

Epidémiologie → Analyser un tableau 2 x 2

B. Branger, 8 mai 2017

- Tableau type

	Colonne 1	Colonne 0	Total
Ligne 1	a	b	a + b = L1
Ligne 0	c	d	c + d = L0
Total	a + c = C1	b + d = C0	Population totale

- Etudes épidémiologiques

→ Tabac +/- (exposition), et cancer poumon +/- (maladie)

* Etude de cohorte

	Malades M+	Non-malades M-	Total
Exposition +	a	b	a + b = E1
Exposition -	c	d	c + d = E0
Total	a + c = M1	b + d = M0	Population totale



On suit une population avec des fumeurs et des non-fumeurs, et on établit un diagnostic de cancers du poumon sur l'évolution

* Lire en ligne

- ✓ Proportion de malades chez les exposés (cancers chez les fumeurs)
 $= P1 = P_{M+}/Expo + = a / a + b = a / E1$
- ✓ Proportion de malades chez les non-exposés (cancers chez les non-fumeurs) = $p0$
 $= P_{M+}/Expo - = c / c + d = c / E0$
- ✓ Test du χ^2 ou test de Fisher. Limite $\chi^2 > 3.84$ pour $p < 0.05$
- ✓ **Différence absolue de "risques"** : différence entre les deux proportions $\Delta AR = P1 - P0$
- ✓ **Différence relative de "risques"** : différence entre les deux proportions / proportion de référence chez les non-fumeurs $\Delta RR = (P1 - P0) / P0$
- ✓ **Risque relatif (RR)** = rapport des deux proportions = $(a/a+b)/(c/c+d) = P1/P0$. Si $RR > 1$: l'exposition est associée à la maladie. Si $RR < 1$: l'absence d'exposition est associée à la maladie. Si pas de différence : $RR = 1$. $RR \neq 1$ ne veut pas "causalité", mais lien ou association. Un RR est accompagné d'un intervalle de confiance (calcul non présenté ici) : si l'intervalle ne comprend pas 1 (à la borne inférieure ou supérieure), RR est dit significativement différent de 1. Résultats similaires au test du χ^2 .
- ✓ Si l'exposition est un médicament: **Nombre de sujets à traiter** (NST ou NTT pour *needed to treat*) : inverse de la différence absolue des proportions ($1/\Delta AR$). Si l'exposition est un médicament, plus le nombre de sujets est faible, plus le médicament est efficace.

* Etude cas-témoins



	Cas Malades M+	Témoins Non-malades M-	Total
Exposition +	a	b	a + b = E1
Exposition -	c	d	c + d = E0
Total	a + c = M1	b + d = M0	Population totale

Les nombres de cas et de de témoins sont "contrôlés" : les totaux des lignes n'ont pas de sens. Les malades sont les cancers du poumon, les témoins n'ont pas de cancer de poumons (et leur sélection peut être problématique).

* Lire en colonne

- ✓ Proportions d'exposés chez les cas : $P_c = P \text{ Expo } + / M+$
- ✓ Proportions d'exposés chez les témoins : $P_t = P \text{ Expo } + / M-$
- ✓ Test du χ^2 ou test de Fisher : Limite $\chi^2 > 3.84$ pour $p < 0.05$
- ✓ Différence absolue de risques : différence entre les deux proportions $\Delta AR = P_c - P_t$
- ✓ Différence relative de risques : différence entre les deux proportions / proportion de référence chez les non-fumeurs $\Delta RR = (P_c - P_t) / P_t$
- ✓ Odds ratio (OR) ou rapport des cotes (RC) = rapport des deux odds d'exposition (a/c et b/d) = $(a*d)/(b*c)$. L'OR n'est pas le rapport des expositions chez les cas et les témoins (ce serait $(a/a+c) / (b/b+d)$). Si OR = 1 : pas de lien. Si OR > 1 : l'exposition est associée à la maladie. Si OR < 1 : l'absence d'exposition est associée à la maladie. OR \neq 1 ne veut pas "causalité", mais lien ou association. Un OR est accompagné d'un intervalle de confiance (calcul non présenté ici) : si l'intervalle ne comprend pas 1 (à la borne inférieure ou supérieure), OR est dit significativement différent de 1. Résultats similaires au test du χ^2 .
- ✓ Si l'exposition est un médicament: Nombre de sujets à traiter (NST ou NTT pour *needed to treat*) : inverse de la différence absolue des proportions d'exposition ($1/\Delta AR$). Si l'exposition est un médicament, plus le nombre de sujets est faible, plus le médicament est efficace.

- Etude transversale

	Malades	Non-malades	Total
Exposition +	a	b	a + b = E1
Exposition -	c	d	c + d = E0
Total	a + c = M1	b + d = M0	Population totale

Exemple : enquête d'un jour des infections nosocomiales (maladie) dans un hôpital; exposition à la sonde urinaire (exposition). Ce n'est pas une enquête de cohorte au sens strict.

Lire selon l'objectif ou la compréhension en colonnes ou en ligne.

- ✓ P M+/Expo + versus P M+/Expo -, OU P Expo +/M+ versus P Expo +/M-
- ✓ Test du χ^2 ou de Fisher
- ✓ **Différence absolue de risques** : différence entre les deux proportions
- ✓ **Différence relative de risque** : différence entre les deux proportions / proportion de référence
- ✓ **Ratio de prévalence (RP)** = rapport des deux proportions = $(a/a+b)/(c/c+d) = P1/P0$. Si $RP > 1$: l'exposition est associée à la maladie. Si $RP < 1$: l'absence d'exposition est associée à la maladie. RR ne veut pas "causalité", mais lien ou association. Un RP est accompagné d'un intervalle de confiance (calcul non présenté ici) : si l'intervalle ne comprend pas 1 (à la borne inférieure ou supérieure, RR est dit significativement différent de 1.

- Analyse de séries appariées

Exemple : dans une maladie chronique, on traite les mêmes sujets avec deux médicaments différents à deux périodes différentes. On examine les paires de résultats pour les deux périodes : soit les deux médicaments améliorent les malades, soit ils n'en améliorent aucun, soit l'un améliore et pas l'autre (et vice versa). Il y a donc quatre paires différentes.

		Médicament 2		
		Améliorés	Non-améliorés	Total
Médicament 1	Améliorés	a	b	a + b
	Non-améliorés	c	d	c + d
Total		a + c	b + d	N(nombre de paires)

- ✓ Proportion de malades améliorés par le médicament 1 : $P1 = (a+b) / N$
- ✓ Proportion de malades améliorés par le médicament 2 : $P1 = (a+c) / N$
- ✓ Test de différence entre les deux proportions de Mac Nemar :
 $\chi^2 \text{ McN} = (b-c)^2 / (b+c)$
- ✓ Odds ration des paires : $OR \text{ app} = c / b$
- ✓ Plus de calculs dans d'autres cours.

- Test diagnostic

	Malades	Non-malades	Total
Test +	a	b	a + b = T1
Test -	c	d	c + d = T0
Total	a + c = M1	b + d = M0	Population totale

- ✓ Sensibilité : proportion de malades ayant le signe : $Se = a/a+c = a / M1$
- ✓ Spécificité : propor. de non-malades n'ayant pas le signe : $Sp = d/b+d = d / M0$
- ✓ Vrais positifs : VP = a. Vrais négatifs : VN = d
- ✓ Faux positifs : FP = b. Faux négatifs = c
- ✓ Valeur prédictive positive : VPP = $a/(a+b) = a / T1$
- ✓ Valeur prédictive négative : VPN = $d/(c+d) = d/T0$
- ✓ Indice de Youden $Y = Se + Sp - 1$
- ✓ Rapport de vraisemblance (si les malades ont plus de test + : RV positif) (*positive likelihood ratio*) : $RV = \text{proportion de VP} / \text{proportion de FP} = ((a/M1) / (b / M 0)) = Se / (1-Sp)$

Les dépistages privilégient la spécificité (diminuer les faux positifs pour ne pas investiguer des non-malades, même si on "passe" à côté de quelques malades), tandis que les diagnostics privilégient la sensibilité (diminuer les faux négatifs, pour ne pas passer à côté du diagnostic, quitte à "embêter " quelques non-malades avec des examens qui seront inutiles.

- Concordance

		Observateur 1		
		Suspect	Normal	Total
Observateur 2	Suspect	a	b	a + b
	Normal	c	d	c + d
Total		a + b	b + d	Population totale (N)

Deux observateurs analysent les mêmes clichés radiologiques et classent les images en suspectes ou normales.

- ✓ Proportion observée de concordance : diagonales a et d, soit $Po = (a+d)/N$
- ✓ Proportion de concordance sous l'hypothèse d'absence de différence : Pc (calcul disponible)
- ✓ Score de concordance : Kappa de Cohen : $\kappa = (Po - Pc)/(1-Pc)$
- ✓ κ varie de 0 à + 1 avec les conclusions possibles
 - < 0 Désaccord
 - 0.00 — 0.20 : Accord très faible
 - 0.21 — 0.40 : Accord faible
 - 0.41 — 0.60 : Accord modéré
 - 0.61 — 0.80 : Accord fort
 - 0.81 — 1.00 : Accord presque parfait

Exemples numériques

- Cohorte (EpiData)

. tables primi césarienne /t /sa /r /rr

Outcome:primi						
Césarienne	1	%	2	%	Total	%
1	117	(77.0)	35	(23.0)	152	(100.0)
2	342	(44.5)	427	(55.5)	769	(100.0)
Total	459	(49.8)	462	(50.2)	921	

Percents: (Row)

Exposure: césarienne = 1

Outcome: primi = 1

Chi²= 53.623 df(1) ; p= 0.0000

RR = 1.73 (95% CI: 1.54-1.95)

tables age30 césarienne /t /sa /r /rr

Outcome:age30						
Césarienne	1	%	2	%	Total	%
1	73	(47.7)	80	(52.3)	153	(100.0)
2	371	(47.4)	411	(52.6)	782	(100.0)
Total	444	(47.5)	491	(52.5)	935	

Percents: (Row)

Exposure: césarienne = 1

Outcome: age30 = 1

Chi²= 0.004 df(1) p= 0.9512

RR = 1.01 (95% CI: 0.84-1.21)

- Cas -témoins (EpiData)

tables primi cesar /o /sa /t /c

Outcome:primi						
cesar	1	%	2	%	Total	%
1	117	{67.2}	35	{27.3}	152	{50.3}
2	57	{32.8}	93	{72.7}	150	{49.7}
Total	174	{100.0}	128	{100.0}	302	

Percents: {Col}

Exposure: cesar = 1

Outcome: primi = 1

Chi²= 46.960 df(1) p= 0.0000

Odds Ratio = 5.45 (95% CI: 3.30-9.00) (Robins,Greenland,Breslow CI)

. tables bishop5 cesar /o /sa /t /c

Outcome:bishop5						
cesar	1	%	2	%	Total	%
1	108	{65.1}	45	{32.8}	153	{50.5}
2	58	{34.9}	92	{67.2}	150	{49.5}
Total	166	{100.0}	137	{100.0}	303	

Percents: {Col}

Exposure: cesar = 1

Outcome: bishop5 = 1

Chi²= 31.158 df(1) p= 0.0000

Odds Ratio = 3.81 (95% CI: 2.36-6.14) (Robins,Greenland,Breslow CI)

- Test diagnostic (OpenEpi)

Analyse de tableau simple

	Malade	Non-malade	Total
Positif	50	10	60
Négatif	5	100	105
	55	110	165

Paramètre	Estimation	Inférieur – Supérieur ICs à 95%	Méthode
Sensibilité	90.91%	(80.42, 96.05 ¹)	
Spécificité	90.91%	(84.07, 94.99 ¹)	
Valeur prédictive positive	83.33%	(71.97, 90.69 ¹)	
Valeur prédictive négative	95.24%	(89.33, 97.95 ¹)	
Exactitude du diagnostic	90.91%	(85.54, 94.41 ¹)	
Rapport de vraisemblance du test positif	10	(8.188 - 12.21)	
Rapport de vraisemblance du test négatif	0.1	(0.06744 - 0.1483)	

Résultats tirés de OpenEpi, version 3, logiciel libre de calcul --DiagnosticTest

- Concordance (MedCalc)

Inter-rater agreement ? X

Test data

Observer A:

Observer B:	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
Class 1	20	2				
Class 2	4	30				
Class 3						
Class 4						
Class 5						
Class 6						

Options

Weighted Kappa Linear weights
 Quadratic weights

Results

Kappa	0,779
Standard error	0,0850
95% CI	0,612 to 0,946

MedCalc 2017

- Autre exemple (SPSS 22.0)

(Proportions en colonnes)

Tableau croisé expo0 * maladie

		maladie		Total	
		0	1		
expo0	0	Effectif	18	6	24
		% dans maladie	72,0%	24,0%	48,0%
1		Effectif	7	19	26
		% dans maladie	28,0%	76,0%	52,0%
Total		Effectif	25	25	50
		% dans maladie	100,0%	100,0%	100,0%

Tests du khi-deux

	Valeur	ddl	Sig. approx. (bilatérale)	Sig. exacte (bilatérale)	Sig. exacte (unilatérale)
khi-deux de Pearson	11,538 ^a	1	,001		
Correction pour continuité ^b	9,696	1	,002		
Rapport de vraisemblance	12,033	1	,001		
Test exact de Fisher				,002	,001
Test de McNemar				1,000 ^c	
N d'observations valides	50				

a. 0 cellules (0,0%) ont un effectif théorique inférieur à 5. L'effectif théorique minimum est de 12,00.

b. Calculée uniquement pour une table 2x2

c. Distribution binomiale utilisée.

		Valeur	Erreur standard asymptotique	T approx. ^b	Signification approx.
Intervalle par	R de Pearson	,480	,124	,000 ^c	,000 ^c
Intervalle					
Ordinal par Ordinal	Corrélation de Spearman	,480	,124	,000 ^c	,000 ^c
Mesure d'accord	Kappa	,480	,124	,001	,001
N d'observations valides		50			

a. L'hypothèse nulle n'étant pas considérée.

b. Utilisation de l'erreur asymptotique standard en envisageant l'hypothèse nulle.

c. Basé sur une approximation normale.

Estimation du risque

	Valeur	Intervalle de confiance à 95%	
		Plus bas	Supérieur
Rapport des cotes pour expo0 (0 / 1)	8,143	2,294	28,901
Pour cohorte maladie = 0	2,786	1,420	5,466
Pour cohorte maladie = 1	,342	,165	,711
N d'observations valides	50		

**NB SPSS : "Rapport des cotes" = Odds ratio (voir le sens de calcul)
 "Pour cohorte " : Risque relatif (et son inverse)**

=====