



Recherche de nouvelles approches de prévention pour le cancer du col utérin: Guide méthodologique

Source:

Alliance pour la Prévention du Cancer du Col Utérin (ACCP)

www.alliance-cxca.org

Les objectifs d'un programme de prévention:

- Réduire l'incidence et la mortalité liées au cancer du col.

Conditions requises:

- Détection de lésions précancéreuses du col (maladie précoce) à l'aide d'un test approprié.
- Traitement de la maladie précoce pour éviter la progression de la maladie.

Test approprié pour déceler un précancer:

- Faible coût
- Sans danger
- Pratique et acceptable
- Fiable
- Performant

Fiabilité du test :

*Capacité du test à donner plusieurs fois des résultats identiques
(que le testeur soit la même personne ou non).*

- La fiabilité dépend de:
 - La manifestation clinique de la maladie précoce
 - La méthode de mesure (subjectivité du test)
 - Les compétences de la personne effectuant le test
 - Le nombre d'étapes pour la réalisation du test
- Faible fiabilité = faible reproductibilité

Les caractéristiques du test:

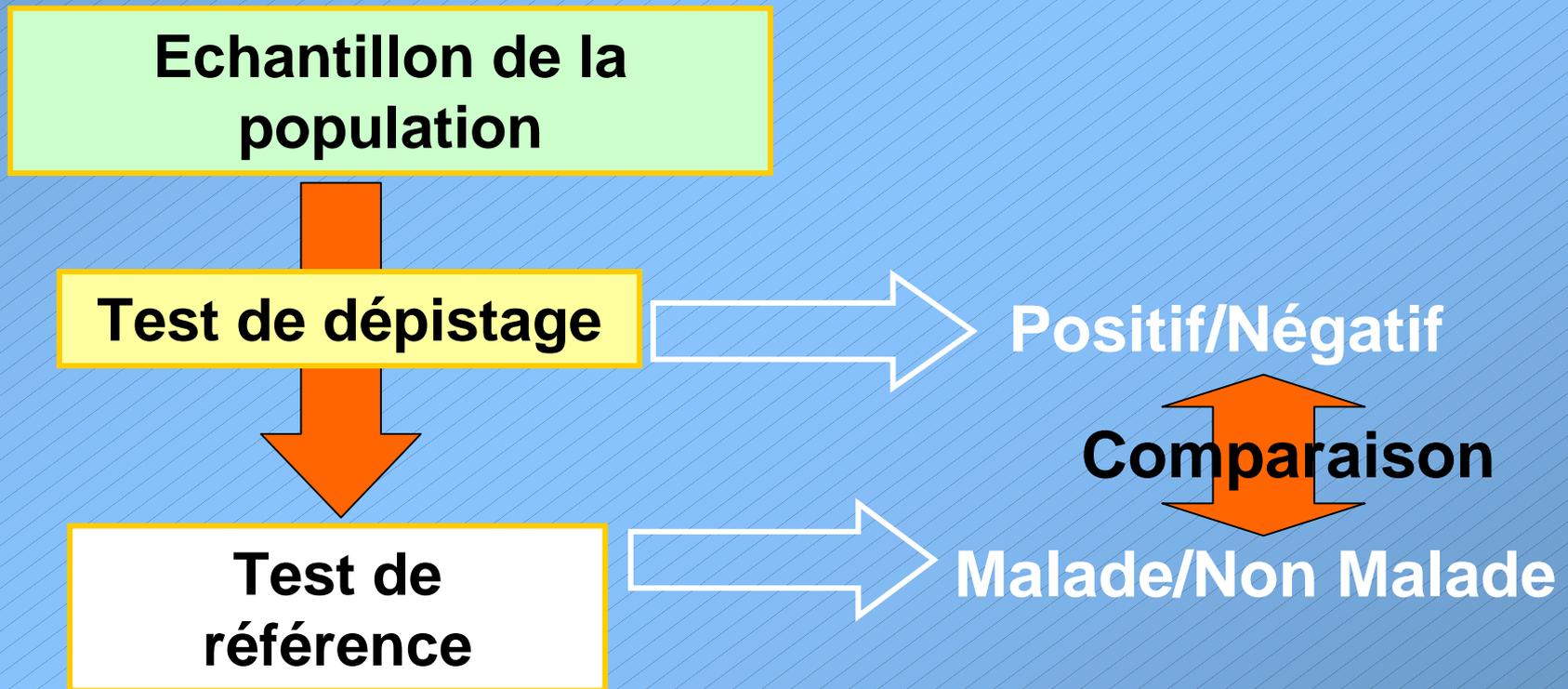
- **Précision: caractéristiques intrinsèques**
 - **Sensibilité (Se):** proportion des femmes ayant réellement la maladie qui ont un test positif.
 - **Spécificité (Sp):** proportion des femmes n'ayant pas la maladie qui ont un test négatif.

Les caractéristiques du test (suite):

- **Utilité clinique: dépend de la prévalence de la maladie.**
 - **Valeur prédictive positive (VPP):** probabilité d'avoir la maladie sachant que le test est positif.
 - **Valeur prédictive négative (VPN):** probabilité de ne pas avoir pas la maladie sachant que le test est négatif.

Estimation des caractéristiques du test:

Etude transversale



Calcul des caractéristiques du test :

		Maladie précoce		
		Présente (+)	Absente (-)	
Test	+	a	b	a+b
	-	c	d	c+d
		a+c	b+d	

Précision

$$Se = a/a+c$$

$$Sp = d/b+d$$

Utilité clinique

$$VPP = a/a+b$$

$$VPN = c/c+d$$

Validité interne:

- **Eviter le biais de mauvaise classification:**
 - Test de référence reconnu et accepté; pas de délai entre le nouveau test et le test de référence.
- **Eviter le biais d'information:**
 - Evaluation des différents tests indépendamment de toute autre information clinique et autres résultats de test.
- **Eviter le biais de vérification:**
 - Le test de référence est appliqué à l'ensemble de la population de l'étude.

Validité externe:

- Caractéristiques des participants de l'étude (par exemple: âge, gamme des stades de la maladie).
- Définition du point de séparation du test (test négatif/test positif).
- Définition de la maladie.

⇒ **Tout cela affecte la capacité de généralisation et la comparaison entre les études.**

Application:

- 100 000 individus, prévalence de la maladie précoce = 1%.
- $Se=90\%$ et $Sp=90\%$ signifie:
 - Détection de 900 cas sur 1 000 vrais cas.
 - Mauvaise classification de 9 900 personnes en bonne santé
- $VPP=8,3\%$ (12 faux positifs pour chaque vrai positif).
- Si la prévalence baisse à 0,1%, $VPP = 0,9\%$

Efficacité des programmes de prévention:

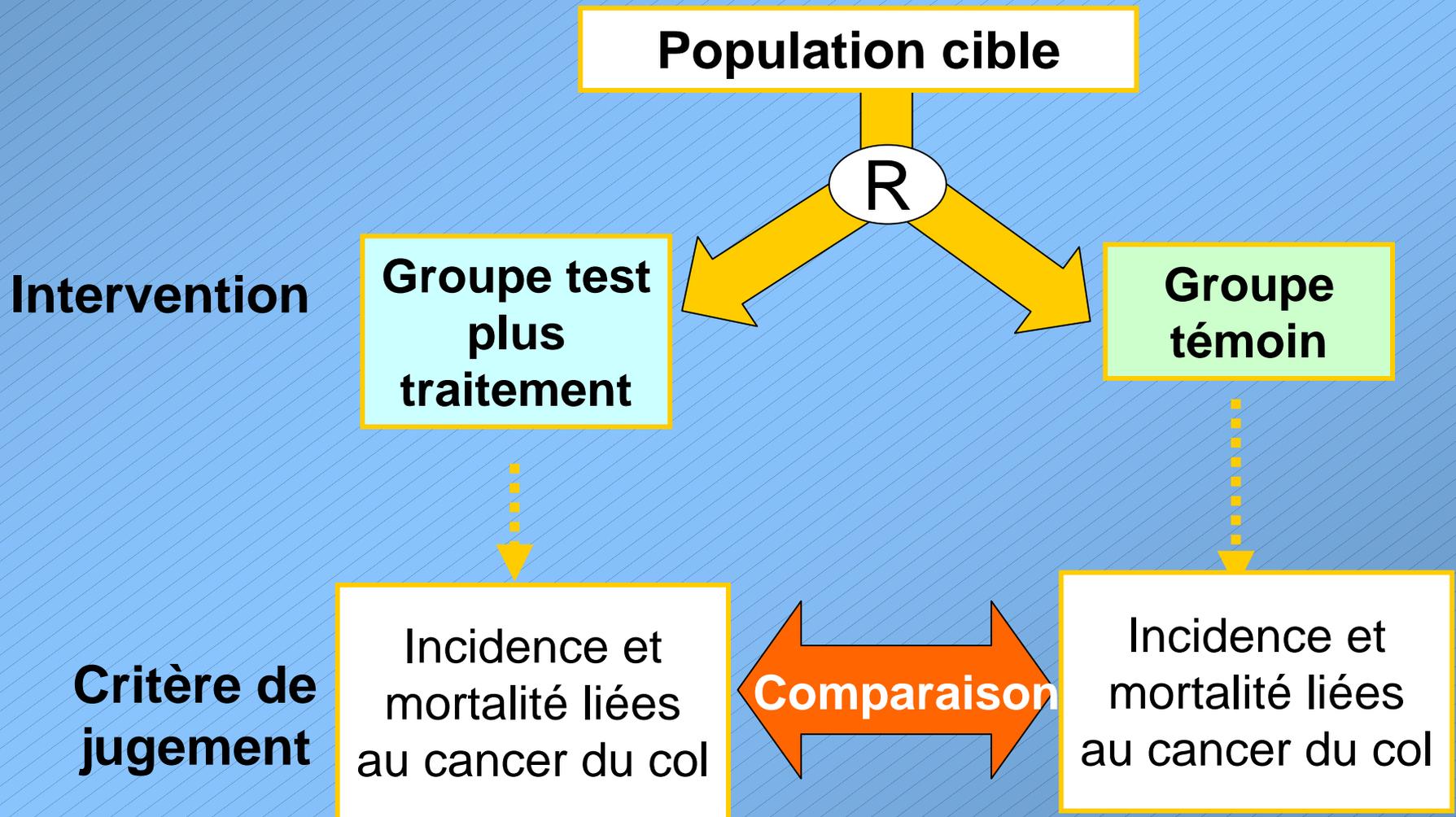
- **Un test approprié ne signifie pas un programme efficace.**
- **On a aussi besoin d'un système efficace de prestation de services:**
 - Bonne couverture du test.
 - Prise en charge appropriée des femmes dépistées positives (limitation du nombre de perdues de vue).
 - Un traitement efficace et acceptable pour un coût raisonnable.

Mesure de l'efficacité théorique et pratique d'un programme:

Types d'étude:

- Etudes expérimentales:
 - Essais contrôlés randomisés
 - Essais contrôlés non randomisés
- Etudes d'observations:
 - Etude de cohorte
 - Etude cas-témoins
 - Etude écologique

Essais contrôlés randomisés:



Essais contrôlés randomisés:

- Répartition aléatoire des personnes/communautés dans l'un ou l'autre groupe pour permettre la comparabilité.
- Standardisation de l'intervention (le test ET les modalités de traitement) pour permettre la comparabilité et la reproductibilité.

C'est la meilleure méthodologie mais elle est très laborieuse.

Considérations d'efficacité pratique:

Le fait que le protocole d'un programme soit efficace théoriquement à travers un essai randomisé ne signifie pas qu'il est efficace pratiquement dans des conditions de programme normales.

- **Résultat à long terme d'un essai randomisé** = résultat d'une application stricte d'un protocole standardisé dans des conditions idéales (efficacité théorique).
- **Efficacité pratique** = les améliorations attendues en santé résultant d'un programme de prestation de services de routine.

Le travail de l'Alliance pour la Prévention du Cancer du Col Utérin (ACCP) :

L'efficacité du test de Papanicolaou à réduire la mortalité du cancer du col est généralement reconnue dans les pays dotés de programmes de dépistage bien organisés mais la réussite de la mise en œuvre est difficile dans les régions à faibles ressources.

- En réponse à ces défis, l'ACCP a mis en place:
 - **Des études transversales** pour évaluer les caractéristiques de tests peu onéreux dans différentes régions.
 - **Des essais randomisés** pour répondre aux questions relevant de l'efficacité théorique des tests de dépistage et des traitements, intégrés aux approches de prestation de services spécifiques.
 - **Des projets pilotes** pour déterminer l'efficacité pratique d'algorithmes alternatifs de prévention dans les pratiques de routine.

Conclusions:

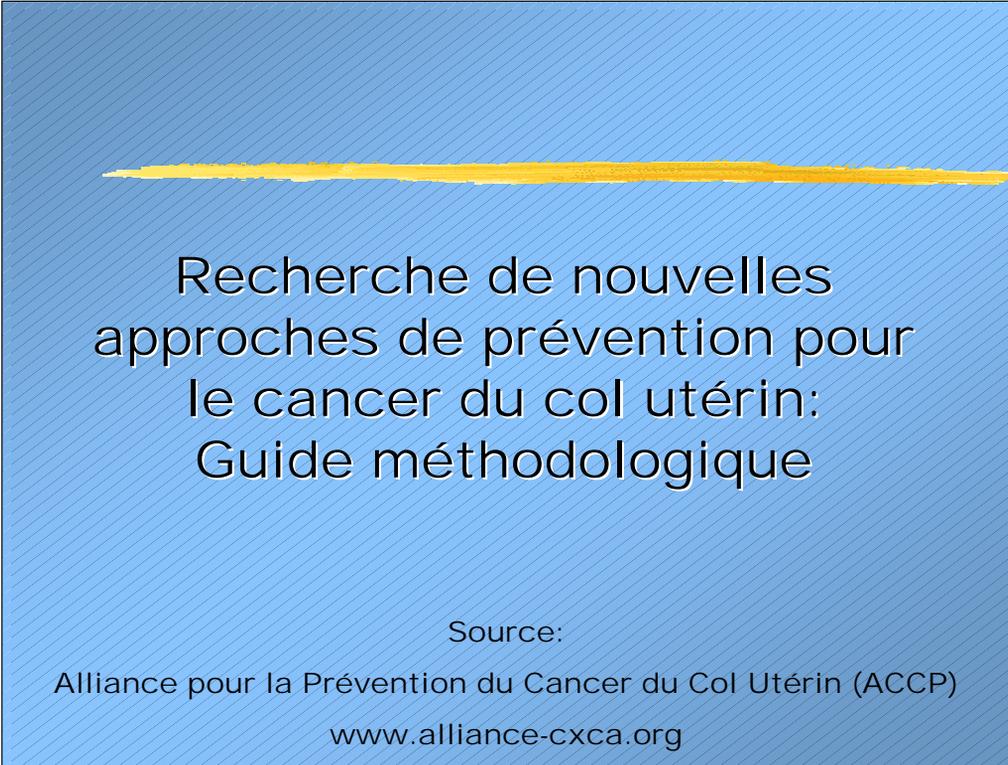
- Déterminer les caractéristiques d'un test nécessite une étude transversale rigoureuse.
- Le choix d'un bon test ne signifie pas nécessairement que l'on obtiendra un programme de prévention efficace.
- Les essais randomisés sont préférables pour évaluer l'efficacité théorique d'un programme, mais exigent la mobilisation d'importantes ressources humaines.
- L'évaluation des projets pilotes et des études d'observation est utile pour déterminer l'efficacité pratique de stratégies de prévention sélectionnées dans un contexte de routine.

Références bibliographiques:

- Hulley SB, Cummings SR (eds.). Designing Clinical Research: An Epidemiologic Approach. Baltimore: Williams and Wilkins, 1988.
- Riegelman RK. Studying a Study and Testing a Test: How to Read the Medical Literature (2nd Edition). Boston: Little, Brown and Company, 1989.
- Sackett DL, Haynes RB, Guyatt GH, Tugwell P. Clinical Epidemiology: A Basic Science for Clinical Medicine. Boston: Little, Brown and Company, 1991.
- Weiss NS. Clinical Epidemiology: The Study of the Outcome of Illness. New York: Oxford University Press, 1986.

Pour plus d'information sur la prévention du cancer du col utérin:

- Alliance pour la Prévention du Cancer du Col Utérin (ACCP) www.alliance-cxca.org
- Les Partenaires de l'ACCP sont:
 - Centre International de Recherche sur le Cancer (CIRC) www.iarc.fr
 - EngenderHealth www.engenderhealth.org
 - JHPIEGO www.jhpiego.org
 - Organisation Panaméricaine de la Santé (OPS) www.paho.org
 - Programme de Technologie Appropriée en Santé (PATH) www.path.org



Recherche de nouvelles approches de prévention pour le cancer du col utérin: Guide méthodologique

Source:

Alliance pour la Prévention du Cancer du Col Utérin (ACCP)

www.alliance-cxca.org

Aperçu général: Cette présentation donne un aperçu des méthodes employées pour établir la valeur de nouvelles approches en matière de prévention du cancer du col utérin.

•*Notes:* Dans de nombreux pays, l'approche traditionnelle pour la prévention du cancer du col consiste en un dépistage par frottis et l'orientation des cas positifs au test vers d'autres structures disposant d'un test diagnostique. Avec cette approche, les femmes positives au second test subissent une biopsie et l'échantillon est envoyé à un laboratoire d'histologie pour confirmation. La femme doit se déplacer une troisième fois pour obtenir les résultats de la biopsie et les recommandations pour la prise en charge, y compris le traitement, le cas échéant—qui peut nécessiter une visite supplémentaire.

•*Notes:* Les nouvelles approches de prévention présentent des alternatives au traditionnel frottis pour détecter les maladies précoces et/ou de nouvelles manières d'associer test et prise en charge des cas positifs.

Les objectifs d'un programme de prévention:

- Réduire l'incidence et la mortalité liées au cancer du col.

Conditions requises:

- Détection de lésions précancéreuses du col (maladie précoce) à l'aide d'un test approprié.
- Traitement de la maladie précoce pour éviter la progression de la maladie.

Aperçu général: Un programme de prévention du cancer du col a pour but de réduire l'incidence et la mortalité par cancer du col.

•*Note pour point 2:* On ne peut réussir à réduire l'incidence et la mortalité que si 3 conditions sont remplies:

- Utiliser un test approprié pour déceler les lésions précancéreuses du col.
- Traiter la maladie précoce.
- Suivre les patientes positives afin de limiter les pertes de vue entre le test et le traitement.

Test approprié pour déceler un précancer:

- Faible coût
- Sans danger
- Pratique et acceptable
- Fiable
- Performant

Aperçu général: La première condition décrite dans la diapo précédente était d'employer un test approprié. Il y a un certain nombre de facteurs qui influencent le choix d'un test approprié.

•*Note pour point 1:* Le coût du test doit être suffisamment faible pour qu'une forte couverture soit possible dans un pays en développement malgré des budgets de santé limités.

•*Note pour point 2:* Le test doit être le moins douloureux et provoquer le moins d'inconfort possibles et n'occasionner aucune complication majeure.

•*Note pour point 3:* Le test doit être pratique et acceptable pour les femmes et les autres membres de la communauté de façon à augmenter la participation des femmes. Ce facteur peut être étudié à travers des études qualitatives au sein de la communauté.

•*Note pour points 4 et 5:* En plus de ces 3 facteurs, le test doit être fiable et être suffisamment performant pour permettre la détection précise de la maladie précoce et une bonne utilité clinique.

Fiabilité du test :

Capacité du test à donner plusieurs fois des résultats identiques (que le testeur soit la même personne ou non).

- La fiabilité dépend de:
 - La manifestation clinique de la maladie précoce
 - La méthode de mesure (subjectivité du test)
 - Les compétences de la personne effectuant le test
 - Le nombre d'étapes pour la réalisation du test
- Faible fiabilité = faible reproductibilité

Aperçu général: La fiabilité détermine le degré avec lequel les mesures répétées d'un test (par la même personne ou non) aboutissent au même résultat.

- Quand le résultat du test pour un même individu varie selon les testeurs, on parle de variabilité inter-observateurs.
- Quand les mesures répétées par la même personne pour un même individu sont différentes, on parle de variabilité intra-observateur.

•*Notes pour point 1:* La fiabilité dépend de:

- La manifestation clinique de la maladie précoce (précancer).
- La méthode de mesure: le test visuel est probablement moins fiable qu'un test biologique parce qu'il est dépendant de la personne qui l'effectue.
- Plus la personne qui réalise le test est compétente, plus le test est fiable. Des procédures standardisées et une formation adéquate des personnes effectuant le test aux procédures standard permettent de minimiser la variabilité.
- Pour les tests qui impliquent plusieurs étapes, le risque d'erreur dans le processus est plus grand. Plus il y a d'étapes, et plus la fiabilité potentielle diminue.
- La fiabilité du test est déterminée grâce à des études de reproductibilité.

•*Notes pour point 2:* Si un test n'est pas très fiable, son application dans différents contextes et par différentes personnes aboutira probablement à des résultats différents.

Les caractéristiques du test:

■ Précision: caractéristiques intrinsèques

- **Sensibilité (Se):** proportion des femmes ayant réellement la maladie qui ont un test positif.
- **Spécificité (Sp):** proportion des femmes n'ayant pas la maladie qui ont un test négatif.

Aperçu général: Une des caractéristiques d'un test est sa capacité à donner le vrai statut vis-à-vis de la maladie (malade ou non) et son utilité dans la pratique clinique à différencier les patientes qui ont probablement la maladie recherchée de celles qui ne l'ont probablement pas.

• *Notes pour point 1:* On emploie traditionnellement différents indicateurs pour déterminer la capacité d'un test à déceler la présence de la maladie précoce (précancer).

- La sensibilité et la spécificité mesurent la capacité d'un test à distinguer correctement les personnes malades des non-malades.
- La sensibilité et la spécificité sont considérées comme des caractéristiques intrinsèques du test car, en principe, elles ne varient pas en fonction de la prévalence de la maladie.
- Habituellement, si une méthodologie adéquate est utilisée pour les estimer, il est possible de comparer les valeurs de la sensibilité et de la spécificité entre différents sites et différents tests. Cette propriété rend la sensibilité et la spécificité utiles à ceux qui recommandent des politiques indiquant quel test est le plus approprié selon les contraintes, exigences et attentes locales.
- La sensibilité et la spécificité sont habituellement corrélées. L'augmentation de l'une implique la diminution de l'autre.
- On ne peut attendre d'un test peu fiable qu'il soit très précis. D'un autre côté, un test avec une très forte fiabilité peut ne pas être très précis, si les mesures répétées sont les mêmes mais incorrectes.

Les caractéristiques du test (suite):

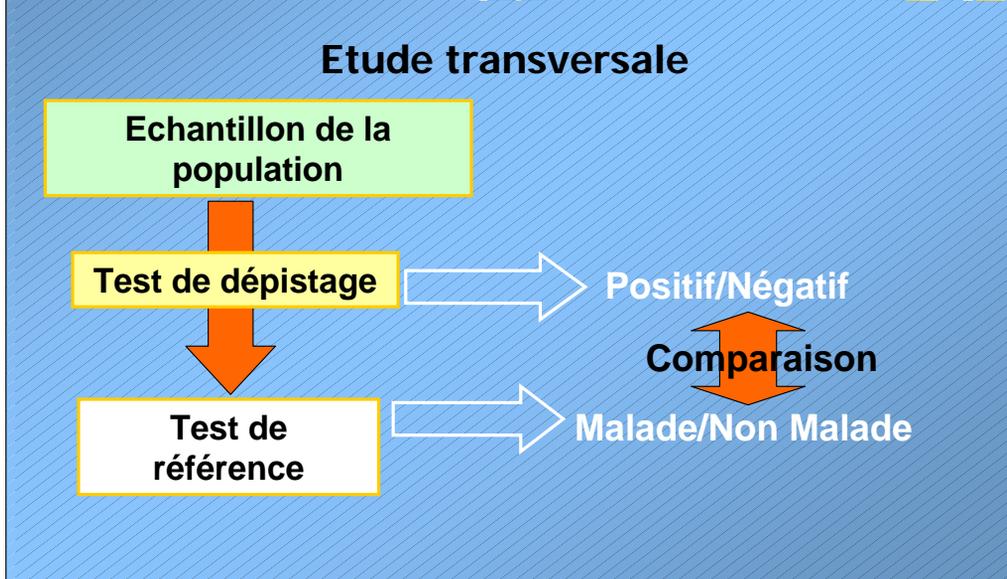
- **Utilité clinique: dépend de la prévalence de la maladie.**
 - **Valeur prédictive positive (VPP):** probabilité d'avoir la maladie sachant que le test est positif.
 - **Valeur prédictive négative (VPN):** probabilité de ne pas avoir pas la maladie sachant que le test est négatif.

Aperçu général: L'utilité clinique est une autre caractéristique importante du test.

•*Notes:* La valeur prédictive mesure la capacité d'un test à prédire la maladie ou l'absence de maladie quand on connaît la prévalence sous-jacente de la maladie dans la population cible. Plus la prévalence de la maladie dans cette population est élevée, plus la valeur prédictive positive du test appliquée à cette population sera élevée.

•*Notes:* Les différences des valeurs prédictives entre les sites d'étude ne peuvent pas être facilement interprétés sans connaître la prévalence de la maladie sous-jacente dans la population étudiée.

Estimation des caractéristiques du test:



Aperçu général : Pour estimer les caractéristiques du test, on mène habituellement des études transversales.

Ces études ont pour but d'estimer la précision d'un test pour détecter les maladies précoces *prévalentes* (par exemple: précancer).

• *Notes*: Avec ce type d'étude, habituellement un échantillon représentatif de la population cible (à risque de la maladie) reçoit indépendamment deux tests:

- Le premier est le test de dépistage que l'on souhaite étudier: les résultats sont classifiés comme positifs (suspicion de maladie précoce) ou comme négatifs (pas de suspicion de maladie précoce).
- Le second est un test de référence qui identifie « correctement » l'existence d'une maladie précoce ou non. Ce doit être un test communément accepté comme étant aussi proche que possible de la réalité c'est-à-dire « la référence ». Ses résultats sont classés selon deux catégories: maladie absente ou présente.

• *Notes*: Les caractéristiques d'un test sont calculées en comparant les résultats d'un test de dépistage à ceux que l'on considère comme « vrais » (c'est-à-dire le test de référence).

Calcul des caractéristiques du test :

		Maladie précoce		
		Présente (+)	Absente (-)	
Test	+	a	b	a+b
	-	c	d	c+d
		a+c	b+d	

Précision

$$Se = a/a+c$$

$$Sp = d/b+d$$

Utilité clinique

$$VPP = a/a+b$$

$$VPN = c/c+d$$

Aperçu général: Ce tableau croisé illustre comment les quatre indicateurs—sensibilité, spécificité et valeur prédictive positive et négative—sont calculés traditionnellement dans des études de caractéristiques de test impliquant une maladie précoce (par exemple: précancer).

•*Notes:* Les quatre cellules du tableau doivent être remplies complètement, c'est-à-dire qu'à chaque résultat de test de dépistage correspond un résultat de test de référence.

•*Notes:* Les formules en bas sont standard pour le calcul d'estimation directe des quatre caractéristiques du test.

Validité interne:

■ Eviter le biais de mauvaise classification:

- Test de référence reconnu et accepté; pas de délai entre le nouveau test et le test de référence.

■ Eviter le biais d'information:

- Evaluation des différents tests indépendamment de toute autre information clinique et autres résultats de test.

■ Eviter le biais de vérification:

- Le test de référence est appliqué à l'ensemble de la population de l'étude.

Aperçu général: Pour que les résultats d'une étude sur les tests caractéristiques soient crédibles, ces études doivent avoir une validité interne. Sinon, l'étude peut engendrer des résultats qui s'écartent systématiquement des vrais valeurs; on parle alors de biais. Cette diapo résume les types de biais qui peuvent affecter les résultats des études portant sur les tests caractéristiques.

• *Note pour point 1:* Il y a biais de mauvaise classification quand le test de référence n'est pas assez précis pour détecter la maladie et les personnes sont « classées » dans la mauvaise catégorie. Pour la détection du précancer du col utérin, un des tests de référence reconnu est la colposcopie, suivie par une biopsie si la colposcopie détecte une quelconque anomalie; c'est-à-dire un test de référence combiné. Pour éviter une mauvaise classification quand on mesure la capacité d'un nouveau test à détecter la maladie prévalente, le nouveau test et le test de référence doivent s'effectuer sans délais l'un après l'autre pour éviter une progression éventuelle de la maladie.

• *Note pour point 2:* Il peut y avoir biais induit si les résultats du test de référence sont déterminés en connaissance soit des résultats du nouveau test, soit de toute information clinique associée au risque de la maladie: on parle de biais d'information. En conséquence, tous les résultats des test de référence doivent être déterminés indépendamment des résultats précédents (à l'aveugle).

• *Note pour point 3:* Il y a biais de vérification quand seulement certains des participants reçoivent le test de référence (par exemple, seulement ceux qui ont un test de dépistage positif). Ce type de biais tend à surestimer la valeur de la sensibilité et à sous-estimer la valeur de la spécificité car on ne « vérifie » pas les personnes dépistées négatives.

Validité externe:

- Caractéristiques des participants de l'étude (par exemple: âge, gamme des stades de la maladie).
- Définition du point de séparation du test (test négatif/test positif).
- Définition de la maladie.

⇒ Tout cela affecte la capacité de généralisation et la comparaison entre les études.

Aperçu général: La validité externe se rapporte à la mesure dans laquelle les résultats d'une étude sont applicables à une population sur laquelle le test sera effectuée.

•*Note pour point 1:* Les études qui incluent les femmes plus âgées, ménopausées, sont susceptibles de fournir des résultats différents des études sur des femmes plus jeunes parce que les différences physiologiques ou anatomiques affectent la performance du test. L'interprétation des différences dans les résultats d'une étude et leur application à une population clinique particulière nécessite la connaissance des caractéristiques de cette population d'étude. Si un test doit être effectué dans la population générale pour identifier la maladie précoce mais que les études ont seulement porté sur des patientes orientées (avec une gamme plus limitée de stades de la maladie), alors le test peut ne pas fonctionner en pratique comme il a été supposé à partir des résultats de l'étude.

•*Note pour point 2:* Les systèmes de classification de test ne sont pas standard. En conséquence, les points de séparation distinguant ce qui est un test positif d'un test négatif dans une étude peuvent différer de ceux d'une autre étude. Cela peut aboutir à des estimations différentes des caractéristiques du test.

•*Note pour point 3:* Les catégories de maladie d'une étude peuvent ne pas correspondre forcément à ceux d'une autre étude et elles doivent toujours être clairement définies quand les résultats sont reportés. Pour les études qui mesurent la précision d'un test qui détecte les maladies du col provoquant un cancer, habituellement la lésion de haut grade est le précurseur adéquat en raison de la forte probabilité de progression des lésions de haut grade vers le cancer. Grouper les femmes qui ont un cancer avec celles qui ont une maladie précancéreuse peut surestimer artificiellement la précision du test car le cancer est plus facile à diagnostiquer que les lésions de haut grade et ces personnes peuvent avoir été orientées vers l'étude de l'extérieur (elles ne font pas partie de la population générale).

•*Note finale:* Les différences dans l'un (quelconque) des facteurs ci-dessus entre les études les rendront moins comparables entre elles et l'on devra prendre soin d'appliquer les résultats d'une étude particulière à une population clinique adéquate (comparable).

Application:

- 100 000 individus, prévalence de la maladie précoce = 1%.
- $Se=90\%$ et $Sp=90\%$ signifie:
 - Détection de 900 cas sur 1 000 vrais cas.
 - Mauvaise classification de 9 900 personnes en bonne santé
- $VPP=8,3\%$ (12 faux positifs pour chaque vrai positif).
- Si la prévalence baisse à 0,1%, $VPP = 0,9\%$

Aperçu général: Il n'y a pas de standard universel pour des valeurs de sensibilité ou de spécificité minimum ou acceptables. Les valeurs acceptables dépendent de l'importance relative de l'identification des vrais cas de maladie par rapport à l'erreur consistant à identifier faussement une femme comme malade quand elle ne l'est pas—et de la prévalence de cette maladie dans cette population.

• *Notes:* Par exemple, si l'on considère une population cible de 100 000 individus avec une prévalence de maladie de 1%, c'est-à-dire 1000 vrais cas:

- Un test avec de bonnes caractéristiques (Se et $Sp = 90\%$) détectera 900 cas sur 1000 vrais cas mais il sera également faussement positif pour 9 900 femmes en bonne santé.
- Ceci aboutit à un VPP de 8,3% c'est-à-dire 12 faux positifs (FP) pour chaque vrai positif (VP) identifié.
- Si la prévalence de la vraie maladie est plus basse (0,1%), la sensibilité et la spécificité restent théoriquement les mêmes mais la VPP descend à 0,9% (c'est-à-dire 111 FP pour chaque VP).
- En revanche, si la prévalence de la maladie est plus forte, la VPP sera également plus élevée et le nombre de FP sera plus faible par rapport au nombre de VP identifié.

• *Note supplémentaire:* Lorsqu'on compare des tests, il faut prendre en considération d'autres facteurs, comme l'impact psychologique d'un résultat faussement positif, la prise en charge de ces tests faussement positifs, le coût relatif pour le système de soin, la réaction de la femme et de la société face à un résultat faux positif par rapport à un résultat faussement négatif et les bénéfices relatifs qu'il y a à identifier une vraie maladie par rapport à la confirmation de l'absence de maladie.

Efficacité des programmes de prévention:

- **Un test approprié ne signifie pas un programme efficace.**
- **On a aussi besoin d'un système efficace de prestation de services:**
 - Bonne couverture du test.
 - Prise en charge appropriée des femmes dépistées positives (limitation du nombre de perdues de vue).
 - Un traitement efficace et acceptable pour un coût raisonnable.

Aperçu général: Divers critères autres que la précision d'un test, son utilité clinique et sa fiabilité permettent à la détection précoce d'améliorer l'issue à long terme. En effet, la détection précoce, si elle n'est pas liée au traitement, n'a pas d'effet à long terme.

• *Note pour point 2:* Pour qu'un programme réduise la mortalité et la morbidité de façon efficace et efficiente, il faut aussi satisfaire d'autres critères élémentaires liés au système de prestation des services:

- La proportion de femmes recevant le test (couverture du test) doit être importante de façon à identifier la majorité des cas prévalents de précancers susceptibles de progresser vers un cancer dans la population.
- Il faut disposer d'un protocole de prise en charge approprié des femmes testées positives. Ce protocole doit permettre à la majorité des lésions détectées d'être finalement traitées (limitation des perdues de vue).
- Le traitement doit être efficace pour le traitement du précancer, acceptable, et disponible à un prix raisonnable (en prenant en compte les budgets de santé limités des pays en développement).

• *Note finale:* Une étude appropriée doit déterminer la mesure dans laquelle ces facteurs, outre la précision du test et sa fiabilité, affectent l'efficacité du programme en général.

Mesure de l'efficacité théorique et pratique d'un programme:

Types d'étude:

- Etudes expérimentales:
 - | Essais contrôlés randomisés
 - | Essais contrôlés non randomisés
- Etudes d'observations:
 - | Etude de cohorte
 - | Etude cas-témoins
 - | Etude écologique

Aperçu général: L'efficacité théorique d'un programme de prévention du cancer du col se mesure à la réduction de la morbidité et de la mortalité dans des conditions idéales ; l'efficacité pratique d'un programme mesure les mêmes résultats à long terme dans des conditions normales de prestation des services.

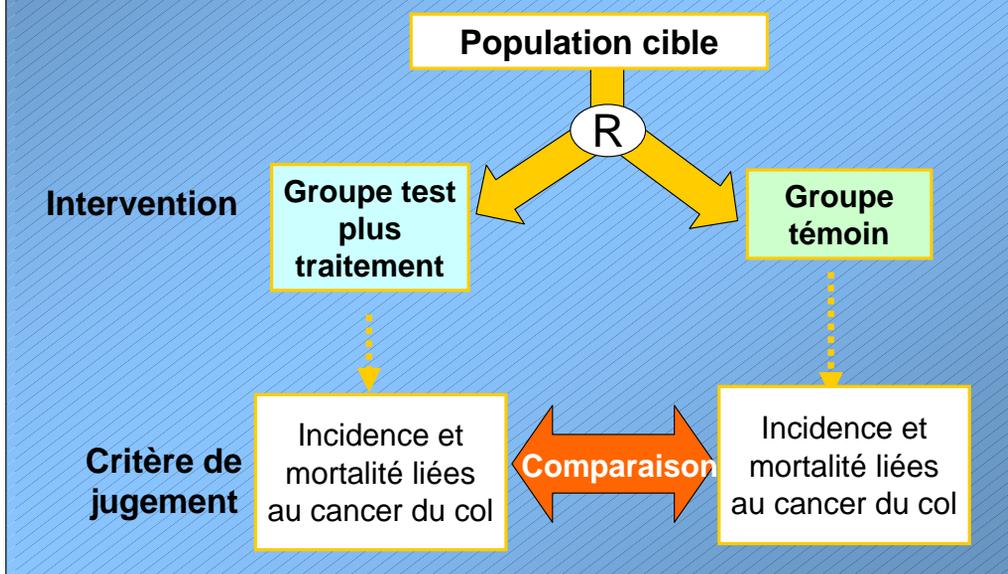
Note pour point 1: Pour mesurer comment le dépistage associé au traitement aboutit dans le temps à une réduction de la morbidité et de la mortalité, on peut employer des études d'intervention ou d'observation. Cette diapo énumère les types d'étude dans chacune des deux principales catégories dans l'ordre décroissant de « rigueur » d'étude. Plus le type d'étude est rigoureux, plus les conclusions qui peuvent être effectuées à partir de ces études sont solides.

Dans les études expérimentales ou d'intervention, les chercheurs contrôlent la façon dont l'intervention est effectuée. Dans le cas de la prévention du cancer du col, l'intervention implique de détecter la maladie grâce au dépistage, d'associer le test au traitement et de traiter efficacement les cas de précancers. De telles études sont élaborées pour tester une hypothétique relation de cause à effet en modifiant le facteur causal (dans ce cas le programme de prévention).

Dans les études d'observation, les investigateurs ne contrôlent pas la façon dont l'intervention est appliquée à la population. Les changements sur l'issue à long terme sont étudiés par rapport aux différences dans le programme et il faut s'efforcer de réduire les sources de biais qui pourraient affecter l'interprétation des résultats, étant entendu que les chercheurs ne contrôlent pas l'intervention.

On peut trouver la description détaillée des différentes méthodes utilisées dans les types d'étude mentionnées plus haut, ainsi que leurs avantages et inconvénients relatifs, dans les manuels d'épidémiologie standard.

Essais contrôlés randomisés:



Aperçu général: Dans un essai contrôlé randomisé, pour optimiser la comparabilité, les personnes admissibles dans la population cible sont réparties *de façon aléatoire* dans un groupe d'intervention (par exemple: test et traitement si appropriés) ou dans un groupe témoin (lequel peut néanmoins recevoir un traitement si un cancer est identifié). Ces groupes d'intervention sont appelés des "branches."

Notes:

- Cette présentation est centrée sur la méthodologie des essais randomisés contrôlés car elle est moins susceptible d'être biaisée.
- Les participants sont suivis pendant plusieurs années de façon à surveiller l'incidence ou la mortalité liées au cancer du col.
- La comparaison de ces résultats à long terme possibles indique la mesure dans laquelle le programme de prévention (test et traitement si indiqué) a été efficace.

Essais contrôlés randomisés:

- Répartition aléatoire des personnes/communautés dans l'un ou l'autre groupe pour permettre la comparabilité.
- Standardisation de l'intervention (le test ET les modalités de traitement) pour permettre la comparabilité et la reproductibilité.

C'est la meilleure méthodologie mais elle est très laborieuse.

Aperçu général: La randomisation est destinée à accroître la probabilité que les facteurs autres que l'intervention elle-même (le programme) affectent de la même façon tous les groupes de l'étude. Ainsi, toutes les différences observées dans les résultats à long terme refléteront plus vraisemblablement l'effet des différences entre les groupes de l'étude.

• *Note pour point 1:* La randomisation peut être effectuée au niveau communautaire plutôt qu'au niveau individuel pour des raisons pratiques et notamment pour éviter la contamination entre les différents groupes.

• *Note pour point 2:* Si l'intérêt principal du programme est son efficacité maximale (son efficacité théorique), les chercheurs doivent s'assurer qu'un protocole standard est appliqué à tous les groupes pour permettre la reproductibilité des modalités d'intervention.

• *Notes complémentaires:*

- Les essais randomisés nécessitent habituellement un suivi à long terme et un large échantillon. La randomisation à l'aveugle pour limiter le biais potentiel résultant de cette connaissance est aussi idéale mais rarement applicable.

- Si les essais randomisés sont considérés comme le type d'étude de référence permettant de déterminer l'efficacité théorique des programmes de détection précoce et de traitement, jusqu'à maintenant les preuves de l'efficacité pratique du test de Papanicolaou à réduire la mortalité liée au cancer du col sont issues d'études d'observation (études cas-témoins et de cohorte).

• *Note finale:* Dans les essais randomisés, l'équipe de recherche contrôle autant d'aspects de l'étude que possible afin d'optimiser la validité interne et donc la crédibilité des résultats. Néanmoins, ceci peut réduire la capacité de généralisation externe parce que les résultats représentent la façon dont le programme fonctionne dans des conditions contrôlées plus « idéales » (efficacité théorique).

Considérations d'efficacité pratique:

Le fait que le protocole d'un programme soit efficace théoriquement à travers un essai randomisé ne signifie pas qu'il est efficace pratiquement dans des conditions de programme normales.

- **Résultat à long terme d'un essai randomisé** = résultat d'une application stricte d'un protocole standardisé dans des conditions idéales (efficacité théorique).
- **Efficacité pratique** = les améliorations attendues en santé résultant d'un programme de prestation de services de routine.

Aperçu général: Il est important de comprendre de quelle façon un essai randomisé nous informe et dans quelle mesure les résultats peuvent différer en pratique.

•*Note pour point 1:* Les essais randomisés sont des études d'efficacité théorique cherchant des améliorations en terme de santé qui résultent de l'application stricte d'un protocole standardisé pour la détection précoce et le traitement de cas identifiés. Ces études mesurent dans des circonstances idéales les bénéfices sanitaires maximums que les programmes impliquant l'utilisation d'un nouveau test, lié au traitement sont supposés atteindre.

•*Note pour point 2:* Comme les programmes ne se déroulent pas de façon rigoureuse et standard en pratique, il est intéressant de connaître la vraisemblable amélioration sanitaire qu'un programme peut apporter, dans des conditions normales impliquant un nouveau test lié à un traitement. Cette information est fournie par des études d'efficacité pratique différentes des essais randomisés.

Le travail de l'Alliance pour la Prévention du Cancer du Col Utérin (ACCP) :

L'efficacité du test de Papanicolaou à réduire la mortalité du cancer du col est généralement reconnue dans les pays dotés de programmes de dépistage bien organisés mais la réussite de la mise en œuvre est difficile dans les régions à faibles ressources.

- En réponse à ces défis, l'ACCP a mis en place:
 - **Des études transversales** pour évaluer les caractéristiques de tests peu onéreux dans différentes régions.
 - **Des essais randomisés** pour répondre aux questions relevant de l'efficacité théorique des tests de dépistage et des traitements, intégrés aux approches de prestation de services spécifiques.
 - **Des projets pilotes** pour déterminer l'efficacité pratique d'algorithmes alternatifs de prévention dans les pratiques de routine.

Aperçu général: L'ACCP recherche différentes méthodes alternatives pour la prévention du col dans des régions à faibles ressources.

• *Notes:* Le frottis est un test de dépistage acceptable dans les programmes de bonne qualité mais son système de prestation de services n'est pas toujours efficace dans les pays en développement (visites multiples, pas de procédures de contrôle qualité etc.). De plus, sa mise en œuvre par le biais d'un bon système de prestation de services peut se révéler coûteuse.

• *Note pour les différents points:* Pour répondre à ce défi, l'ACCP a élaboré des études portant sur différents tests (IVA, IVL, test de recherche de l'ADN du VPH, cytologie de qualité) et sur les façons de lier le test au traitement.

- La précision et la fiabilité des tests sont étudiées à l'aide d'études transversales dans plusieurs pays.
- L'efficacité théorique est étudiée à l'aide de grands essais randomisés comparant différents algorithmes de tests et de traitements. Les résultats de ces essais randomisés seront disponibles dans les prochaines années pour la morbidité et d'ici à 10 ans pour leur effet sur la mortalité.
- L'efficacité pratique est évaluée à travers des programmes pilotes dans différents endroits.

Conclusions:

- Déterminer les caractéristiques d'un test nécessite une étude transversale rigoureuse.
- Le choix d'un bon test ne signifie pas nécessairement que l'on obtiendra un programme de prévention efficace.
- Les essais randomisés sont préférables pour évaluer l'efficacité théorique d'un programme, mais exigent la mobilisation d'importantes ressources humaines.
- L'évaluation des projets pilotes et des études d'observation est utile pour déterminer l'efficacité pratique de stratégies de prévention sélectionnées dans un contexte de routine.

Aperçu général: Pour conclure, il y a 4 étapes à envisager dans la recherche de nouvelles stratégies de prévention du cancer du col:

- Note pour point 1:* Un nouveau test doit être évalué par une étude transversale rigoureuse pour déterminer sa précision.
- Note pour point 2:* Le test doit être lié au traitement dans le cadre d'un système de prestation de services de soins réalisable pour déterminer sa valeur dans la prévention du cancer du col.
- Note pour point 3:* Le choix du meilleur algorithme de test et de traitement doit être fait à l'aide d'essais randomisés mais ils exigent la mobilisation d'importantes ressources humaines.
- Note pour point 4:* Les évaluations des projets pilotes et des études d'observation peuvent fournir des preuves de l'efficacité pratique de l'algorithme choisi de prestation des services dans un contexte de routine.

Références bibliographiques:

- Hulley SB, Cummings SR (eds.). Designing Clinical Research: An Epidemiologic Approach. Baltimore: Williams and Wilkins, 1988.
- Riegelman RK. Studying a Study and Testing a Test: How to Read the Medical Literature (2nd Edition). Boston: Little, Brown and Company, 1989.
- Sackett DL, Haynes RB, Guyatt GH, Tugwell P. Clinical Epidemiology: A Basic Science for Clinical Medicine. Boston: Little, Brown and Company, 1991.
- Weiss NS. Clinical Epidemiology: The Study of the Outcome of Illness. New York: Oxford University Press, 1986.

Pour plus d'information sur la prévention du cancer du col utérin:

- Alliance pour la Prévention du Cancer du Col Utérin (ACCP) www.alliance-cxca.org
- Les Partenaires de l'ACCP sont:
 - Centre International de Recherche sur le Cancer (CIRC) www.iarc.fr
 - EngenderHealth www.engenderhealth.org
 - JHPIEGO www.jhpiego.org
 - Organisation Panaméricaine de la Santé (OPS) www.paho.org
 - Programme de Technologie Appropriée en Santé (PATH) www.path.org