

## CHAPITRE 12

# La rigueur dans l'utilisation de méthodes mixtes pour les évaluations d'interventions de santé en Afrique subsaharienne : où en est-on ?

*Lara Gautier<sup>1</sup>  
Mathieu Seppey<sup>2</sup>  
Amina Achaïbou<sup>3</sup>*

<sup>1</sup> École de santé publique de l'Université de Montréal, Institut de recherche en santé publique de l'Université de Montréal et CESSMA/Institut de recherche pour le développement, Université Paris Diderot-Paris VII, Paris Sorbonne Cité

<sup>2</sup> École de santé publique de l'Université de Montréal et Institut de recherche en santé publique de l'Université de Montréal

<sup>3</sup> École de santé publique de l'Université de Montréal

## RÉSUMÉ

Ces dernières années, l'usage de méthodes mixtes pour évaluer les interventions de santé en Afrique subsaharienne a considérablement augmenté. Malgré le fort potentiel des méthodes mixtes pour l'évaluation de la complexité de ce type d'interventions, le risque est de « faire du qualitatif à rabais » ou de conduire des analyses statistiques peu rigoureuses. Pourtant, peu de chercheurs s'intéressent à la rigueur des méthodes mixtes. Nous proposons donc une analyse de la littérature scientifique de janvier 2010 à décembre 2015. Nous avons identifié 268 articles sur PubMed, Web of Science et CINAHL. Leurs données ont été synthétisées sur Excel et analysées avec l'outil *Mixed Methods Appraisal Tool* de Pluye et autres (2011). Nos résultats montrent que la qualité des articles évaluant des interventions de santé mondiale en utilisant les méthodes mixtes est variable. Du côté quantitatif, des liens de causalité sont parfois établis sans tenir compte de certains biais. Toutefois, la plupart des insuffisances sont constatées du côté qualitatif : pratiquement aucun article ne justifie le choix de l'échantillonnage ou de la stratégie de recrutement des participants et peu de critères visant à assurer la rigueur méthodologique sont utilisés, mis à part certains types de triangulation. Ainsi, la crédibilité des résultats, la confirmation auprès des participants ou la réflexivité du chercheur sont souvent des critères de rigueur laissés pour compte.

## MOTS-CLÉS

Méthodes mixtes, rigueur, interventions, santé mondiale, Afrique subsaharienne

## INTRODUCTION

Ces dernières années, l'usage de méthodes mixtes pour évaluer les interventions de santé a considérablement augmenté (Pluye, 2012). Ceci s'explique probablement par la forte valorisation de cette approche méthodologique par la communauté scientifique en santé. Pluye (2012) définit l'approche des méthodes mixtes comme « la combinaison d'une méthode qualitative et d'une méthode quantitative » par opposition à la combinaison de deux méthodes du même type, c'est-à-dire la « méthode multiple » (Suárez-Herrera et Blain, 2012, p. 27). On reconnaît généralement trois principaux apports des méthodes mixtes : l'intérêt de pouvoir interpréter des résultats quantitatifs grâce à une analyse complémentaire employant les méthodes qualitatives; la nécessité d'employer des méthodes quantitatives pour généraliser de manière statistique des résultats qualitatifs; et enfin, la possibilité « de mieux comprendre un phénomène nouveau (méthodes qualitatives) et d'en mesurer l'ampleur, l'évolution, les causes et les effets (méthodes quantitatives) » (Suárez-Herrera et Blain, 2012, p. 27). Les bénéfices de la combinaison des méthodes qualitatives et quantitatives reposent également sur leur potentiel d'accroissement de la validité et de la crédibilité des résultats d'une étude : en effet, plus les résultats sont confirmés par des sources variées de données, plus ils sont dignes de confiance (Johnson et Onwuegbuzie, 2004).

Toutefois, plusieurs problèmes se posent concernant l'utilisation des méthodes mixtes : on relève notamment l'écueil d'un simple ajout marginal d'une deuxième méthodologie (qualitative ou quantitative) qui est souvent de faible qualité. En outre, les données collectées sont

susceptibles de se recouper de façon imparfaite, ceci empêchant la triangulation de ces données : on peut parfois faire face à un décalage entre les données de questionnaires/instruments quantitatifs validés et bien contrôlés par le chercheur, et les données qualitatives d'entrevues/observations/groupes de discussion qui reflètent des logiques propres aux informateurs clés (Lehoux, 2015). Enfin, l'intégration des données quantitatives et qualitatives peut s'avérer complexe et parfois inappropriée : les choix épistémologiques peuvent s'avérer difficiles à intégrer, et sont même parfois susceptibles de s'opposer (*ibid.*).

Pourtant, peu de chercheurs s'intéressent actuellement à la rigueur des méthodes mixtes. Ceci tient probablement au fait que les méthodes mixtes constituent un champ méthodologique en développement, encore peu enseigné dans les universités. Pourtant, ce champ méthodologique est de plus en plus demandé depuis une dizaine d'années par les bailleurs de fonds et par plusieurs comités de révision par les pairs (Guével et Pommier, 2012). Cet engouement pousse parfois des chercheurs insuffisamment formés en méthodes mixtes, qualitatives ou quantitatives à utiliser ces différentes méthodes. La plupart de ceux-ci ne seront donc pas totalement aptes à respecter la logique des méthodes mixtes (c'est-à-dire l'intégration d'analyses quantitatives et qualitatives et non la simple superposition de celles-ci), ni celle des méthodes quantitatives (c'est-à-dire la description des caractéristiques d'un phénomène et/ou la mesure des associations causales via des calculs statistiques dont les résultats seront plus ou moins puissants selon la taille de l'échantillon et sa représentativité), ou celle des méthodes qualitatives (c'est-à-dire l'approfondissement de la compréhension d'un phénomène). Ainsi, malgré le fort potentiel des méthodes mixtes pour l'évaluation d'interventions complexes, le risque est de ne pas intégrer les résultats et de faire des collectes et des analyses de données qualitatives ou statistiques peu rigoureuses.

Face à ces écueils potentiels, l'objectif de cette étude est d'apprécier la qualité des articles scientifiques qui rapportent des évaluations d'interventions de santé mondiale dans les pays d'Afrique subsaharienne à l'aide de l'outil MMAT (*Mixed Methods Appraisal Tool*) (Pluye et autres, 2011).

## MÉTHODE

Nous proposons une analyse de la littérature scientifique en anglais de janvier 2010 à décembre 2015, ce qui permet un aperçu des articles publiés les six dernières années. Nous avons effectué des recherches dans les bases de données bibliographiques suivantes : PubMed, Web of Science et CINAHL. Les mots clés utilisés étaient les suivants : « méthodes mixtes », (« santé mondiale » OU « santé internationale »), « intervention\* OU programme\* OU projet\* », tout pays d'Afrique subsaharienne.

### Sélection des articles

Nous avons procédé par étapes pour sélectionner les articles. Après avoir éliminé les doublons, nous avons réalisé une sélection par titre, par résumé, puis par manuscrit complet. Les critères d'inclusion étaient les suivants : l'étude est rédigée en anglais; il s'agit une évaluation d'une intervention en santé mondiale (évaluation d'implantation ou évaluation d'impact); l'intervention a été menée dans un ou plusieurs pays d'Afrique subsaharienne; l'étude fait usage de méthodes mixtes. Nous avons également appliqué les critères d'exclusion suivants : thèmes concernant la santé animale et les zoonoses, la qualité de vie, la gestion des déchets et la santé environnementale; articles méthodologiques; protocoles d'études utilisant les méthodes mixtes; revues de littérature; études présentant un devis de « méthodes multiples » (par ex., incluant méthodes qualitatives et analyses cliniques).

### Extraction et évaluation des données

À partir d'une feuille Excel, nous avons construit une base de données pour analyser de manière systématique l'ensemble des articles sélectionnés. Nous avons extrait les données suivantes des articles : auteurs, année, pays cible, et domaine d'intervention en santé mondiale (par ex., santé maternelle, VIH, vaccination, renforcement des capacités, etc.). Une fois que ceci a été fait, nous avons évalué la qualité des articles en appliquant à chacun les 13 critères du *MMAT* de Pluye et autres (2011) traduits en français (Robert, 2015) (tableau 1).

Il convient de noter que les titres des sections quantitatives « III. ii » et « III. iii » ont été modifiés par rapport à la version originale de l'outil : le « III. ii » avait pour titre « Quantitatif sans randomisation » tandis que le « III. iii » s'intitulait « Quantitatif descriptif ». Nous avons décidé de modifier les titres de ces parties afin de pouvoir inclure les articles de méthodes mixtes dans lesquels les auteurs ne faisaient pas de comparaison intergroupe sans se limiter pour autant à des statistiques descriptives (c'est-à-dire des études rétrospectives transversales sans comparaison intergroupes). Nous expliquons dans la discussion que ce choix est imparfait et qu'une section supplémentaire pour cette catégorie méthodologique mériterait d'être créée pour pouvoir appliquer des critères plus adéquats. En outre, comme indiqué dans l'outil original (Pluye et autres, 2011) et toujours concernant la section « III. iii », le critère 10.c n'était pas toujours applicable. En effet, les auteurs n'administraient pas nécessairement un questionnaire à la population à l'étude, et parfois il n'y avait pas de participants à proprement parler (par ex., dans le cas des analyses quantitatives de documents). Ainsi, pour ces articles, nous n'avons pas mesuré la conformité à ce 10<sup>e</sup> critère : le score final de qualité dans ce cas n'est calculé que sur la base de 12 critères.

Pour des raisons de commodité, nous avons appliqué un simple score binaire (1 ; 0) pour chacun des critères. Nous avons ensuite fait la somme des 1 et 0 pour les 12 ou 13 critères afin d'obtenir le score final de qualité de chaque article, ramené en pourcentage (sur 12 ou 13).

## Synthèse des données

Tout en lisant chaque article évalué, nous prenions des notes pour pouvoir éclairer nos scores de quelques exemples concrets. Une fois que le score en pourcentage était obtenu, nous avons analysé les données descriptives de chaque article (date de publication, pays d'intervention, domaine d'intervention, et scores finaux en pourcentage) à l'aide de tableaux croisés-dynamiques. Nous avons aussi réalisé des comparaisons par année afin d'identifier d'éventuelles évolutions dans la qualité des articles. Nous avons ensuite rapporté les principaux résultats par écrit : ceux-ci sont présentés dans la section suivante.

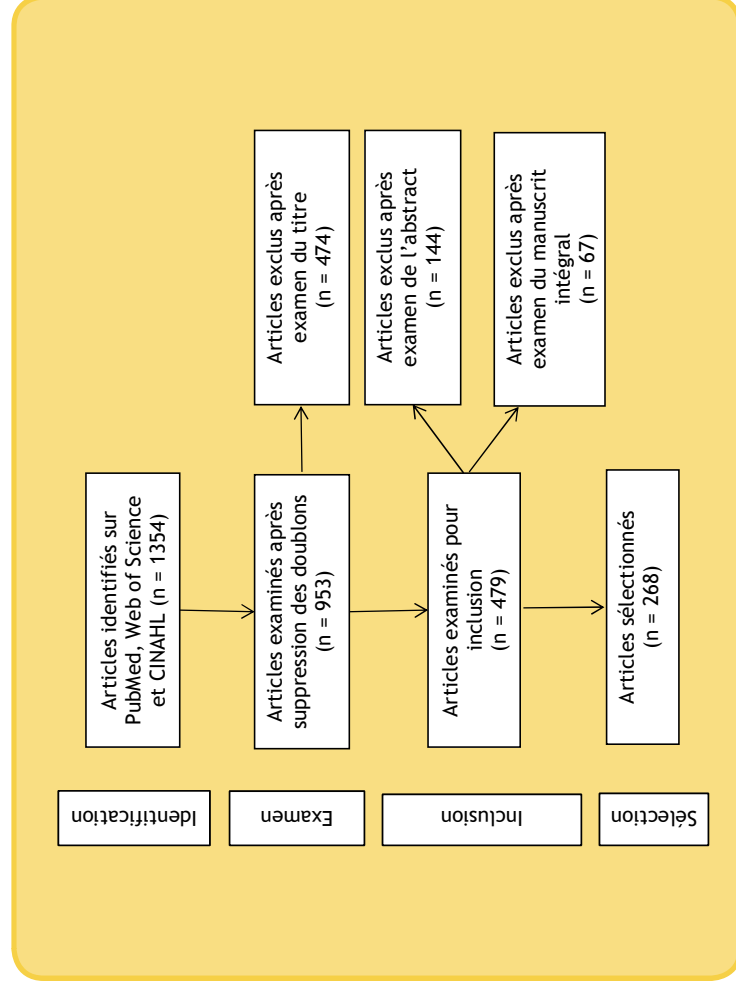
LES 13 CRITÈRES DU MMAT (ADAPTÉ DE PLUYE ET AUTRES, 2011  
ET DE LA TRADUCTION RÉALISÉE PAR ROBERT, 2015)

## Tableau 1

<p>I. QUESTIONS PRÉLIMINAIRES</p> <p>NB : L'évaluation de la qualité peut ne pas être poursuivie si la réponse est « Non » ou « Ne sait pas » à l'une ou aux deux questions.</p>	<p>1. Y a-t-il une question de recherche qualitative et/ou quantitative claire (ou un objectif de recherche clair) ?</p> <p>2. Les données collectées permettent-elles de répondre à la question de recherche (ou à l'objectif de recherche) ? (par ex., la période de suivi est-elle suffisamment longue pour permettre à l'effet de se produire [pour les études longitudinales] ?)</p>
<p>II. PARTIE QUALITATIVE</p>	<p>3. Les sources de données (par ex., participants) permettent-elles de répondre à la question de recherche ?</p> <p>4. Le processus d'analyse des données qualitatives permet-il de répondre à la question de recherche ?</p> <p>5. Les auteurs ont-ils suffisamment pris en compte la manière dont les résultats sont liés au contexte (par ex., le cadre dans lequel les données ont été collectées) ?</p> <p>6. Les auteurs ont-ils suffisamment pris en compte la manière dont leur présence influence les résultats (par ex., leurs interactions avec les participants) ?</p>
<p>III. i. PARTIE QUANTITATIVE</p> <p>AVEC SÉLECTION ALÉATOIRE CONTRÔLÉE</p>	<p>7a. Y a-t-il une description claire du processus de randomisation et/ou d'un processus approprié de distribution dans les groupes ?</p> <p>8a. Y a-t-il une description claire du processus de sélection à l'aveugle ?</p> <p>9a. Les données sont-elles complètes (&gt; 80 %) ?</p> <p>10a. Le nombre de perdus de vue est-il faible (&lt; 20 %) ?</p> <p>7b. Le processus de sélection des participants permet-il de minimiser les facteurs de confusion ?</p>
<p>III. ii. PARTIE QUANTITATIVE</p> <p>SANS SÉLECTION ALÉATOIRE</p> <p>AVEC COMPARAISON DE GROUPES</p>	<p>8b. Les instruments de mesure de l'exposition (ou de l'intervention) et des effets sont-ils appropriés (origine claire, validité connue, outil valide; et absence de contamination entre les groupes si pertinent) ?</p> <p>9b. Dans les groupes comparés (exposés/ non exposés; avec ou sans l'intervention; cas/ contrôles), les participants sont-ils comparables, ou les chercheurs prennent-ils en compte (en les contrôlant) ces différences dans leurs analyses ?</p> <p>10b. Les données sont-elles complètes (&gt; 80 %) et/ou le taux de réponse est-il acceptable (&gt; 60 %), ou un taux acceptable de suivi pour les études de cohorte (selon la durée du suivi) ?</p>
<p>III. iii. PARTIE QUANTITATIVE</p> <p>SANS COMPARAISON DE GROUPES</p>	<p>7c. La stratégie d'échantillonnage permet-elle de répondre à la question de recherche ?</p> <p>8c. L'échantillon est-il représentatif de la population à l'étude ?</p> <p>9c. Les instruments de mesure sont-ils appropriés (origine claire, validité connue, outil valide) ?</p> <p>10c. Le taux de réponse est-il acceptable (&gt; 60 %) ?</p>
<p>IV. MÉTHODES MIXTES</p>	<p>11. Le devis de recherche mixte permet-il de répondre à la question de recherche ?</p> <p>12. L'intégration des données qualitatives et quantitatives (ou des résultats) permet-elle de répondre à la question de recherche ? (NB : Ceci n'est pas une question double car, dans les recherches mixtes, les données qualitatives et quantitatives et/ou les résultats qualitatifs et quantitatifs peuvent être intégrés.)</p> <p>13. Les auteurs ont-ils suffisamment pris en compte les limites associées au processus d'intégration (par ex., divergence des données ou des résultats qualitatifs et quantitatifs dans le cas d'un devis de triangulation) ?</p>

Figure 1

SCHÉMA DES ÉTAPES DE NOTRE REVUE DE LA LITTÉRATURE (ADAPTÉ DE MOHER, LIBERATI, TETZLAFF, ALTMAN ET PRISMA GROUP, 2009)



## RÉSULTATS

À partir de 1354 articles initialement identifiés à partir des trois bases de données, nous avons sélectionné 268 références (voir figure 1) qui ont été synthétisées sur une feuille Excel et analysées à l'aide des 12 ou 13 critères du MMAT.

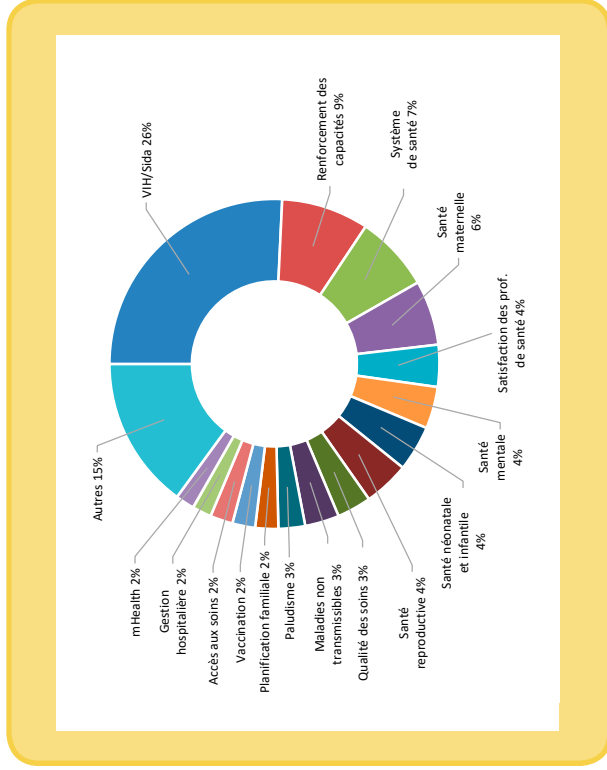
## Des articles représentatifs des tendances d'interventions de santé mondiale en Afrique subsaharienne

Malgré la baisse depuis 2010 des financements sur le VIH/Sida, plus d'un quart des articles employant les méthodes mixtes en Afrique subsaharienne (69/268, soit 26 %) portaient sur l'évaluation d'interventions de prévention ou de lutte contre le VIH/Sida et ses effets (voir figure 2). Les autres champs récurrents les plus fréquents étaient : le renforcement des capacités des professionnels de santé, les interventions visant l'amélioration du système de santé et les interventions de santé maternelle. Malgré une visibilité et des financements internationaux assez tardifs, la santé mentale est un thème qui revenait somme toute assez fréquemment dans les articles (11/268 soit un peu plus de 4 %) par rapport à d'autres thèmes plus « classiques » comme le paludisme et la planification familiale (2 à 3 %). Les programmes contre les maladies non transmissibles (maladies cardiovasculaires et cancer du col de l'utérus, principalement) sont relativement peu évalués, malgré le fardeau que peuvent représenter ces maladies et malgré la récente augmentation des financements pour prévenir et traiter ces problèmes de santé (Murray et Dieleman, 2013). L'évaluation de programmes innovants comme ceux en santé mobile (mHealth, 1,9 % des références sélectionnées) est aussi plus importante par rapport à l'évaluation d'interventions concernant d'« anciennes » maladies comme la tuberculose (1,5 %).

Concernant les pays les plus fréquemment ciblés par des évaluations d'interventions employant des méthodes mixtes, les pays anglophones sont surreprésentés par rapport aux pays francophones (à l'exception du Burkina Faso). Les pays attirant le plus de financement (Afrique du Sud, Ouganda, Tanzanie, Malawi, Kenya, Ghana, etc.) sont aussi ceux dans lesquels un nombre important d'évaluations utilisant les méthodes mixtes sont réalisées (figure 3) (OCDE, 2016).

Figure 2

THÈMES PRINCIPAUX  
DES 268 ARTICLES

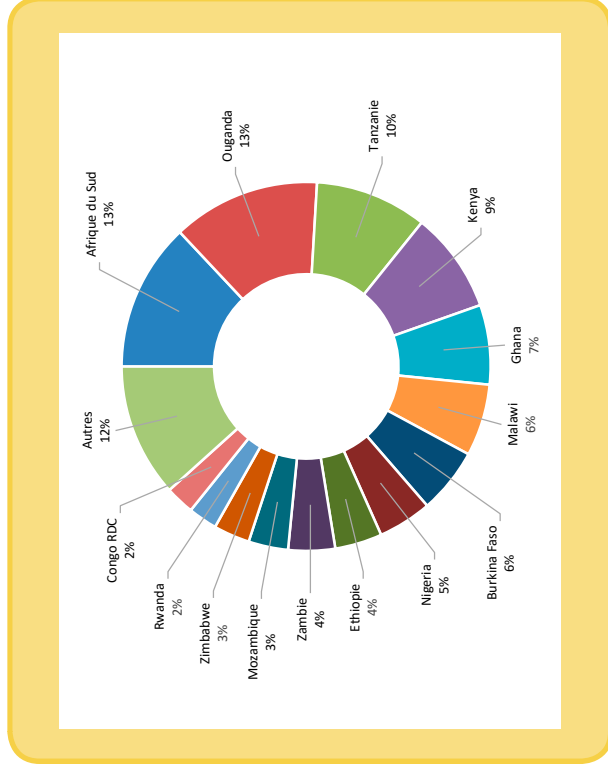


## Description et commentaire des méthodes quantitatives et qualitatives utilisées

Les méthodes utilisées pour les parties quantitatives de ces articles étaient en général peu sophistiquées : seulement 16 % (43/268) des études utilisaient des comparaisons de groupes (sélectionnés de façon randomisée ou non; catégories « III. i » et « III. ii »). Toutes les autres (225/268) entraient dans la catégorie « III. iii ». Seules 11 études (4 %) comportaient un devis d'essai contrôlé randomisé (application des quatre critères de la catégorie « III. i »), tandis que 32 (12 %) étaient des études cas-témoin, des études de cohortes ou des études transversales analytiques (application des quatre critères de la catégorie « III. ii »).

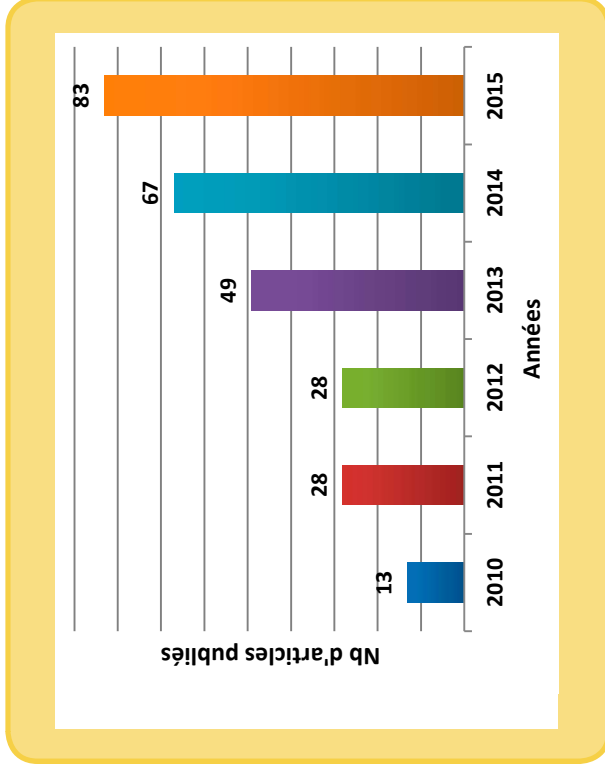
Figure 3

RÉPARTITION  
DES 268 ARTICLES PAR PAYS



Quant à l'évolution temporelle de l'usage des méthodes mixtes dans l'évaluation des interventions de santé mondiale en Afrique, elles suivent une courbe de croissance constante. Le nombre d'articles publiés a doublé entre 2010 et 2011, et augmenté entre 25 à 40 % entre 2012 et 2015 (figure 4). Il est à noter que près d'un tiers des articles sélectionnés dans cette étude ont été publiés en 2015 exclusivement. Sur la base de la croissance notée dans cette revue de littérature, il est à prévoir une augmentation d'au moins 25 % aussi pour 2016. Comme évoqué précédemment, ceci est probablement dû à l'accroissement de l'intérêt pour les méthodes mixtes de la part des revues à comité de pairs ces cinq dernières années.

**Figure 4**  
ÉVOLUTION DU NOMBRE  
D'ARTICLES PUBLIÉS PAR  
ANNÉE



Dans les 225 références auxquelles nous avons appliqué les critères de la catégorie « III. iii », les devis d'étude étaient extrêmement variés : séries de cas, rapports de cas individuels, études de prévalence, etc. Mises à part les 11 études contrôlées (incluant quelques essais contrôlés randomisés par grappes) où l'échantillon était aléatoire, les stratégies d'échantillonnage étaient pour la plupart intentionnelles (ou en fonction de la « disponibilité » des participants présents sur le lieu d'étude), souvent pour faire correspondre l'échantillon de la partie qualitative – ce qui donnait lieu à des limites reconnues comme importantes pour tout épidémiologiste ou statisticien : biais de sélection et taille de l'échantillon limitée. Les questionnaires quantitatifs employaient souvent des échelles psychométriques (par ex., échelle de Likert) qui étaient ensuite analysées à l'aide de modèles statistiques (par ex., *odd*

*ratios*) et l'utilisation de logiciels d'analyse statistique (par ex., SPSS). Dans les évaluations d'impact, des modèles de régression linéaire ou de régression multivariée étaient utilisés.

Pour ce qui est de la partie qualitative des articles sélectionnés, le recours aux entrevues individuelles semi-structurées y était largement majoritaire (77 %). Plus d'un tiers des articles (36 %) faisaient usage de discussions de groupe, en complément ou en remplacement des entrevues individuelles. Les autres méthodes de collecte de données qualitatives, par exemple l'observation participante ou non-participante, étaient rarement mentionnées. Une exception notable est l'article d'Irimu (2014) qui emploie non seulement un devis ethnographique, mais aussi une approche de recherche-action participative. Cette approche théorique se justifie tout à fait dans le cadre d'un devis mixte. Nous avons aussi à plusieurs reprises repéré l'usage de la théorie ancrée dans les articles de méthodes mixtes (par ex., Hodge et Roby, 2010; Juma, Edwards et Spitzer, 2014) et de la phénoménologie (Hawkes et autres, 2013), probablement parce que ces approches théoriques sont vraisemblablement revenues « à la mode » ces dernières années dans les milieux de recherche en santé publique américains<sup>1</sup>. Le mode d'échantillonnage était presque systématiquement intentionnel. En ce qui concerne le traitement et l'interprétation des données, on observe une surutilisation de l'analyse thématique et du *framework analysis*. Il s'agit dans la plupart des cas de parties qualitatives descriptives qui ne disent pas leur nom. Dans la grande majorité des cas, l'emploi d'un logiciel de type *computer-assisted qualitative data analysis software* (CAQDAS) comme NVivo ou Atlas.ti était mentionné (par ex., : Hardon et autres, 2012, 2013; Robyn et autres, 2014), parfois comme seule description de l'analyse – ce qui est problématique parce que la simple mention de ce type de logiciels ne permet pas d'expliquer la manière dont ils sont utilisés pour l'analyse. En effet, ces logiciels ne sont que des outils informatiques qui assistent et facilitent le codage, le regroupement de fragments de texte, etc.

1. Exemple des auteurs affiliés à l'Université Johns Hopkins dont les publications en santé sont répertoriées sur PubMed/Medline : depuis sept ans (depuis 2010), pas moins de 40 articles mentionnent l'usage de cette approche méthodologique, par rapport à seulement huit articles publiés au cours des sept années précédentes (2003-2009).

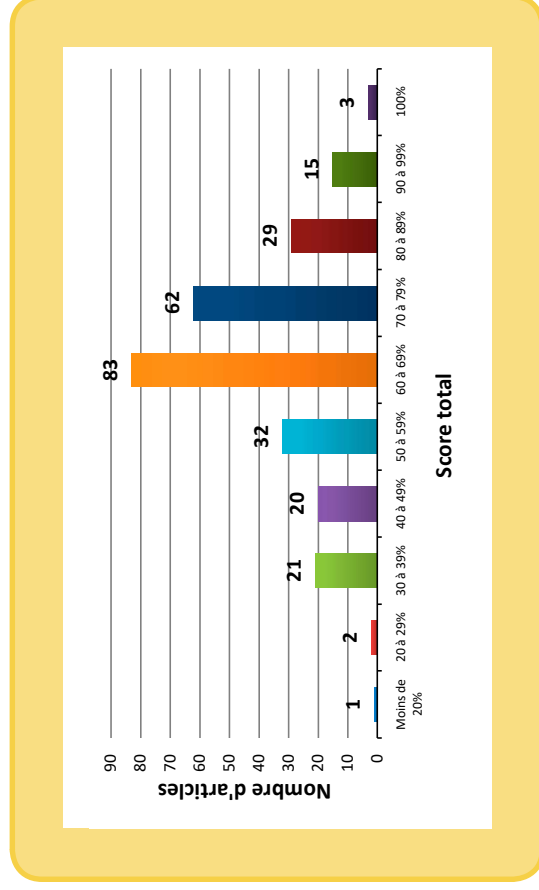


## Conformité aux critères du MMAT

Suite à notre évaluation, 72 % des articles (192/268) ont obtenu un score final supérieur à 60 % (voir figure 5), une valeur seuil pouvant être considérée comme acceptable. La plupart des articles se situaient en effet à des niveaux ni bons ni mauvais, mais acceptables, avec des scores oscillant entre 60 et 79 %. Les articles avec un score supérieur à 80 %, c'est-à-dire ceux qui remplissaient au moins 10/12 critères ou 11/13 critères, étaient peu nombreux : 18 % seulement (47/268). Le critère le moins fréquemment présent était celui portant sur la reconnaissance des difficultés et limites liées au processus d'intégration des aspects qualitatifs et quantitatifs qui est au cœur des méthodes mixtes : dans seulement 19 % des articles, les auteurs ont tenu compte de ce 13<sup>e</sup> critère.

Figure 5

### RÉPARTITION DES ARTICLES PAR SCORE TOTAL



Une analyse détaillée montre tout d'abord que les questions de recherche de méthodes mixtes ou les questions qualitatives et quantitatives n'étaient pas toujours claires. Quand elles ne l'étaient absolument pas et que les méthodes mixtes ne transparaissaient aucunement dans les objectifs, un score de 0 était attribué à cette partie considérée comme essentielle (7 % des articles). Dans certains articles, la formulation pouvait sembler inhabituelle : on mesurait avec un questionnaire la satisfaction de patients vis-à-vis de services octroyés par l'intervention, tandis qu'on explorait avec du qualitatif une relation causale. Souvent, faire du quantitatif n'avait aucun sens pour analyser un phénomène purement qualitatif (par ex., exploration des perceptions de la satisfaction des bénéficiaires vis-à-vis de services de soutien psychologique aux séropositifs). Enfin, certaines questions de recherche étaient clairement hypothético-déductives. Par exemple, un article proposait d'examiner les prédicteurs de la séroprévalence VIH dans une cohorte (qui recevait une intervention de conseil et dépistage du VIH) « avec une attention particulière pour l'affiliation religieuse comme conduit possible du risque VIH » (Hawkes et autres, 2013).

Le choix des méthodes mixtes n'était pas non plus explicite dans certains cas (4,5 % des articles). Compte tenu de l'approche employée dans ces articles, qui apparaissait tout de même correspondre à toutes les caractéristiques d'un devis mixte, nous avons décidé de continuer à évaluer les articles qui avaient un score de 0 à l'un ou l'autre de ces critères.

Seulement une minorité d'articles (moins d'un tiers d'entre eux) précisaient les devis spécifiques de méthodes mixtes (séquentiel exploratoire, séquentiel explicatif, devis concomitant, etc.). L'intégration des résultats avait rarement lieu : la plupart du temps (environ 75 % des articles), les résultats quantitatifs étaient présentés en premier et ils étaient suivis des résultats qualitatifs. L'intégration avait plus souvent lieu au stade de la discussion, et donc au moment de l'interprétation. Dans certains articles la discussion se trouvait scindée elle aussi en deux parties (Anastasi et autres, 2015).

## Manque de représentativité des résultats quantitatifs

Pour la partie quantitative (III), nous avons observé que de nombreux biais n'étaient pas rapportés par les auteurs. En premier lieu, le biais d'auto-administration était rarement mentionné (Bechange, 2010). En second lieu, les biais de désirabilité sociale (par ex., questionnaires administrés par les professionnels de santé à leurs propres patients) n'étaient pas pris en compte dans certains articles (Bansil et autres, 2014; Kahabuka, Kvale, Moland et Hinderaker, 2011) ou bien ils étaient mentionnés, mais de façon furtive (Chang et autres, 2013).

En ce qui concerne les études d'impact, les liens de causalité sont parfois établis sans tenir compte de facteurs de confusion ou de la possible colinéarité entre les variables explicatives. Par exemple, dans Hawkes et autres (2013), les auteurs ont mesuré l'association entre le statut (auto-rapporté) de catholique et la non-utilisation d'un préservatif (considéré comme facteur de risque de contamination au VIH). Or, il n'y avait aucun contrôle des facteurs de confusion, le raisonnement faisant alors penser au syllogisme suivant : « les participants séropositifs disent pour la plupart qu'ils sont chrétiens et vont à l'église catholique, donc il y a une association entre la religion catholique et la séropositivité » (Hawkes et autres, 2013).

Comme évoqué précédemment, les stratégies d'échantillonnage pouvaient apparaître discutables (dans 36 % des 225 articles dont la partie quantitative entrait dans la catégorie « III. iii » sans comparaison de groupes); la raison principale était que les stratégies d'échantillonnage étaient les mêmes pour les portions qualitative et quantitative de l'étude. Quant aux essais contrôlés randomisés, dans 36 % des articles la procédure de dissimulation des répartitions (*allocation concealment*) était peu détaillée. Dans les études de cohortes, les participants exposés à l'intervention et ceux non exposés n'étaient pas toujours recrutés dans la même population. Enfin, pour les études cas-témoins, le recrutement des participants n'était pas toujours réalisé indépendamment de l'intervention. Dans 34 % des articles de type « III. ii » (c'est-à-dire sans sélection aléatoire, mais avec comparaison intergroupes), quand les groupes comparés présentaient des caractéristiques différentes, les

auteurs ne tenaient pas compte de ces différences dans l'analyse, qui ne faisait pas le contrôle de ces variables.

Parmi les 225 articles, plus de la moitié (55 %) utilisaient un échantillon non représentatif des populations à l'étude. Les résultats étaient similaires pour les études censées être plus proches de l'étalement or des évaluations scientifiques (c'est-à-dire les essais contrôlés randomisés) : 38 % des articles ayant une partie quantitative de type « III. ii » présentaient le même problème. En outre, les échantillons étaient souvent petits (provenant de l'échantillonnage raisonné relié à la partie qualitative), ce qui limitait la portée et la puissance des résultats quantitatifs.

En revanche, la description des mesures était généralement bien expliquée. Dans les parties quantitatives de type « III. i » et « III. ii » des 43 articles utilisant ce type de méthodes, les variables étaient convenablement définies et les mesures étaient clairement exposées et adéquates par rapport à la question de recherche. Dans les parties de type « III. iii », plus de 80 % des cas présentaient des outils de mesure référencés, validés et bien détaillés. Les résultats étaient aussi le plus souvent complets (plus de 80 % de mesures complètes de résultats pour les parties de type « III. i » et « III. ii »), et on observait un taux de réponse supérieur à 60 % dans 68 % des parties de type « III. iii » pour lesquelles ce critère était applicable.

## Limites des parties qualitatives

Si les aspects quantitatifs des études mixtes incluses présentaient différentes lacunes, des limitations étaient surtout constatées pour les aspects qualitatifs. Plusieurs auteurs préféraient employer une méthode mixte, sans pour autant fournir d'explications quant à la collecte de données qualitatives ou rendre explicites les résultats qualitatifs (par ex., Coetzee et Kagee, 2013).

Dans 27 % des articles, nous avons estimé que les sources de données qualitatives n'étaient pas pertinentes. Notamment, les auteurs utilisent fréquemment les groupes de discussion pour explorer les normes sociales concernant des phénomènes culturellement sensibles comme le refus de la vaccination (Cockcroft et autres, 2014), la circoncision volontaire

(Hatzold et autres, 2014), et les services à domicile de conseil et dépistage des maladies sexuellement transmissibles (Jurgensen et autres, 2013). Il est habituellement recommandé de réaliser des entrevues individuelles dans ce type de recherche pour explorer les perspectives individuelles qui sont souvent hétérogènes et moins connues que les normes sociales. Dans le cas de Jurgensen et autres (2013), la réflexion critique des auteurs par rapport à l'hétérogénéité des réponses des participants au groupe de discussion apparaît limitée. Dans Hawkes et autres (2013), la méthode de collecte de données – des entretiens structurés – apparaît peu adaptée à l'approche phénoménologique choisie : la phénoménologie cherche à obtenir une compréhension approfondie d'une expérience vécue (Van Manen, 1944) alors qu'un entretien structuré permet difficilement d'atteindre cet objectif.

Dans la partie qualitative de 35 % des articles, il n'y avait pas de prise en compte suffisante du contexte et de la manière dont celui-ci affecte la collecte, les résultats et l'interprétation des données. Les développements portant sur la réflexivité du (ou des) chercheur(s) étaient davantage absents : dans près des trois quarts des articles (74 %) les auteurs ne rapportaient aucune démarche réflexive. Par opposition, certains articles (Ridde et Olivier de Sardan, 2015) faisaient la démonstration d'une forte réflexivité sur l'usage et l'apport des méthodes mixtes en santé mondiale. Contrairement à ce qui est attendu en recherche qualitative, il y avait également peu d'explication/justification de l'échantillonnage ou de la stratégie de recrutement des participants. À plusieurs reprises, la démonstration de crédibilité des résultats était l'insertion de la phrase « un chercheur qualitatif expérimenté a réalisé la collecte [ou l'analyse] des données », ce qui est nécessaire mais insuffisant. De façon générale, nous avons relevé que les critères visant à assurer la rigueur méthodologique en recherche qualitative étaient peu pris en compte par les auteurs des articles inclus. Mis à part l'emploi de certains types de triangulation (par ex., triangulation des résultats quantitatifs et qualitatifs, ou triangulation des sources de données provenant d'entretiens et de groupes de discussion), la crédibilité des résultats par rapport au contexte observé, et la confirmation auprès des participants, n'étaient que très rarement pris en compte ou mentionnés par les auteurs.

## DISCUSSION

Nos résultats montrent que la qualité méthodologique des études employant les méthodes mixtes en santé mondiale en Afrique subsaharienne est variable d'après ce qui est rapporté dans les articles concernant les aspects qualitatifs, quantitatifs et mixtes des études incluses dans notre revue. Du côté quantitatif, des liens de causalité sont parfois établis sans tenir compte de certains biais. La plupart des limitations sont constatées du côté qualitatif : pratiquement aucun article ne justifie l'échantillonnage ou la stratégie de recrutement des participants et peu de critères visant à assurer la rigueur méthodologique sont utilisés (Whittemore, Chase et Mandle, 2001). Ainsi, la crédibilité des résultats, la confirmation auprès des participants ou la réflexivité du chercheur sont souvent des critères de rigueur qui ne sont pas mentionnés.

En accord avec Pluye (2012), nous confirmons que les articles avec méthodes mixtes en santé mondiale font rarement référence à des travaux méthodologiques en santé mondiale. Les ouvrages méthodologiques les plus couramment cités sont ceux de Creswell (Creswell, Plano Clark, Gutmann et Hanson, 2003; Creswell, 2013), et dans une moindre mesure celui de Tashakkori et Teddlie (2003).

Globalement, les auteurs des articles avec méthodes mixtes documentent peu les informations correspondant aux critères de rigueur de la recherche qualitative. En revanche, les informations correspondant aux critères de rigueur des méthodes quantitatives sont souvent fournies; en particulier, les analyses étaient beaucoup mieux décrites que dans les parties qualitatives. Ceci semble traduire que la recherche qualitative peut être moins connue par les chercheurs en santé publique que les méthodes épidémiologiques et statistiques.

En outre, réaliser des évaluations en utilisant les méthodes mixtes peut s'avérer coûteux à la fois en temps et financièrement. Malgré le fort intérêt de telles études, le contexte de nombreux pays d'Afrique subsaharienne est tel qu'il peut être plus difficile de les mettre en œuvre par rapport aux pays à haut revenu, par manque de matériel, de ressources humaines ou de moyens financiers. En témoigne la surreprésentation de l'Afrique du Sud, seul pays de cette région ayant atteint un niveau de revenu intermédiaire de tranche supérieure.

Ce manque de ressources peut aussi expliquer les faibles scores de certaines études, que ce soit pour la section quantitative, la section qualitative, ou la section méthodes mixtes. L'accès aux données peut notamment s'avérer problématique dans le contexte des pays d'Afrique subsaharienne, dans lesquels l'accès physique et la confiance entre les participants et l'équipe de recherche font parfois défaut. En particulier, lorsqu'il s'agit de travailler avec des populations vulnérables ou rurales, l'accès aux données, tant qualitatives que quantitatives, peut être difficile. De ces défis résultent des choix d'échantillonnage (taille et source de données) qui peuvent apparaître comme limités, mais qui sont en réalité justifiables compte tenu du contexte de l'étude. Une possibilité pour être plus « juste » dans l'évaluation de la rigueur dans un tel contexte serait de noter les critères liés à ces dimensions (critères n°3 et n°7c) en tenant compte du cadre dans lequel se déroule la collecte (zone reculée, période post-conflit, populations vulnérables, etc.) et des ressources mises à disposition des chercheurs (projet financé localement, par opposition à un projet financé par des bailleurs étrangers avec de plus gros moyens). Bien sûr, cela implique que ces informations soient spécifiées dans les articles, ce qui n'est pas systématique. En cas de doute, s'en remettre à l'avis d'un second évaluateur pourrait aider à prendre une décision éclairée.

## La logique des méthodes mixtes vs celle des « méthodes multiples »

Pour la grande majorité des articles inclus dans cette revue, la logique même des méthodes mixtes n'est pas respectée et s'apparente plutôt à ce que Teddlie et Tashakkori (2006) appellent des méthodes quasi-mixtes. Celles-ci mobilisent des données qualitatives et quantitatives sans pour autant les intégrer (Teddlie et Tashakkori, 2006). La majorité des articles sélectionnés documentent simplement dans différentes sections les analyses qualitatives et quantitatives. À l'extrême, certaines études intitulées « mixtes » par leurs auteurs ont été exclues de cette revue de littérature puisqu'elles présentaient les résultats qualitatifs et quantitatifs dans différents articles. Différents points de vue s'affrontent toutefois quant au niveau d'intégration requis

pour l'appellation « méthode mixte » (Onwuegbuzie et Johnson, 2006). Certains auteurs préfèrent réserver l'appellation « méthode mixte » pour un niveau plus élevé d'intégration des méthodes qualitatives et quantitatives (Yin, 2006), tandis que d'autres, adoptant une approche plus utilitariste, mettent plutôt de l'avant la simple valeur ajoutée de l'ajout d'une autre méthode (ceci indépendamment de son intégration) (Onwuegbuzie et Leech, 2004). Le manque de consensus international sur ce niveau d'intégration, et les développements méthodologiques des méthodes mixtes, peuvent expliquer les disparités méthodologiques dans l'utilisation des méthodes mixtes; ces disparités allaient de la présentation des résultats qualitatifs et quantitatifs dans différents articles ou différentes sections, à une intégration de ces résultats en une seule analyse. Se spécialisant dans l'une ou l'autre des méthodologies de recherche (quantitative ou qualitative), certains chercheurs peuvent trouver difficile de sortir de leur cadre méthodologique pour adopter ce qui représenterait un troisième type de méthodologie de recherche, la méthodologie mixte.

## Le MMAT, un outil perfectible

Tout en étant un outil populaire pour analyser les études avec méthodes mixtes (23 000 visites Web entre 2013 et 2017), le MMAT est validé et testé pour sa fidélité, mais reste une « preuve de concept » perfectible. Différents problèmes peuvent être soulevés lors de l'utilisation du MMAT. Premièrement, la classification des études n'est pas exhaustive. On repère en effet différents types d'études non inclus dans les différentes catégories de critères : les études pré-post test et les régressions multivariées sans comparaison intergroupes, par exemple. Ce type d'études pourrait être ajouté dans la catégorie des études non randomisées (catégories « IV ») comme sous-section « sans groupe contrôle », les autres types d'études mentionnées nécessitant la comparaison avec un groupe contrôle. Un deuxième point à noter est que les divers critères ne sont pas toujours adéquatement décrits. Plusieurs critères s'avèrent très larges (ouvrant à une interprétation hétérogène) et nécessitent davantage de précision, par exemple les critères liés à la validité interne (catégories « III. ii » et « IV. iii »). Pour faciliter l'utilisation de l'outil, il serait pertinent de développer davantage la

définition de la validité. Cet exercice de précision représente toutefois un défi si l'on souhaite éviter de tomber dans le piège d'un critère trop détaillé et complexe (par ex., énumération de biais possibles : biais de l'enquêteur, biais d'auto-déclaration, etc.). Au contraire, plusieurs critères présentent aussi différents éléments à considérer en même temps, ce qui complexifie l'évaluation de ces derniers : par exemple, on demande dans le critère « III. ii » d'évaluer la clarté de la méthode de collecte de données, le type de données collectées, les changements concernant la collecte, et l'adéquation entre l'analyse et la question de recherche. L'accumulation d'information à l'intérieur du même critère complexifie son analyse puisque cette dernière consiste simplement à affirmer si le critère est présent ou absent. Une question se pose donc quant à une possible pondération des critères pour faciliter l'utilisation de l'outil. Une pondération permettrait de donner un score plus important à des critères perçus comme tels, ceci amenant toutefois un problème de subjectivité; or, cet aspect constitue le troisième problème rencontré lors de l'utilisation du MMAT. Plusieurs critères parlent de « clarté », de « pertinence », de « manière appropriée », etc., des termes pouvant facilement être interprétés de différentes manières selon les auteurs ou les utilisateurs de l'outil, ce qui rend ce dernier moins fiable (Souto et autres, 2015). L'évaluation des critères peut donc s'avérer très subjective, d'où l'importance d'un double examen/consensus durant/

avant l'évaluation des articles. La standardisation des résultats de l'outil devient donc difficile. Quatrièmement, peu d'accent est mis sur différents critères de rigueur, spécifiquement concernant les méthodes qualitatives : la crédibilité, la minutie, l'intégrité, l'authenticité, etc. (Whittemore et autres, 2001). Le MMAT préconise les critères de rigueur tels que la réflexivité, la transférabilité et la transparence procédurale, qui sont importants, sans toutefois représenter l'ensemble de ce que devrait être une bonne étude qualitative.

## CONCLUSION

En conclusion, deux conditions principales sont habituellement associées à la réussite d'une étude employant des méthodes mixtes : disposer d'un temps suffisamment long pour s'approprier ces méthodes et avoir une équipe de recherche transdisciplinaire composée de membres habitués à travailler ensemble (Pluye, 2012). Nos résultats suggèrent que les limitations des études ne remplissant pas ces deux conditions sont nombreuses. Par conséquence, il est important de développer, soutenir et promouvoir l'enseignement des méthodes mixtes à tous les niveaux, y compris dans la formation continue des chercheurs seniors en santé mondiale.

## Bibliographie

- Anastasi, E., M. Borchert, O. M. R. Campbell, E. Sondorp, F. Kaducu, O. Hill, D. Okeng, V.N. Odong et I. Lange (2015). « Losing women along the path to safe motherhood: why is there such a gap between women's use of antenatal care and skilled birth attendance? A mixed methods study in northern Uganda », *BMC Pregnancy and Childbirth*, vol. 15, n°287.
- Bansil, P., S., Wittet, J. L. Lim, J. L. Winkler, P. Paul et J. Jeronimo (2014). « Acceptability of self-collection sampling for HPV-DNA testing in low-resource settings: a mixed methods approach », *BMC Public Health*, vol. 14, n° 596.
- Bechange, S. (2010). « Determinants of project success among HIV/AIDS non-governmental organizations (NGOs) in Rakai, Uganda », *The International Journal of Health Planning and Management*, vol. 25, n°3, p. 215-230.
- Chang, L. W., V. Njie-Carr, S. Kalenge, J. F. Kelly, R.C. Bollinger et S. Alamo-Talisuna (2013). « Perceptions and acceptability of mHealth interventions for improving patient care at a community-based HIV/AIDS clinic in Uganda: A mixed methods study », *AIDS Care*, vol. 25, n° 7, p. 874-880.
- Cockcroft, A., M. U. Usman, O. F. Nyamucherera, H. Emori, B. Duke, N. A. Umar et N. Andersson (2014). « Why children are not vaccinated against measles: a cross-sectional study in two Nigerian States », *Archives of Public Health = Archives Belges de Santé Publique*, vol. 72, n° 1, p. 48-48.
- Coetzee, B. et A. Kagee (2013). « The Development of an Inventory to Assess the Structural Barriers to Clinic Attendance and Pill-taking Amongst Users of Antiretroviral Therapy », *Aids and Behavior*, vol. 17, n°1, p. 319-328.

- Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*, Thousand Oaks, SAGE publications.
- Creswell, J. W., V. L. Plano Clark, M. L. Gutmann, et W. E. Hanson (2003). « Advanced mixed methods research designs », A. Tashakkori et C. Teddlie, *Handbook of Mixed Methods in Social and Behavioral Research*, Thousand Oaks, SAGE publications, p. 209-240.
- Guével, M.-R. et J. Pommier (2012). « Recherche par les méthodes mixtes en santé publique: enjeux et illustration », *Santé Publique*, vol. 24, n° 1, p. 23-38.
- Hardon, A., G. B. Gomez, E. Vernooij, A. Desclaux, R. K. Wanyenze, O. Ky-Zerbo, E. Kageha, I. Namakhoma, J. Kinsman, C. Spronk, E. Meij, M. Neuman et C. M. Obermeyer (2013). « Do support groups members disclose less to their partners? The dynamics of HIV disclosure in four African countries », *Bmc Public Health*, vol. 13, n° 589.
- Hardon, A., E. Vernooij, G. Bongololo-Mbera, P. Cherutich, A. Desclaux, D. Kyaddondo, O. Ky-Zerbo, M. Neuman, R. K. Wanyenze et C. M. Obermeyer (2012). « Women's views on consent, counseling and confidentiality in PMTCT: a mixed-methods study in four African countries », *Bmc Public Health*, vol. 12, n° 26.
- Hatzold, K., W. Mavhu, P. Jasi, K. Chatora, F. M. Cowan, N. Taruberekera, O. Mugurungi, K. Ahanda et E. Njeuhmeli (2014). « Barriers and Motivators to Voluntary Medical Male Circumcision Uptake among Different Age Groups of Men in Zimbabwe: Results from a Mixed Methods Study », *Plos One*, vol. 9, n° 5, p. e85051.
- Hawkes, M., E. S. Sivasivugha, S. K. Ngigi, C. K. Masumbuko, J. Brophy et Z. T. Kibendelwa (2013). « HIV and religion in the Congo: a mixed-methods study », *Current HIV Research*, vol. 11, n° 3, p. 246-253.
- Hodge, D. R. et J. L. Roby (2010). « Sub-Saharan African women living with HIV/AIDS: an exploration of general and spiritual coping strategies », *Social Work*, vol. 55, n° 1, p. 27-37.
- Irimu, G. W., A. Greene, D. Gathara, H. Kihara, C. Maina, D. Mbori-Ngacha, D., D. Zurovac, M. Santau, J. Todd et M. English (2014). « Explaining the uptake of paediatric guidelines in a Kenyan tertiary hospital--mixed methods research », *BMC Health Services Research*, vol. 14, n° 119.
- Johnson, R. B. et A. J. Onwuegbuzie (2004). « Mixed methods research: A research paradigm whose time has come », *Educational Researcher*, vol. 33, n° 7, p. 14-26.
- Juma, P. A., N. Edwards et D. Spitzer (2014). « Kenyan nurses involvement in national policy development processes », *Nursing Research and Practice*, n° 236573.
- Jurgensen, M., I. F., Sandoy, C. Michelo, K. Fylkesnes, S. Mwangala et A. Blystad (2013). « The seven Cs of the high acceptability of home-based VCT: Results from a mixed methods approach in Zambia », *Social Science & Medicine*, vol. 97, p. 210-219.
- Kahabuka, C., G. Kvale, K. M. Moland, et S. G. Hinderaker (2011). « Why care-takers bypass Primary Health Care facilities for child care - a case from rural Tanzania », *BMC Health Services Research*, vol. 11, n° 315.
- Lehoux, P. (2015). *Séance 14: Devis mixtes, avenir de la recherche qualitative et bilan du cours*. Présentation du cours d'Introduction à la recherche qualitative, Montréal, École de santé publique de l'Université de Montréal.
- Moher, D., A. Liberati, J. Tetzlaff, D. G. Altman et PRISMA Group (2009). « Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement », *Annals of Internal Medicine*, vol. 151, n° 4, p. 264-269.
- Murray, C. J. L. et J. L. Dieleman (2013). *Global health funding reaches new high as funding priorities shift*. Seattle, Institute for Health Metrics and Evaluation.
- OCDE. (2016). *Development aid at a glance. Statistics by region. 2. Africa*. Paris, Organisation de Coopération et de Développement Economique.
- Onwuegbuzie, A. J. et R. B. Johnson (2006). « The validity issue in mixed research », *Research in the Schools*, vol. 13, n° 1, p. 48-63.
- Onwuegbuzie, A. J. et N. L. Leech (2004). « Enhancing the interpretation of significant findings: The role of mixed methods research », *The Qualitative Report*, vol. 9, n° 4, p. 770-792.
- Pluye, P. (2012). « Méthodes mixtes et santé mondiale », J. C. Suárez-Herrera et M. J. Blain, *La recherche en santé mondiale : Perspectives socio-anthropologiques*, Québec, Association francophone pour le savoir, p. 129-35.
- Pluye, P., E. Robert, M. Cargo, A. O' Cathain, F. Griffiths, F. Boardman, M. P. Gagnon et M. C. Rouseau (2011). *Proposal: A mixed methods appraisal tool for systematic mixed studies reviews*, Montréal, McGill University, Department of Family Medicine.
- Ridde, V. et J.-P. Olivier de Sardan (2015). « A mixed methods contribution to the study of health public policies: complementarities and difficulties », *Bmc Health Services Research*, vol. 15, 57.
- Robert, E. (2015). *Mixed Method Appraisal Tool pour l'évaluation de la qualité des études de la revue réaliste*. Montréal, Université de Montréal.

- Robyn, P. J., T. Baernighausen, A. Souares, A. Traore, B. Bicaba, A. Sie et R. Sauerborn (2014). « Provider payment methods and health worker motivation in community-based health insurance: A mixed-methods study », *Social Science & Medicine*, vol. 108, p. 223-236.
- Souto, R. Q., V. Khanassov, Q. N. Hong, P. L. Bush, I. Vedel et P. Pluye (2015). « Systematic mixed studies reviews: Updating results on the reliability and efficiency of the mixed methods appraisal tool », *International Journal of Nursing Studies*, vol. 52, n° 1, p. 500-501.
- Suárez-Herrera, J. C. et M.-J. Blain (2012). *La recherche en santé mondiale: perspectives socio-anthropologiques*. Québec, Association francophone pour le savoir.
- Tashakkori, A. et C. Teddlie (2003). *Handbook of Mixed Methods in Social & Behavioral Research*, Thousands Oaks, SAGE publications.
- Teddlie, C. et A. Tashakkori (2006). « A general typology of research designs featuring mixed methods », *Research in the Schools*, vol. 13 n°1, p. 12-28.
- Van Manen, M. (1944). « Practicing phenomenological writing », *Phenomenology + Pedagogy*, vol. 2, n° 1, p. 36-69.
- Whittemore, R., S. K. Chase, et C. L. Mandle (2001). « Validity in qualitative research », *Qualitative Health Research*, vol. 11, n° 4, p. 522-537.
- Yin, R. K. (2006). « Mixed methods research: Are the methods genuinely integrated or merely parallel? », *Research in the Schools*, vol. 13, n°1, p. 41-47.