, but a longer timeframe would be needed to examine its association with heart attacks.

The issue of timeframe is important to enable meaningful analysis of the relationships between exposures and outcomes to be conducted. This often requires at least several years, especially when looking at health outcomes, but it depends on the research question and outcomes being examined.

Cross-sectional analyses allow no time to see an effect, since the exposures and outcomes are assessed at the same time, so those would get a "no" response.

Question 8. Different levels of the exposure of interest

If the exposure can be defined as a range (examples: drug dosage, amount of physical activity, amount of sodium consumed), were multiple categories of that exposure assessed? (for example, for drugs: not on the medication, on a low dose, medium dose, high dose; for dietary sodium, higher than average U.S. consumption, lower than recommended consumption, between the two). Sometimes discrete categories of exposure are not used, but instead exposures are measured as continuous variables (for example, mg/day of dietary sodium or BP values).

In any case, studying different levels of exposure (where possible) enables investigators to assess trends or dose-response relationships between exposures and outcomes—e.g., the higher the exposure, the greater the rate of the health outcome. The presence of trends or dose-response relationships lends credibility to the hypothesis of causality between exposure and outcome.

For some exposures, however, this question may not be applicable (e.g., the exposure may be a dichotomous variable like living in a rural setting versus an urban setting, or vaccinated/not vaccinated with a one-time vaccine). If there are only two possible exposures (yes/no), then this question should be given an "NA," and it should not count negatively towards the quality rating.

Question 9. Exposure measures and assessment

Were the exposure measures defined in detail? Were the tools or methods used to measure exposure accurate and reliable—for example, have they been validated or are they objective? This issue is important as it influences confidence in the reported exposures. When exposures are measured with less accuracy or validity, it is

harder to see an association between exposure and outcome even if one exists. Also as important is whether the exposures were assessed in the same manner within groups and between groups; if not, bias may result.

For example, retrospective self-report of dietary salt intake is not as valid and reliable as prospectively using a standardized dietary log plus testing participants' urine for sodium content. Another example is measurement of BP, where there may be quite a difference between usual care, where clinicians measure BP however it is done in their practice setting (which can vary considerably), and use of trained BP assessors using standardized equipment (e.g., the same BP device which has been tested and calibrated) and a standardized protocol (e.g., patient is seated for 5 minutes with feet flat on the floor, BP is taken twice in each arm, and all four measurements are averaged). In each of these cases, the former would get a "no" and the latter a "yes."

Here is a final example that illustrates the point about why it is important to assess exposures consistently across all groups: If people with higher BP (exposed cohort) are seen by their providers more frequently than those without elevated BP (nonexposed group), it also increases the chances of detecting and documenting changes in health outcomes, including CVD-related events. Therefore, it may lead to the conclusion that higher BP leads to more CVD events. This may be true, but it could also be due to the fact that the subjects with higher BP were seen more often; thus, more CVD-related events were detected and documented simply because they had more encounters with the health care system. Thus, it could bias the results and lead to an erroneous conclusion.

Question 10. Repeated exposure assessment

Was the exposure for each person measured more than once during the course of the study period? Multiple measurements with the same result increase our confidence that the exposure status was correctly classified. Also, multiple measurements enable investigators to look at changes in exposure over time, for example, people who ate high dietary sodium throughout the followup period, compared to those who started out high then reduced their intake, compared to those who ate low sodium throughout. Once again, this may not be applicable in all cases. In many older studies, exposure was measured only at baseline. However, multiple exposure measurements do result in a stronger study design.

Question 11. Outcome measures

Were the outcomes defined in detail? Were the tools or methods for measuring outcomes accurate and reliable—for example, have they been validated or are they objective? This issue is important because it influences confidence in the validity of study results. Also important is whether the outcomes were assessed in the same manner within groups and between groups.

An example of an outcome measure that is objective, accurate, and reliable is death—the outcome measured with more accuracy than any other. But even with a measure as objective as death, there can be differences in the accuracy and reliability of how death was assessed by the investigators. Did they base it on an autopsy report, death certificate, death registry, or report from a family member? Another example is a study of whether dietary fat intake is related to blood cholesterol level (cholesterol level being the outcome), and the cholesterol level is measured from fasting blood samples that are all sent to the same laboratory. These examples would get a "yes." An example of a "no" would be self-report by subjects that they had a heart attack, or self-report of how much they weigh (if body weight is the outcome of interest).

Similar to the example in Question 9, results may be biased if one group (e.g., people with high BP) is seen more frequently than another group (people with normal BP) because more frequent encounters with the health care system increases the chances of outcomes being detected and documented.

Question 12. Blinding of outcome assessors

Blinding means that outcome assessors did not know whether the participant was exposed or unexposed. It is also sometimes called "masking." The objective is to look for evidence in the article that the person(s) assessing the outcome(s) for the study (for example, examining medical records to determine the outcomes that occurred in the exposed and comparison groups) is masked to the exposure status of the participant. Sometimes the person measuring the exposure is the same person conducting the outcome assessment. In this case, the outcome assessor would most likely not be blinded to exposure status because they also took measurements of exposures. If so, make a note of that in the comments section.

As you assess this criterion, think about whether it is likely that the person(s) doing the outcome assessment would know (or be able to figure out) the exposure status of the study participants. If the answer is no, then blinding is adequate. An example of adequate blinding of the outcome assessors is to create a separate committee, whose members were not involved in the care of the patient and had no information about the study participants' exposure status. The committee would then be provided with copies of participants' medical records, which had been stripped of any potential exposure information or personally identifiable information. The committee would then review the records for prespecified outcomes according to the study protocol. If blinding was not possible, which is sometimes the case, mark "NA" and explain the potential for bias.

Question 13. Followup rate

Higher overall followup rates are always better than lower followup rates, even though higher rates are expected in shorter studies, whereas lower overall followup rates are often seen in studies of longer duration. Usually, an acceptable overall followup rate is considered 80 percent or more of participants whose exposures were measured at baseline. However, this is just a general guideline. For example, a 6-month cohort study examining the relationship between dietary sodium intake and BP level may have over 90 percent followup, but a 20-year cohort study examining effects of sodium intake on stroke may have only a 65 percent followup rate.

Question 14. Statistical analyses

Were key potential confounding variables measured and adjusted for, such as by statistical adjustment for baseline differences? Logistic regression or other regression methods are often used to account for the influence of variables not of interest.

This is a key issue in cohort studies, because statistical analyses need to control for potential confounders, in contrast to an RCT, where the randomization process controls for potential confounders. All key factors that may be associated both with the exposure of interest and the outcome—that are not of interest to the research question—should be controlled for in the analyses.

For example, in a study of the relationship between cardiorespiratory fitness and CVD events (heart attacks and strokes), the study should control for age, BP, blood cholesterol, and body weight, because all of these factors are associated both with low fitness and with CVD events. Well-done cohort studies control for multiple potential confounders.

Some general guidance for determining the overall quality rating of observational cohort and cross-sectional studies

The questions on the form are designed to help you focus on the key concepts for evaluating the internal validity of a study. They are not intended to create a list that you simply tally up to arrive at a summary judgment of quality.

Internal validity for cohort studies is the extent to which the results reported in the study can truly be attributed to the exposure being evaluated and not to flaws in the design or conduct of the study—in other words, the ability of the study to draw associative conclusions about the effects of the exposures being studied on outcomes. Any such flaws can increase the risk of bias.

Critical appraisal involves considering the risk of potential for selection bias, information bias, measurement bias, or confounding (the mixture of exposures that one cannot tease out from each other). Examples of confounding include co-interventions, differences at baseline in patient characteristics, and other issues throughout the questions above. High risk of bias translates to a rating of poor quality. Low risk of bias translates to a rating of good quality. (Thus, the greater the risk of bias, the lower the quality rating of the study.)

In addition, the more attention in the study design to issues that can help determine whether there is a causal relationship between the exposure and outcome, the higher quality the study. These include exposures occurring prior to outcomes, evaluation of a dose-response gradient, accuracy of measurement of both exposure and outcome, sufficient timeframe to see an effect, and appropriate control for confounding—all concepts reflected in the tool.

Generally, when you evaluate a study, you will not see a "fatal flaw," but you will find some risk of bias. By focusing on the concepts underlying the questions in the quality assessment tool, you should ask yourself about the potential for bias in the study you are critically appraising. For any box where you check "no" you should ask, "What is the potential risk of bias resulting from this flaw in study design or execution?" That is, does this factor cause you to doubt the results that are reported in the study or doubt the ability of the study to accurately assess an association between exposure and outcome?

The best approach is to think about the questions in the tool and how each one tells you something about the potential for bias in a study. The more you familiarize yourself with the key concepts, the more comfortable you will be with critical appraisal. Examples of studies rated good, fair, and poor are useful, but each study must be assessed on its own based on the details that are reported and consideration of the concepts for minimizing bias.

Last Updated March 2014

RÉSUMÉ

Introduction

Le diagnostic de rupture du ligament croisé antérieur (LCA) repose essentiellement sur l'examen clinique et l'Imagerie par Résonance Magnétique (IRM). Il y a beaucoup de désaccord concernant le test optimal et l'IRM coûte cher, est dépendante du radiologue et a des délais souvent longs. La laximétrie pourrait permettre d'obtenir un diagnostic précis en évitant ces inconvénients.

Objectif

Évaluer, sur base de la littérature, la plus-value du GNRB® dans le diagnostic de lésion du LCA.

Méthode

Quatre bases de données ont été épluchées à la recherche d'articles comparant le GNRB® à tout autre outil diagnostique sur des sujets avec des LCA pathologiques. Une sélection d'article a été réalisée en suivant des critères d'inclusion et d'exclusion prédéfinis.

Résultats

525 références ont été récoltées dans les différentes bases de données puis triées pour ne garder que les 10 articles comparant le GNRB® à différents outils sur des sujets pathologiques ; le Telos®, l'IRM, le KT-1000®, l'examen clinique et le Rolimeter®.

Conclusion

La majorité des études soutient que le GNRB® possède une valeur diagnostique avérée. Celle-ci surpasse fréquemment celle des outils auxquelles il a été comparé. Malgré cela, le GNRB® semble ne pas être en mesure de remplacer l'IRM mais leur utilisation combinée pourrait augmenter l'exactitude du diagnostic.

Design

Revue narrative systématisée.