

I Astuces pour l'analyse critique d'article scientifique

Pièges et astuces pour analyser un article d'imagerie cardiovasculaire multimodale



T. PEZEL
Service de Cardiologie,
CHU Lariboisière, PARIS;
Unité Inserm-UMR 942, PARIS.



S. TOUPIN
Siemens Healthineers IHU
LIRYC (Institut de Rythmologie
et de Modélisation cardiaque),
BORDEAUX.

L'imagerie cardiovasculaire multimodale connaît une période d'engouement sans précédent avec des développements techniques majeurs proposés par les constructeurs. Cet attrait pour l'imagerie multimodale est d'ailleurs bien présent auprès des jeunes internes de cardiologie. En effet, d'après un sondage récent réalisé par le Collège des cardiologues en formation (CCF) de la Société Française de Cardiologie, l'imagerie cardiovasculaire multimodale serait la première surspécialité (FST – option) choisie par les internes devant la rythmologie et la cardiologie interventionnelle...

Ainsi, les articles d'imagerie cardiovasculaire sont de plus en plus présents dans la littérature et lors des congrès nationaux et internationaux. Cependant, cet univers qui navigue entre la cardiologie et la radiologie présente ses propres codes et sa méthodologie spécifique, avec de véritables astuces à connaître pour faire la part des choses entre "véritable avancée diagnostique pour le patient" et "effet de style visuel".

Nous vous proposons d'aborder ensemble les grands pièges à vérifier fréquemment dans les articles d'imagerie cardiovasculaire.

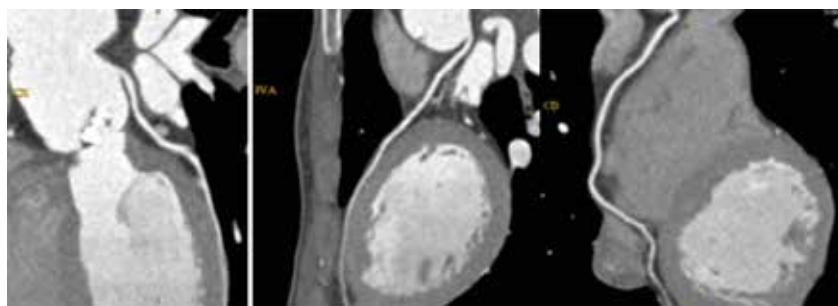
Rester vigilant devant l'évaluation d'une nouvelle technique sur volontaires sains

Lorsqu'on souhaite évaluer une nouvelle méthode d'imagerie, comme par exemple une nouvelle séquence IRM, 3 étapes chronologiques sont possibles : études sur modèle animal, études sur volontaires sains et études sur patients.

Ainsi, les études cliniques sur volontaires sains sont fréquentes pour valider une nouvelle méthode diagnostique, mais nous devons rester très vigilants

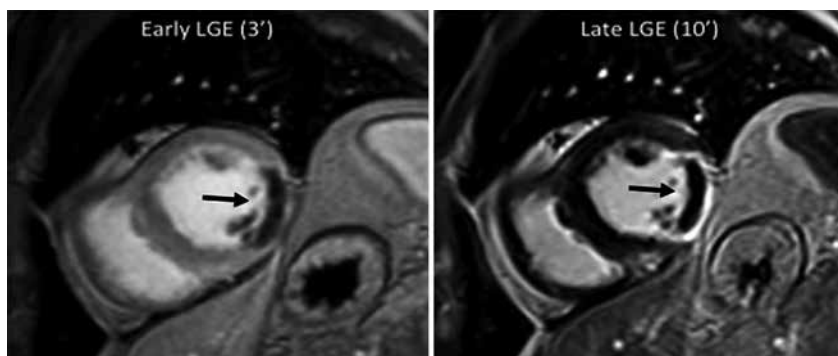
dans l'analyse de ce type d'étude. En effet, contrairement à un recrutement systématique et consécutif de patients au sein d'un service qui ne permet pas à l'investigateur de sélectionner ces patients, le recrutement de volontaires sains permet une grande liberté. Ainsi, l'investigateur pourra choisir de présenter dans son article les "meilleurs volontaires sains", autrement dit ceux donnant les meilleures images.

De plus, les conditions d'examen des volontaires sains sont beaucoup plus faciles à contrôler que celles des patients malades (**encadré 1**).



- Pour l'acquisition d'images coronaires de qualité, le patient doit idéalement être :
 - capable de **retenir sa respiration** suffisamment longtemps (évite les artefacts respiratoires);
 - avec un **rythme cardiaque le plus régulier possible et non tachycarde** (évite les problèmes de synchronisation ECG qui peuvent entraîner des artefacts de mouvement);
 - plutôt **mince** (donne de meilleures images que chez un patient obèse).
- Ainsi, il semble évident qu'il sera plus facile de remplir toutes ces conditions d'exams chez un **volontaire sain mince**, de 25 ans, capable de maintenir parfaitement l'apnée. L'acquisition réalisée dans des conditions optimales permet d'obtenir des images de qualité.
- *A contrario*, la qualité d'acquisition des images ne sera pas la même chez un **patient** âgé de 80 ans, avec une BPCO limitant sa capacité d'apnée, en surpoids et avec un antécédent de fibrillation auriculaire.
- **En résumé** : les études d'imagerie réalisées sur volontaires sains sont absolument capitales comme "preuve de concept", mais elles doivent absolument être validées ensuite chez des patients.

Encadré 1 : Conditions d'examen des volontaires sains beaucoup plus faciles à contrôler : exemple d'une étude évaluant une nouvelle modalité de coroscaner sur volontaires sains (iconographie d'illustration du service de radiologie du CHU Lariboisière, Paris).



- Dans une étude qui évaluerait une nouvelle séquence IRM pour diagnostiquer l'infarctus du myocarde avec *no-reflow*, le **gold standard** serait la séquence de rehaussement tardif (*LGE – late gadolinium enhancement*), qui doit être réalisée 10 minutes après l'injection de gadolinium (*late LGE*).
- Or, si les auteurs utilisent comme **gold standard** une séquence de rehaussement tardif réalisée 3 minutes après l'injection de gadolinium (*early LGE*), permettant de diminuer la durée d'examen, l'étude n'est pas valide et l'évaluation de la nouvelle séquence s'accompagne d'un risque de **biais de classement de mesure**.

Encadré 2 : S'assurer que le test de référence est bien réalisé selon les recommandations pour être le **gold standard** : exemple d'une étude évaluant une nouvelle séquence IRM pour diagnostiquer l'infarctus du myocarde avec *no-reflow* (iconographie d'illustration pour montrer l'importance du temps d'acquisition d'après une review de Bouleti et al. ACVD 2015).

Vérifier scrupuleusement le choix du test diagnostique de référence (**gold standard**)

Il est important de rappeler que l'enjeu d'une étude diagnostique est de montrer que la nouvelle technique est la plus performante possible, comparée à une technique de référence (*gold standard*). Cependant, il faut s'assurer qu'il s'agit bien de la technique de référence à utiliser dans ce contexte pour être sûr du résultat énoncé. En pratique, la plupart des techniques sont d'abord validées sur modèle animal afin de prendre l'anatomopathologie comme référence. Puis ensuite, pour les études cliniques, on utilise des techniques de référence validées, mais encore faut-il s'assurer que ce *gold standard* a été bien choisi et réalisé (**encadré 2**).

Vérifier le mode d'évaluation des examens diagnostiques

Dans les articles évaluant des tests diagnostiques d'imagerie, il est capital que les auteurs détaillent dans la partie "Méthode" comment l'examen d'imagerie a été interprété et surtout par qui ! Tout d'abord, nous devons vérifier le nombre d'opérateurs ainsi que le niveau d'expérience qui doit être précisé avec : nombre d'années d'expérience, présence du niveau de diplôme en imagerie cardiaque (français, européens – *EACVI certification, levels I-III* – ou américains – *SCMR/SCCT certification levels I-III*).

Ensuite, le *process* d'évaluation entre les opérateurs doit être détaillé : étude réalisée en aveugle ou non, quelle marche à suivre en cas de discordance entre 2 opérateurs (présence d'un 3^e pour trancher?).

Attention, **le choix des opérateurs est très important** dans ce type d'étude car il entraîne des conséquences directes sur la portée des résultats dans la pratique quotidienne :

– en effet, si la nouvelle technique a été mesurée uniquement par de grands

Astuces pour l'analyse critique d'article scientifique

experts avec une expérience importante : la validité interne de l'étude (fiabilité des résultats) sera bonne, mais l'étude souffrira d'une diminution de la validité externe (applicabilité des résultats à la pratique quotidienne), car dans la vraie vie tous les opérateurs ne sont pas des experts ;

– à l'inverse, si la nouvelle technique est évaluée par des experts de différents niveaux, la validité externe sera bonne car cela correspond à la situation rencontrée tous les jours. Cependant, on risque d'avoir des résultats moins fiables avec une diminution de la reproductibilité inter-observateur.

Enfin, une mesure de la reproductibilité inter-observateur et intra-observateur permet une évaluation précise de la validité interne de l'étude, autrement dit à quel point les résultats présentés sont fiables.

Savoir vérifier la prévalence de la maladie dans l'étude

La **prévalence de la maladie** étudiée dans l'étude correspond au nombre de patients diagnostiqués malades par le *gold standard* chez tous les patients évalués (exemple : prévalence de 10 % de myocardite devant tous les patients se présentant pour douleur thoracique et troponine élevée aux urgences).

Cette valeur présente dans la partie "Résultats" doit absolument être précisée par les auteurs afin de la comparer à la prévalence connue de cette maladie dans la population étudiée, qui sera précisée dans l'introduction ou la partie "Discussion" de l'article. En effet, nous

devons absolument retenir que cette prévalence de la maladie dans l'étude va directement influencer la mesure des valeurs prédictives positive (VPP) et négative (VPN) du nouveau test.

Ainsi, **si la prévalence de la maladie dans l'étude est très différente de la prévalence connue dans la littérature, nous ne sommes plus autorisés à conclure sur les VPP et VPN** qui peuvent être artificiellement améliorées ! Dans cette situation, on ne peut plus conclure que sur les paramètres suivants : sensibilité, spécificité, rapports de vraisemblance positif (RVP) et négatif (RVN).

Vérifier le délai de réalisation de l'examen par rapport à la pathologie étudiée

L'imagerie cardiaque, et tout particulièrement l'IRM cardiaque, a souvent pour objectif de réaliser une caractérisation tissulaire du myocarde afin de distinguer un infarctus du myocarde d'une myocardite, d'un takotsubo ou d'une amylose cardiaque...

Cependant, il est important de toujours raisonner sur **la chronologie physiopathologique de la pathologie étudiée**.

Prenons l'exemple de l'œdème myocardique que l'on retrouve à la fois dans l'infarctus du myocarde, la myocardite ou le takotsubo. Il est important de savoir que **l'œdème myocardique n'est visible au maximum qu'un mois après le début des symptômes**. Ainsi, une étude qui évaluerait un nouveau test diagnostique pour le diagnostic de myocardite, réalisé 2 à 8 semaines après l'épisode cli-

nique, risquerait fortement de manquer le diagnostic d'œdème myocardique de la myocardite pour tous les patients évalués après 1 mois. Ce type de biais de mesure est fréquent, avec un risque de sous-estimer la performance du nouveau test diagnostique.

Conclusion

La littérature des études d'imagerie cardiovasculaire multimodale connaît une progression marquée avec des *impact factors* de journaux d'imagerie cardiaque en augmentation forte et constante ces dernières années... Cependant, il existe de nombreux biais et spécificités méthodologiques de ces études, qui doivent être connus afin de faire le tri dans cette littérature très visuelle parfois plus graphique que scientifique.

POUR EN SAVOIR PLUS

- COHEN JF, KOREVAAR DA, ALTMAN DG *et al.* STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open*, 2016;6:e012799.
- WHITING PF, RUTJES AWS, WESTWOOD ME *et al.* QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*, 2011;155:529-536.
- BOSSUYT PM, REITSMA JB, BRUNS DE *et al.* STARD 2015: An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies. *Radiology*, 2015;277:826-832.
- BOULETI C, MEWTON N, GERMAIN S. The no-reflow phenomenon: State of the art. *Arch Cardiovasc Dis*, 2015;108:661-674.
- PEZEL T. *Réussite à la lecture critique d'article scientifique*. Paris, ed. Estem-Vuibert. 2018 ; 4^e édition.