

Des tableaux de contingence 3 x 2 pour les tests diagnostiques ?

Tom Poelman, Vakgroep Huisartsgeneeskunde en Eerstelijnsgezondheidszorg, UGent

Texte sous la responsabilité de la rédaction néerlandophone

Lorsque le médecin prend une décision médicale, il se base souvent sur les informations apportées par les tests diagnostiques. La recherche sur la précision des tests diagnostiques n'est pas moins indispensable que les études menées sur l'efficacité des traitements. Les études comparant un test à évaluer (test index) et un test dont la précision a été suffisamment démontrée (test de référence) constituent une étape importante à cet égard. Les résultats d'une étude diagnostique sont habituellement présentés sous forme de tableau de contingence 2 x 2 qui permet de calculer la sensibilité, la spécificité, les valeurs prédictives d'un test positif et négatif, les rapports de vraisemblance positif et négatif, les forces probante et excluante.

Tout comme les études portant sur l'efficacité des traitements, les études diagnostiques sont également sujettes à diverses formes de biais^{1,2}. Les questionnaires tant **STARD** (Standards for the Reporting of Diagnostic Accuracy) que **QUADAS** (Quality Assessment of Diagnostic Accuracy Studies)³ qui évaluent respectivement le rapport et la qualité des études diagnostiques attirent l'attention sur différents types de biais dans la sélection des patients et dans les modalités et le moment de la réalisation du test index et du test de référence.

Une méta-analyse récente⁴ compare la précision diagnostique de l'angiographie coronaire non invasive par CT scan (test index) et celle de la coronarographie conventionnelle (test de référence). Les investigateurs ont été confrontés à des résultats non évaluables dans 109 des 120 études. Il s'agissait le plus souvent de rétrécissements vasculaires à l'angiographie par CT scan dont les investigateurs n'étaient pas en mesure de déterminer la pertinence clinique pour le patient (sténose de 50% ou plus). De ce fait, il était impossible de dire si le test index était clairement positif ou négatif. Tant STARD que QUADAS questionnent le rapport des résultats « ininterprétables » et « indéterminés » du test index, pour 24 des 109 études. Il n'est pas clair que les résultats non évaluables ont été traduits au niveau du patient et de quelle manière ils l'ont été. Le questionnaire STARD se penche non seulement sur le rapport mais aussi sur la manière dont les résultats non évaluables sont traités.

		Test de référence	
		Positif	Négatif
Test index	Positif	Vrai positif	Faux positif
	Non évaluable	Non évaluable avec test de référence positif	Non évaluable avec test de référence négatif
	Négatif	Faux négatif	Vrai négatif

Source: Schuetz GM, Schlattmann P, Dewey M. Use of 3x2 tables with an intention to diagnose approach to assess clinical performance of diagnostic tests: meta-analytical evaluation of coronary CT angiography studies. *BMJ* 2012;345:e6717.

La synthèse méthodique montre que 26 études n'ont pas tenu compte des vaisseaux sanguins non évaluables, que 23 études ont exclu les patients présentant des résultats non évaluables, que 26 études les ont tous considérés comme positifs et que 7 études les ont tous considérés comme négatifs. D'un point de vue clinique, il vaut évidemment mieux considérer les résultats

non évaluables comme positifs et procéder à des examens complémentaires. Mais avec cette approche, la précision diagnostique du test index sera faussement augmentée. Seules 3 études ont utilisé un tableau de contingence 3 x 2⁵ (voir tableau) pour séparer les résultats non évaluables. Pour 23 études, les auteurs de la synthèse méthodique ont pu établir eux-mêmes un tableau de contingence 3 x 2 parce que le résultat de la coronarographie invasive (test de référence) était également connu pour les résultats non évaluables. Cette analyse montre que la précision diagnostique est significativement plus faible ($p < 0,05$) dans les tableaux de contingence 3 x 2 que dans les tableaux de contingence 2 x 2. Le calcul de la sensibilité et de la spécificité dans un tableau de contingence 3 x 2 permet d'effectuer une « analyse en intention de diagnostiquer », par analogie avec une analyse en intention de traiter dans des études d'intervention. Lorsque, dans le tableau de contingence 3 x 2 tous les résultats non évaluables sont considérés comme positifs, le nombre de vrais positifs et le nombre de faux positifs augmentent, avec pour conséquence une augmentation de la sensibilité et une diminution de la spécificité. A l'inverse, si tous les résultats non évaluables sont considérés comme négatifs, le nombre de vrais négatifs et le nombre de faux négatifs augmentent, avec pour conséquence une diminution de la sensibilité et une augmentation de la spécificité. Les auteurs de la méta-analyse sont donc arrivés à une sensibilité qui varie de 98,3 (IC à 95% de 96,9 à 99,0) à 92,9 (IC à 95% de 88,8 à 95,5) et à une spécificité qui varie de 78,4 (IC à 95% de 71,6 à 84,0) à 90,5 (IC à 95% de 86,8 à 93,2). Pour avoir une image réelle (sans surestimation) de la précision diagnostique, il faut donc considérer les résultats non évaluables comme de faux négatifs lorsque le test de référence est positif et comme de faux positifs lorsqu'il est négatif. L'importance clinique de l'analyse en intention de diagnostiquer avec un tableau de contingence 3 x 2, par comparaison à une analyse avec un tableau de contingence 2 x 2, est bien illustrée lors du calcul du rapport de vraisemblance négatif. Il est respectivement de 0,09 (IC à 95% de 0,06 à 0,15) et de 0,02 (IC à 95% de 0,01 à 0,04). La force excluante chute de 50 à 11 en analyse en intention de diagnostiquer avec une table 3 x 2. Il s'avère donc, compte tenu des résultats non évaluables rencontrés en pratique, que l'angiographie par CT scan permettra moins d'exclure une pathologie coronaire.

Dans les analyses de Minerva portant sur les études diagnostiques, nous n'avons que rarement⁶ tenu compte de la manière dont les investigateurs avaient traité les résultats non évaluables. Or il est important de vérifier si les résultats non évaluables sont correctement traités, et certainement pour les tests diagnostiques, les personnes effectuant le test index devant avoir une certaine expertise pour pouvoir l'interpréter correctement.

Conclusion

La transparence dans le rapport des résultats d'une étude diagnostique au moyen d'un tableau de contingence 3 x 2 avec inclusion des résultats non évaluables et la réalisation d'une analyse en intention de diagnostiquer permettent d'avoir une idée plus correcte de la précision d'un test diagnostique.

Références voir site web www.minerva-ebm.be