



HAL
open science

**Les enjeux de l'usage des méthodes quantitatives dans
l'évaluation d'impact d'interventions en promotion de la
santé : le cas d'une intervention pour améliorer
l'empowerment en santé sexuelle d'immigrés originaires
d'Afrique subsaharienne en France**

Andrainolo Ravalihasy

► **To cite this version:**

Andrainolo Ravalihasy. Les enjeux de l'usage des méthodes quantitatives dans l'évaluation d'impact d'interventions en promotion de la santé : le cas d'une intervention pour améliorer l'empowerment en santé sexuelle d'immigrés originaires d'Afrique subsaharienne en France. Médecine humaine et pathologie. Université Paris Cité, 2023. Français. NNT : 2023UNIP5274 . tel-04760531

HAL Id: tel-04760531

<https://theses.hal.science/tel-04760531v1>

Submitted on 30 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Paris Cité

**École doctorale Pierre Louis de santé publique : épidémiologie et
sciences de l'information biomédicale (ED 393)**

Infection, Anti-microbien, Modélisation, Évolution (UMR-S 1137)

**Les enjeux de l'usage des méthodes quantitatives dans
l'évaluation d'impact d'interventions en promotion de la
santé :**

*le cas d'une intervention pour améliorer l'empowerment en santé sexuelle
d'immigrés originaires d'Afrique subsaharienne en France*

Par Andrainolo RAVALIHASY

Thèse de doctorat en Prévention et promotion de la santé

Dirigée par Yazdan YAZDANPANAHAH
et co-dirigée par Valéry RIDDE

Présentée et soutenue publiquement le 13/12/2023

Composition du jury

Jade GHOSN, Professeur des Universités, Université Paris Cité	Président du jury
Perrine ROUX, Directrice de Recherche, Université Aix-Marseille	Rapporteuse
Laura TEMIME, Professeure des Universités, CNAM	Rapporteuse
Erica DI RUGGIERO, Professeure Associée, Université de Toronto	Examinatrice
Kévin JEAN, Maître de Conférence des Universités, CNAM	Examinateur
Yazdan YAZDANPANAHAH, Professeur des Universités, Université Paris Cité	Directeur de thèse
Valéry RIDDE, Directeur de Recherche, Université Paris Cité	Co-directeur de thèse
Lidia KARDAŚ-SŁOMA, Docteure, Université Paris Cité	Membre invité

Laboratoires d'accueil

Cette thèse a été réalisée au sein de l'équipe DeSCID (Decision Sciences in Infectious Disease Prevention, Control and Care) de l'unité de recherche IAME (Infection, Antimicrobials, Modelling, Evolution) et de l'équipe SageSud (Santé, vulnérabilités et relations de genre au Sud) de l'unité de recherche Ceped (Centre Population et Développement).



Remerciements

La réalisation de cette thèse a été impulsée par la fébrilité liée à l'inconnu, a été alimentée par la passion liée à la découverte et soutenue par l'effort nécessaire pour mener son terme. Du début à la fin, plusieurs personnes ont concouru de près ou de loin à sa réalisation, leur témoigner ma gratitude est la moindre des choses.

Je remercie mes directeurs de thèse, Yazdan Yazdanpanah et Valéry Ridde, d'avoir accepté d'encadrer cet étudiant silencieux, hésitant et d'avoir contribué ainsi à me (re)construire en tant que chercheur. Votre bienveillance et l'opportunité d'apprendre de vos travaux et de vos champs complémentaires m'auront permis de penser et repenser cette thèse de manière à (essayer d') en tirer le meilleur. Yazdan, tu as été présent à des moments cruciaux et déterminants. Valéry, tu as aiguisé mon sens critique et tu as su m'aiguiller tout en me laissant la liberté nécessaire pour laisser mon travail s'exprimer. Je joins bien sûr à ces remerciements Lidia Kardaś-Słoma qui a accompagné de très près chaque pas que je faisais dans la réalisation de cette thèse. Ton calme, ta patience et ta compréhension m'ont permis d'avancer sereinement dans mes travaux.

Je suis également reconnaissant envers les membres du jury pour leur disponibilité et pour avoir accepté d'évaluer ce travail. Je remercie Perrine Roux et Laura Temime d'avoir accepté d'être rapporteuses de cette thèse et apporter leurs regards complémentaires. Je remercie également Erica Di Ruggiero, Kévin Jean et Jade Ghosn d'avoir accepté d'examiner ce travail.

Je tiens à remercier Annabel Desgrées du Loû. Annabel, merci de ton soutien depuis mes premiers pas dans la recherche. Je ne saurais trouver les mots pour te remercier de toutes les opportunités que tu m'as offertes en m'intégrant à plusieurs de tes projets de recherche : du fond du cœur, merci !

Je remercie tout particulièrement Rigas Arvanitis, directeur du Ceped, qui m'a extirpé de mon bureau pour réaliser cette thèse dans les meilleures conditions possible grâce à sa bienveillance. Cette bienveillance me rappelle bien sûr l'environnement de travail dont je dispose au Ceped. Cela fait maintenant dix ans que les rencontres scientifiques, amicales et humaines qui y ont lieu nourrissent mes réflexions et ont inspiré mon envie de travailler dans le monde de la recherche.

En parlant de laboratoire et de bienveillance, l'équipe DeSCID du laboratoire IAME a été particulièrement tolérant envers mon statut de doctorant fantôme et a su comprendre la difficulté de concilier mes fonctions au sein du Ceped et de doctorant. Mes pensées vont particulièrement à Sylvie et Anthony, répondant toujours présents à mes sollicitations. C'est un réel plaisir de travailler avec vous et je souhaite vivement poursuivre cette collaboration.

Tellement de personnes m'ont aidé et contribué à ce travail. Nathalie, Pearl, Manuella, Isabelle, Emilie, Laurence, Annie ; j'ai bénéficié de votre expertise et encore mieux, de votre temps pour réaliser mes travaux. Merci !

Grâce au projet MAKASI, j'ai rencontré tant de personnes qui m'ont aidé à nourrir mes réflexions. Une mention spéciale à Karna : tu sais à quel point nos échanges formels et informels ont pu nourrir ce travail. De même, Anne, Julia, Iris, Séverine, Corinne, Veroska, Ruth, Patricia, Mounia, Qays, Flore, Jean-Noël, Maria et toutes celles et ceux qui ont contribué à ce projet, et donc indirectement à cette thèse : encore merci.

Mes premiers pas dans la recherche ont commencé grâce au projet Parcours. C'était une belle expérience où j'ai pu bénéficier de critiques, de défis, d'encadrement, d'encouragements et d'inspirations grâce aux personnes que j'y ai rencontré. Vous vous reconnaitrez. Merci à tous.

A toutes les personnes envers qui j'ai pu me permettre d'être spontané, honnête et confier à quel point il pouvait être compliqué de concilier travail, thèse et vie de famille : vos éclats de rire, vos blagues, nos sarcasmes ont alimenté mon courage.

Ceci est l'aboutissement de tout un parcours, je remercie toutes les personnes qui m'ont aidé à tracer ce chemin, je ne pourrai pas tous vous citer mais il est certain que je vous suis reconnaissant.

Merci à ma famille depuis tout ce temps, aujourd'hui et encore demain : Mamako, Tombo, Saina, Mihaja, Tialy, Kanty, Ingrid, Donia, et ma Poupoune Alexa.

Mamako, encore une fois : merci d'avoir été stricte et sévère quand il le fallait et même quand il ne le fallait pas.

Une pensée pour Papa, qui reste un exemple d'effort, de discipline et de réussite.

Titre : Les enjeux de l'usage des méthodes quantitatives dans l'évaluation d'impact d'interventions en promotion de la santé : le cas d'une intervention pour améliorer l'empowerment en santé sexuelle d'immigrés originaires d'Afrique subsaharienne en France

Résumé :

Les programmes élaborés dans le champ de la promotion de la santé requièrent des méthodes d'évaluation permettant de prendre en compte le contexte, les besoins et la participation des acteurs, y compris les communautés premières concernées par ces programmes. Ainsi, la pertinence des méthodes quantitatives telles que mobilisées actuellement est souvent remise en question par le fait qu'elles sont plus souvent utilisées pour constater les impacts des programmes de santé et non les comprendre. Selon leur utilisation, ces méthodes peuvent pourtant être mises à profit pour prendre en compte des éléments du contexte de mise en œuvre de ces programmes.

Cette thèse porte sur l'usage des méthodes quantitatives dans l'évaluation d'impact des interventions en promotion de la santé en prenant pour étude de cas l'intervention MAKASI, une intervention proposant d'améliorer l'*empowerment* en santé sexuelle chez des immigrants originaires d'Afrique subsaharienne vivant en France. Cette thèse propose dans un premier temps un recensement critique des méthodes utilisées pour évaluer l'impact des interventions dans ce type de contexte. Dans un second temps, elle décrit les méthodes (indicateurs et méthodes d'estimation) et les choix qui ont été faits dans l'usage de ces méthodes dans le cas de l'intervention MAKASI. Elle remet ensuite en perspective le rôle de l'évaluateur vis-à-vis de ces choix et du contexte de l'intervention. Ce travail montre les enjeux propres à l'usage des indicateurs et les enjeux propres à l'usage des méthodes d'estimation pour prendre en compte des éléments de contexte de l'intervention, notamment les inégalités sociales et structurelles de santé. Enfin, cette thèse montre l'importance de la participation et l'engagement de l'évaluateur dans les dispositifs prévus dans les programmes de santé pour orienter les choix d'évaluation. Cet engagement est un enjeu supplémentaire se distinguant des débats en faveur et en défaveur de l'usage des méthodes quantitatives telles que mobilisées actuellement dans l'évaluation des interventions en promotion de la santé, pour montrer dans quelles conditions cet usage est pertinent.

Mots clefs : Promotion de la santé, Evaluation d'impact, Méthodes quantitatives, Recherche interventionnelle, Empowerment

Title: The challenges of using quantitative methods to assess the impact of health promotion interventions: a case study based on an intervention to improve sexual health empowerment among immigrants from sub-Saharan Africa in France.

Abstract:

Health promotion programs require evaluation methods that take into account the context, the needs and the participation of stakeholders and communities concerned by these programs. Thus, the relevance of quantitative methods, as currently used, is often challenged due to the fact that they are more often used to observe the impact of health programs than to understand them. Depending on how they are applied, these methods can be used to take into account the context in which these programs are implemented.

This thesis focuses on the use of quantitative methods to assess the impact of health promotion interventions by taking the MAKASI intervention, which aims to improve sexual health empowerment among immigrants from sub-Saharan Africa living in France, as a case study. Firstly, a methodological review was conducted to map the methods used to evaluate the impact of such interventions. Secondly, the methods (indicators and estimation methods) and the choices made in the use of these methods in the MAKASI intervention are described. Then, a reflexive analysis puts into perspective the role of the evaluator in relation to these choices and the context of the intervention. This thesis shows the challenges inherent in the use of indicators and the challenges inherent in the use of estimation methods when accounting for intervention contexts, especially social and structural health inequalities. Finally, this thesis shows the extent to which the evaluator's participation and commitment in health programs activities can shape evaluation choices. The evaluator's commitment is an issue that contributes to highlight the circumstances under which the use of quantitative methods is relevant for health promotion interventions evaluation.

Keywords: Health promotion, Impact evaluation, Quantitative methods, Interventional studies, Empowerment

Table des matières

Table des matières.....	8
Liste des sigles et abréviations.....	12
Liste des illustrations	13
Liste des figures.....	13
Liste des équations	13
Productions scientifiques dans le cadre de la thèse.....	14
Articles dans des revues à comité de lecture en premier auteur.....	14
Chapitre de livre	14
Communications orales	15
Communications affichées	15
Productions scientifiques non liées à la thèse	16
Introduction.....	19
Des programmes complexes et des évaluations complexes : un impératif pour réduire l'épidémie de VIH.....	22
Objectifs et structure de la thèse.....	23
Partie – I : Etat de l'art sur l'évaluation d'impact.....	26
Chapitre – 1 : L'usage des méthodes quantitatives dans l'évaluation d'impact des interventions en santé	27
1. L'émergence de l'expérimentation aléatoire comme méthode de référence de l'évaluation d'impact	27

2.	La logique contrefactuelle et les plans d'étude d'évaluation d'impact.....	31
a.	La causalité contrefactuelle n'est pas que statistique.....	32
b.	Les plans quasi-expérimentaux	36
3.	Les enjeux de validité des résultats au-delà des enjeux statistiques	39
4.	D'autres modèles causaux pour l'évaluation des interventions complexes ?.....	42
5.	Conclusion	49
Chapitre – 2 : Les enjeux d'évaluation dans la lutte contre l'épidémie de VIH.....		51
1.	Vers des programmes de plus en plus complexes.....	51
a.	Les approches biomédicales de la prévention combinée	53
b.	Les approches comportementales de la prévention combinée	54
c.	Les approches structurelles de la prévention combinée	54
d.	Combiner les approches biomédicales, comportementales et structurelles	55
2.	Recommandations en termes d'évaluation	56
3.	Le projet MAKASI : un exemple concret de programme de prévention combinée du VIH58	
4.	Hypothèses et objectifs de la recherche doctorale	58
Partie – II : Revue critique des méthodes utilisées pour évaluer l'impact des programmes de prévention combinée sur la propagation du VIH		61
Chapitre – 3 : La difficulté et la nécessité de comparer les différentes méthodes quantitatives mobilisées pour évaluer l'impact des programmes de prévention combinée du VIH		62
Chapitre – 4 : Quelles méthodes quantitatives pour évaluer l'impact des interventions dans la lutte contre le VIH ?.....		65

Partie – III : Étude de cas de la mise en œuvre de l'évaluation dans le cas de l'intervention MAKASI.....	68
Chapitre – 5 : L'étude MAKASI.....	69
1. Contexte de l'étude.....	69
2. Conception et mise en œuvre de l'intervention.....	69
a. Théorie du changement : les composantes actives de l'intervention.....	69
b. Théorie de la mise en œuvre de l'intervention	71
3. Matériels et méthodes.....	71
a. Population d'étude	71
b. Plan d'étude d'évaluation d'impact	72
c. Données récoltées	74
d. Analyses réalisées dans le cadre de la thèse	75
Chapitre – 6 : Les enjeux des indicateurs de mesure de l'impact de l'intervention.....	76
Chapitre – 7 : Les enjeux des méthodes d'estimation d'impact de l'intervention	79
Chapitre – 8 : Analyse réflexive de la mise en œuvre des méthodes quantitatives dans le cadre du projet MAKASI	83
Partie – IV : Discussion générale.....	86
Apports identitaires des évaluateurs.....	88
Apports intercommunautaires	89
Apports interdisciplinaires.....	90
Place des enjeux des méthodes dans les courants de l'évaluation	92
Le cas particulier de la modélisation prospective.....	94

Application à la promotion de la santé	96
Conclusion.....	98
Bibliographie	101
Annexes	115
Annexe 1	116
Annexe 2.....	124
Annexe 3	143
Annexe 4.....	155
Annexe 5	184
Liste des éléments sous droits	212

Liste des sigles et abréviations

ANRS : Agence Nationale de Recherche sur le Sida et les hépatites virales

CONSORT : CONSolidated Standards Of Reporting Trails

EISS : Empowerment en santé sexuelle

IASS : Immigré originaire d’Afrique subsaharienne

ICH : International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use / Conseil international d'harmonisation des exigences techniques pour l'enregistrement des médicaments à usage humain

MAKASI : Mobilisation des Afro Karibéens pour l’Appui à la Santé et à l’Installation

OMS : Organisation mondiale de la Santé

PEP : Prophylaxie post exposition

PPCV : Programme de prévention combinée du VIH

PrEP : Prophylaxie pré exposition

SIDA : Syndrome d’immunodéficience acquise

SUTVA : Stable Unit Treatment Value Assumption

VIH : Virus de l’immunodéficience humaine

Liste des illustrations

Liste des figures

Figure 1 : Illustration du modèle à résultats potentiels de Rubin	34
Figure 2 : Illustration d'une approche orientée par la théorie de l'intervention basée sur l'approche dite « intervening mechanism approach »	45
Figure 3 : Illustration d'une approche orientée par la théorie de l'intervention basée sur l'approche dite « moderating mechanism approach »	46
Figure 4 : Illustration d'une approche orientée par la théorie de l'intervention basée sur l'approche dite « Integrative Process/Outcome Evaluation »	47
Figure 5 : Description des approches combinées de la prévention du VIH	53
Figure 6 : Illustration de l'utilisation de l'approche campbellienne dans les programmes de prévention	57
Figure 7 : Théorie du changement de l'intervention MAKASI	70
Figure 8 : Théorie de la mise en œuvre de l'intervention MAKASI	71
Figure 9 : Schéma expérimental du projet MAKASI	73
Figure 10 : Diagramme de flux de la participation à l'étude MAKASI	74
Figure 11 : Evolution du score moyen global et par dimension d'EISS chez les participants au cours du suivi dans l'intervention	78

Liste des équations

Équation 1 : Illustration de l'estimation de l'effet d'un dispositif dans le cadre du modèle à effets potentiels	35
------------------------------------------------------------------------------------------------------------------------	----

Productions scientifiques dans le cadre de la thèse

Articles dans des revues à comité de lecture en premier auteur

Ravalihasy, A., Rude, N., Yazdanpanah, Y., Kardas-Sloma, L., Desgrées du Loû, A., Gosselin, A., & Ridde, V. (2021). Development and Validation of an HIV/AIDS Empowerment Scale for Impact Intervention Evaluation. An Example from the MAKASI Intervention. *American Journal of Health Education*, 52(5), 296-306. <https://doi.org/10.1080/19325037.2021.1955230>

Ravalihasy, A., Kardaś-Słoma, L., Yazdanpanah, Y., & Ridde, V. (2022). Quantitative methods used to evaluate impact of health promotion interventions to prevent HIV infections : A methodological systematic review protocol. *Systematic Reviews*, 11(1), 87. <https://doi.org/10.1186/s13643-022-01970-z>

Ravalihasy, A., Ante-Testard, P. A., Kardas-Sloma, L., Yazdanpanah, Y., De Allegri, M., & Ridde, V. (2023). Quantitative Methods Used to Evaluate Impact of Combination HIV Prevention Intervention : A Methodological Systematic Review. *AIDS and Behavior*. <https://doi.org/10.1007/s10461-023-04000-8>

Ravalihasy, A. L'évaluateur et l'évaluation d'impact d'une intervention comme levier potentiel de l'empowerment au niveau communautaire : un exercice réflexif à partir de l'intervention MAKASI [en révision]

Ravalihasy, A. Modélisation de l'impact d'une intervention de prévention du VIH basée sur la réduction des vulnérabilités et des inégalités sociales sur le risque de transmission chez les immigrants originaires d'Afrique subsaharienne vivant en Île-de-France [en préparation]

Chapitre de livre

Ravalihasy, A., Coulibaly, K., & Bousmah, M.-Q. (2023). La mesure de l'empowerment en santé sexuelle : Réflexion critique sur les indicateurs. In A. Desgrées Du Loû & A. Gosselin, *Vers l'empowerment en santé : Recherches communautaires autour du projet Makasi* (p. 35-44). ANRS | MIE.

Communications orales

Coulibaly, K., **Ravalihasy, A.** Comment mesurer l’empowerment en santé sexuelle des immigrés précaires ? Le cas de Makasi. Séminaire de l’Institut Convergence Migrations – Département Santé, Aubervilliers, France, Janvier 2021

Ravalihasy, A. Mesurer l’empowerment pour remettre la participation et les besoins des communautés au cœur de l’évaluation des programmes de santé. 24ème Conférence mondiale de l’UIPES sur la Promotion de la Santé, Montréal, Québec, Canada, Mai 2022

Ravalihasy, A. Converting implementation issues into quantitative findings to challenge complex interventions evaluation: insights from a systematic review. AIDSImpact, Stockholm, Suède, Juin 2023

Communications affichées

Ravalihasy, A. Décider, interagir, prendre conscience face au VIH : une échelle de mesure de l’empowerment en santé sexuelle chez les immigrés originaires d’Afrique subsaharienne vivant en France. Afraviv (10e conférence internationale francophone VIH/Hépatites), Dakar, Sénégal, Novembre 2020

Ravalihasy, A. Une échelle de mesure de l’empowerment en santé sexuelle : un outil adapté et adaptable aux interventions et intervenants auprès des immigrés originaires d’Afrique subsaharienne exposés au VIH. Afraviv (11e conférence internationale francophone VIH/Hépatites), Marseille, France, Avril 2022

Ravalihasy, A. Structural interventions as game-changer for vulnerable populations? : insights from a case modeling study. AIDSImpact, Stockholm, Suède, Juin 2023

Productions scientifiques non liées à la thèse

- Audibert, M., N'Landu, A., Ravit, M., Raffalli, B., **Ravalihasy, A.**, Ridde, V., & Dumont, A. (2020). Forfait obstétrical et inégalités dans l'accès aux soins maternels en Mauritanie. *Revue économique*, 71(6), 1045-1067.
- Coulibaly, K., Bousmah, M.-Q., **Ravalihasy, A.**, Taéron, C., Mbiribindi, R., Senne, J.-N., Gubert, F., Gosselin, A., & Desgrées du Loû, A. (2023). Bridging the knowledge gap of biomedical HIV prevention tools among sub-saharan african immigrants in France. Results from an empowerment-based intervention. *SSM - Population Health*, 23, 101468. <https://doi.org/10.1016/j.ssmph.2023.101468>
- Coulibaly, K., Gosselin, A., Carillon, S., **Ravalihasy, A.**, Melchior, M., Ridde, V., Desgrées du Loû, A. (2022). Is empowerment in sexual health measurable? A scoping review of definitions and measurement indicators. *Health Promotion International*, 37(5), daac139. <https://doi.org/10.1093/heapro/daac139>
- Desgrées du Loû, A., Coulibaly, K., Zoumenou, I., Gosselin, A., Carillon, S., **Ravalihasy, A.**, Eid, J. (2023). La participation sociale, levier d'empowerment pour les immigrés précaires. *Revue Européenne des migrations Internationales*. [à paraître]
- Gosselin, A., **Ravalihasy, A.**, Pannetier, J., Lert, F., Desgrées du Loû, A. (2020). When and why? Timing of post-migration HIV acquisition among sub-Saharan migrants in France. *SEXUALLY TRANSMITTED INFECTIONS*, 96(3), 227-231. <https://doi.org/10.1136/sextrans-2019-054080>
- Pilecco, F. B., Guillaume, A., **Ravalihasy, A.**, Desgrées du Loû, A (2020). Induced Abortion and Migration to Metropolitan Paris by Sub-Saharan African Women : The Role of Intendedness of Pregnancy. *JOURNAL OF IMMIGRANT AND MINORITY HEALTH*, 22(4), 682-690. <https://doi.org/10.1007/s10903-019-00956-9>
- Ravit, M., **Ravalihasy, A.**, Audibert, M., Ridde, V., Bonnet, E., Raffalli, B., Roy, F.-A., N'Landu, A., & Dumont, A. (2020). The impact of the obstetrical risk insurance scheme in Mauritania on maternal healthcare utilization : A propensity score matching analysis. *HEALTH POLICY AND PLANNING*, 35(4), 388-398. <https://doi.org/10.1093/heapol/czz150>
- Ravalihasy A**, Faye A, Diallo AI, G Ibrahima, Ridde V. A social acceptability scale: validation in the context of government measures to curb the COVID-19 pandemic in Senegal [en révision]

Introduction

« *We have the duty of formulating, of summarising, and of communicating our conclusions, in intelligible form, in recognition of the right of other free minds to utilize them in making their own decisions.* »

(Nous avons le devoir de formuler, résumer, et communiquer nos conclusions, de manière compréhensible, en reconnaissance du droit des autres esprits libres à les utiliser dans leurs propres décisions.

[Traduction libre])

Ronald Aylmer Fisher, *Statistical Methods and Scientific Induction*, 1955

Plus de soixante-dix ans se sont écoulés depuis que les états membres de l'Organisation Mondiale de la Santé (OMS) ont intégré dans la définition de la santé les états de bien-être physique, mental et social¹ (Organisation mondiale de la Santé, 1946). Cette définition démontre la nécessité de porter les actions de santé publique au-delà des facteurs directs de la détérioration de la santé des populations, tels que les agents pathogènes, pour s'assurer de la bonne santé des populations. Elle implique la reconnaissance d'une portée sociale de la santé qui s'inscrit dorénavant dans des actions coordonnées non exclusives au secteur biomédical (Organisation mondiale de la Santé, 1978). Les enjeux de santé des populations s'imbriquent dans des enjeux socioéconomiques et structurels, faisant de la santé une ressource pouvant impacter et être impactée par les dynamiques qui en résultent. Les programmes et dispositifs de santé qui en découlent ambitionnent de proposer des actions concrètes à la jonction de ces enjeux à travers une concertation et une participation des parties concernées. A travers ces programmes, on s'attend d'une part à une prise en compte de manière spécifique des déterminants de la santé dans leur diversité (Dahlgren & Whitehead, 1991, 2021), et d'autre part à une participation plus effective des premiers concernés, les individus comme les communautés, dans la mise en œuvre de ces programmes depuis leur conception jusqu'à leur déploiement (Organisation mondiale de la Santé, 1986). Ces actions de promotion de la santé revêtent donc une dimension politique, multisectorielle et impliquent différents acteurs.

Dans cette perspective, on assiste au déploiement de programmes complexes, ou du moins à une lecture complexe de ces programmes (Petticrew, 2011) tant par les imbrications de leurs contenus, ou leur mise en œuvre dans différents contextes, que par la diversité des acteurs impliqués. Adopter ces stratégies complexes s'avère pourtant nécessaire et pertinente (Diderichsen et al., 2001) dans un monde où on constate, de manière globalisée, une

¹ La santé se définit comme suit dans la constitution de l'OMS : « La santé est un état de complet bien-être physique, mental et social, et ne consiste pas seulement en une absence de maladie ou d'infirmité. »

distribution différenciée de l'occurrence des maladies en fonction des inégalités socioéconomiques et structurelles au niveau des populations (Beckfield et al., 2013; Bleich et al., 2012; Gwatkin, 2017). De telles stratégies présentent des défis dans leur réalisation, notamment dans la mobilisation et la redistribution des ressources de santé, et requièrent de pouvoir identifier les niveaux auxquels s'opèrent les inégalités (Diderichsen et al., 2019; Francis-Oliviero et al., 2020). Malgré ces défis, elles présentent le potentiel d'envisager la réalisation de politiques de santé socialement justes (Marmot & Bell, 2012).

Dans ce contexte, le choix des programmes de santé à mettre en œuvre s'effectue logiquement et idéalement en mettant en perspective les ressources à allouer et redistribuer avec la pertinence et l'efficacité des programmes en question (Gertler et al., 2016b). Ces questions évaluatives sont particulièrement importantes dans la caution qu'elles apportent dans la prise de décisions liée au déploiement d'un programme (Alkin & Taut, 2021; Devaux-Spatarakis et al., 2021). Ainsi, indépendamment des usages possibles des informations résultant des évaluations, ces dernières ont pour première vocation de rendre compte de la pertinence des activités et ressources mobilisées dans le cadre de ces programmes. En ce qui concerne l'évaluation de l'impact de ces derniers sur la santé des populations, l'usage de méthodes quantitatives comparant des personnes bénéficiant et ne bénéficiant pas des dispositifs est plébiscité comme le témoignent plusieurs manuels sur le sujet (Gertler et al., 2016a; Linfield & Posavac, 2018; Newcomer et al., 2015; Rossi et al., 2019). Souvent, ces méthodes permettent uniquement de constater qu'un programme a été efficace dans une configuration donnée. Elles sont mobilisées selon une séquence où les programmes en question sont testés, ensuite évalués dans des conditions optimales puis en conditions réelles avant d'être déployés à l'échelle d'une population. La pertinence de cette procédure d'évaluation, qui repose sur l'aptitude des méthodes à estimer l'ampleur des impacts des programmes sur la santé, apurée de l'effet du contexte de mise en œuvre, motive en partie les critiques à son encontre (Chen,

2010). D'une part, les estimations d'impact obtenues selon ces procédures se focalisent sur les mécanismes des composantes des programmes, mettant au second plan ceux tributaires de leur mise en œuvre et qui sont pourtant utiles à leur déploiement à une échelle populationnelle ou leur répliation dans un autre contexte. D'autre part, la séquence d'évaluation telle que recommandée dans cette configuration ne peut être exécutée systématiquement selon la nature des programmes et les ressources disponibles. Ces limites sont à l'origine de débats et d'autres propositions méthodologiques concernant l'évaluation d'impact des programmes de santé publique (Chen, 2010, 2015d; Pawson & Tilley, 1997), en complément des méthodes précédentes qui ne sont désormais plus considérées comme référence. L'usage de ces derniers fait particulièrement l'objet de réserve chez les acteurs qui défendent une vision des programmes de santé publique où les individus et les communautés sont mis à contribution pour changer les comportements et les environnements sociaux de manière à améliorer la santé² (de Salazar & Hall, 2009; Potvin & McQueen, 2009; Rootman et al., 2001). Si le fait de faire reposer l'évaluation de ces programmes sur des données probantes reste légitime, il est nécessaire de repenser la notion de preuve et la considération des contextes (de Salazar & Hall, 2009; McQueen & Anderson, 2001; O'Neill, 2003; Potvin & McQueen, 2009).

Des programmes complexes et des évaluations complexes : un impératif pour réduire l'épidémie de VIH

Les recherches et actions effectuées dans l'histoire de la lutte contre la propagation du virus de l'immunodéficience humaine (VIH) et du syndrome d'immunodéficience acquise (SIDA) montrent que les solutions potentielles ne peuvent résulter que de la combinaison d'approches biomédicales, sociales et structurelles, faisant de la prévention du VIH un domaine d'étude

² On fait référence ici aux principes propres au domaine de la promotion de la santé tels qu'énoncés par la Charte d'Ottawa (Organisation mondiale de la Santé, 1986)

majeur dans le champ de la promotion de la santé (Piot et al., 2008). Les recommandations concernant la conception et la mise en œuvre de programmes de lutte contre le VIH soulignent l'importance des enjeux sociaux et structurels essentiels dans la lutte contre l'épidémie. Ainsi, la maîtrise de la propagation du virus repose actuellement sur le potentiel que représentent les programmes de prévention combinée du VIH (PPCV) (Joint United Nations Programme on HIV/AIDS, 2007, 2021). Ces programmes cherchent à mettre à profit dans une stratégie combinée les avantages des outils biomédicaux, des interventions comportementales et sociales, et des approches structurelles. Le potentiel que représente ces programmes est accompagné de défis lors de toutes les activités qui en découlent. En ce qui concerne l'évaluation de l'impact sur la santé des populations, les recommandations prônent la pertinence d'une diversité de méthodes à mobiliser de manière à évaluer au mieux la pertinence de chaque composante des programmes, en particulier mais également dans l'ensemble, au regard des contextes de leur déploiement (Joint United Nations Programme on HIV/AIDS, 2010). Ces évaluations sont particulièrement importantes afin de disposer de plus d'informations sur les stratégies les plus à même d'être efficaces en fonction des populations concernées et des enjeux locaux. Il est attendu des évaluations d'impact des PPCV qu'elles fournissent des informations opérationnelles dans la mesure du possible.

Objectifs et structure de la thèse

Les attentes vis-à-vis des programmes de santé publique, notamment dans la lutte contre le VIH, ne font que rendre saillants les limites de l'usage qui est fait des méthodes quantitatives telles que mobilisées actuellement dans l'évaluation d'impact (Cambon & Alla, 2019). Dans cette thèse, nous postulons que ces méthodes sont pertinentes à condition que leur usage soit cohérent avec les objectifs et le contexte des programmes. A travers ce travail, nous montrons

les enjeux propres aux indicateurs et les enjeux propres aux méthodes d'estimation pour prendre en compte le contexte et les objectifs des programmes dans l'évaluation de leur impact. Enfin, nous montrons l'importance de l'engagement de l'évaluateur dans le dispositif des programmes de santé pour orienter les choix d'évaluation. Ainsi, l'usage et les choix liés à ces méthodes sont également complexes, leur pertinence étant tributaire de la cohérence avec les objectifs et la mise en œuvre de ces programmes. Bien que les estimations produites par ces méthodes puissent parfois être dénuées de l'effet des contextes pour mieux apprécier l'impact des contenus de l'intervention, les choix dans la mise en œuvre des méthodes quantitatives peuvent tenir compte de ces contextes.

Pour réaliser ces objectifs, nous proposons une revue systématique de la littérature sur les méthodes quantitatives d'évaluation d'impact des PPCV et une étude de cas à partir d'un PPCV. De manière plus spécifique, nous prenons le cas de l'intervention MAKASI, une intervention d'*empowerment* en santé sexuelle à destination des immigrés originaires d'Afrique subsaharienne (IASS) vivant des situations de précarité en France.

Cette thèse est structurée comme suit :

- Une première partie introductive présente les généralités sur les méthodes d'évaluation d'impact et leurs enjeux en général, et dans la lutte contre l'épidémie de VIH en particulier
- La deuxième partie propose un recensement critique des méthodes utilisées pour évaluer l'impact des PPCV sur la propagation du VIH.
- Dans une troisième partie, nous présentons une étude de cas de la mise en œuvre de l'évaluation dans le cas de l'intervention MAKASI. Nous étudions deux cas portant sur les indicateurs et les méthodes quantitatives d'évaluation d'impact : la validation d'une échelle de mesure de l'*empowerment* en santé sexuelle et un exercice de

modélisation d'impact de l'intervention sur la transmission du VIH au niveau des bénéficiaires de l'intervention. Nous mettons ensuite en perspective les choix liés à l'usage de ces méthodes avec l'interaction de l'évaluateur avec le dispositif d'évaluation.

- La quatrième partie servira à débattre les résultats de la thèse et en tirer les implications pour l'évaluation des programmes complexes en santé publique.

Partie – I : Etat de l’art sur l’évaluation d’impact

« Remember, always, that everything you know, and everything everyone knows, is only a model. Get your model out there where it can be viewed.

Invite others to challenge your assumptions and add their own. »

(Souvenez-vous, toujours, que tout ce que vous connaissez, et tout ce que tout le monde connaît, est juste un modèle. Exposez votre modèle pour qu’il soit accessible. Invitez d’autres personnes à remettre en question vos hypothèses et présenter les leurs. [Traduction libre])

Donella Meadows, *Thinking in Systems: A primer*, 2008

Chapitre – 1 : L’usage des méthodes quantitatives dans l’évaluation d’impact des interventions en santé

1. L’émergence de l’expérimentation aléatoire comme méthode de référence de l’évaluation d’impact

La notion de *politique basée sur des données probantes*³ héritée de la *médecine basée sur les données probantes*⁴ a, au fil des années, conquis le monde des acteurs impliqués dans le choix et le déploiement des politiques et dispositifs socio-sanitaires (Baron, 2018). Le principe de politique basée sur des données probantes repose sur l’accumulation préalable de preuves quant à l’efficacité d’un dispositif avant sa mise en œuvre au sein d’une population (Chalmers, 2003; Gertler et al., 2016b; Sackett et al., 1996). La notion de preuves dans ce contexte consiste en l’attribution à une politique donnée de la capacité à améliorer la situation des bénéficiaires. Dans l’idéal, cette notion serait quantifiée en comparant la situation des personnes bénéficiant du dit dispositif avec ces mêmes personnes si elles n’en bénéficiaient pas. Cette comparaison introduit un raisonnement contrefactuel : mesurer ce qui serait arrivé dans les mêmes conditions et au même moment en présence ou en absence du dispositif.

En pratique, elle se quantifie en comparant la situation de personnes bénéficiant du dispositif considéré avec d’autres personnes n’en bénéficiant pas. Pour que les différences observées entre ces deux groupes puissent être attribuées à l’efficacité du dispositif, ces deux groupes doivent être comparées de manière à ce que leurs caractéristiques ne puissent pas altérer cette

³ Traduction de la notion d’ « evidence-based policy ». Selon l’Encyclopædia Britannica, les politiques basées sur les données probantes sont fondées sur l’utilisation de preuves empiriques obtenues à l’aide de méthodes scientifiques.

⁴ Traduction de la notion d’ « evidence-based medicine ». Selon l’Encyclopædia Britannica, cette notion est relative à la considération des preuves scientifiques les plus récentes et pertinentes dans les prises de décision quant aux soins à prodiguer à un patient.

attribution (Reichardt, 2019a). Le raisonnement contrefactuel est appliqué en utilisant des méthodes permettant de faire l'hypothèse que la situation des groupes comparés serait identique en absence du dispositif à évaluer et qu'elles réagiraient de la même manière en sa présence. Les méthodes mobilisées sont donc censées permettre de constituer de tels groupes et/ou de corriger les biais liés à la constitution de ces groupes et pouvant affecter l'estimation de l'impact du dispositif. L'ampleur de la différence de situation entre ces personnes, lorsqu'elle est jugée suffisante, constitue alors une preuve que le dispositif est bien la cause d'un changement de situation chez les bénéficiaires.

Actuellement, les expérimentations aléatoires constituent une méthode réputée au point d'être considérées comme méthode de référence⁵ parmi celles mobilisées pour la réalisation d'analyses contrefactuelles (Reichardt, 2019h; Sackett et al., 1996). Si on prend l'exemple le plus simple de son utilisation pour tester l'effet d'un dispositif, la méthode consisterait à sélectionner aléatoirement un échantillon de personnes et faire bénéficier le dispositif, toujours de manière aléatoire, à la moitié des personnes sélectionnées, la différence de situations entre les deux groupes conséquents constituant l'impact du dispositif. Cette manière de procéder permet de rendre les membres de l'échantillon interchangeables dans l'allocation du dispositif rendant ainsi viable le raisonnement contrefactuel car l'attribution du dispositif tout comme son effet deviennent également interchangeables aux individus, bien qu'en réalité cet effet n'est interchangeable qu'en moyenne⁶. La dissémination de cette méthode, dite de randomisation, pour évaluer des dispositifs dans différentes disciplines et secteurs d'activité trouve ses origines dans les travaux de Ronald Aylmer Fisher, bien que des exemples de son usage précèdent ses travaux (Hall, 2007). La promotion de son usage a rencontré la réticence

⁵ On fait référence ici au terme de « gold-standard »

⁶ L'effet du dispositif est calculé comme une différence globale (donc moyenne) entre le groupe bénéficiant et le groupe ne bénéficiant pas du dispositif. Attribuer cet effet à chaque individu revient donc à attribuer à chaque individu une moyenne calculée sur l'ensemble de l'échantillon

des contemporains de Fisher, cependant les innovations statistiques qui ont accompagné ses travaux sur la randomisation ont abouti à sa large diffusion dans le monde scientifique, en particulier dans le domaine de l'évaluation.

Le procédé de randomisation, dont la simplicité n'est qu'apparente, a permis d'innover la manière de traiter statistiquement les erreurs de mesure dans les expérimentations scientifiques (Armitage, 2003; Hall, 2007). Avant l'utilisation répandue de la randomisation dans la constitution des échantillons, les procédures utilisées consistaient à minimiser les erreurs en contrôlant tant que peut se faire les conditions d'expérimentations et la similarité entre les échantillons à comparer. Ces procédures ne permettaient pas de suffisamment garantir la comparabilité des échantillons et des conditions expérimentales, entre autres sur des facteurs non observés, ni de quantifier les erreurs effectuées pour garantir la réplicabilité des expérimentations. La procédure de randomisation proposée par Fisher a répondu à ces enjeux en proposant une méthode qui, premièrement, permet d'éliminer théoriquement, ou plutôt pratiquement de minimiser les erreurs liées au matériel expérimental et celles liées au chercheur. D'une part, la randomisation concède à tous les échantillons pouvant être constitués les mêmes propriétés, permettant ainsi d'effectuer une estimation sans biais de l'effet du dispositif (Rubin, 1972). D'autre part, ces erreurs deviennent purement aléatoires (donc le fruit du hasard) grâce à la randomisation qui leur confère par la même occasion des propriétés pouvant être décrites en utilisant les méthodes proposées par la théorie des probabilités⁷. Deuxièmement, cette manière de constituer les échantillons et comparer les groupes permet de calculer dans quelles mesures les différences observées peuvent être dues au hasard. La théorie des probabilités peut être utilisée pour calculer la fraction d'échantillons qui auraient abouti à une différence d'une même ampleur ou plus parmi tous ceux qui auraient pu être

⁷ Selon l'Encyclopædia Britannica, la théorie des probabilités est une branche des mathématiques dédiée à l'étude des phénomènes aléatoires

constitués. Ainsi, on obtiendrait la probabilité que le hasard à lui seul puisse être responsable des différences observées ou des estimations calculées⁸ (Rubin, 1972). Cette manière d'effectuer des estimations s'est avérée être importante par la suite car elle permet de constater si les effets d'un dispositif sont dus au hasard même dans le cas où leur ampleur n'est pas flagrante (Baron, 2018).

Ces propriétés ont valu à la randomisation une attractivité et une adoption, notamment, auprès d'épidémiologistes à partir des années trente devant l'absence d'évaluation standardisée des traitements à appliquer aux patients face à des maladies sévissant à cette époque (Baron, 2018; Doll, 1998). La nécessité de fournir des preuves quant à l'efficacité de traitements ont vu apparaître des exemples de premières publications concernant des expérimentations basées sur la randomisation en médecine (Francis Jr et al., 1955; Medical Research Council, 1948). Les motivations quant à l'usage de ces procédures en médecine semblent plus relever de l'éthique et de l'intégrité scientifique que de la précision statistique (Armitage, 2003; Doll, 1998). Alors que les premiers arguments en faveur de son utilisation reposaient sur des propriétés statistiques, son adaptation au monde biomédical était accompagnée d'arguments d'ordre éthique reposant sur l'égalité des chances des patients à bénéficier d'un traitement et l'évitement des biais conscients et inconscients des expérimentateurs dans l'allocation du traitement. Néanmoins, les résultats obtenus pour justifier la pertinence d'un dispositif, comme l'usage de la streptomycine dans le traitement de la tuberculose ou le vaccin contre la poliomyélite, ont persuadé au point de motiver son institutionnalisation par la *Food and Drug Administration* aux Etats-Unis en 1962 (Junod, 2008). De nos jours, l'utilisation de ces essais randomisés contrôlés (ECR)⁹ pour fournir les preuves d'efficacité d'un dispositif médical est

⁸ Nous retrouvons ici la notion de significativité statistique quantifiée par le p-value (probability value), paramètre souvent évoqué et utilisé pour donner une valeur quantitative à cette significativité

⁹ Les essais randomisés contrôlés, autre dénomination attribuée aux expérimentations aléatoires dans le monde biomédical, font référence aux méthodes expérimentales dans lesquelles on compare des groupes recevant le dispositif à évaluer avec des groupes ne le recevant pas, les groupes étant constitués aléatoirement

devenue une norme internationale comme le témoignent les bonnes pratiques cliniques incluses dans les protocoles de mise en œuvre fournis par le Conseil International d'Harmonisation des exigences techniques pour l'enregistrement des médicaments à usage humain (ICH)¹⁰ ou les publications successives de recommandations quant à la transparence des rapports sur les études expérimentales menées à travers le *CONSORT statement* (Begg et al., 1996; Moher et al., 2001, 2010; Schulz et al., 2010).

Bien que son usage au-delà des interventions médicamenteuses ne soit pas aussi répandue, les ECR sont utilisés depuis longtemps. On peut observer des exemples de son usage, notamment dans les interventions sociales des années trente à nos jours, laissant des résultats mitigés quant à sa pertinence dans d'autres domaines que celui de la médecine (Baron, 2018). Malgré les exemples de son usage plus nombreux qu'à ses débuts et des évaluateurs en faveur de son usage (Campbell, 1971; Chalmers, 2003), des critiques quant à sa pertinence dans l'évaluation au-delà de la recherche biomédicale se font plus nombreux (Cartwright & Hardie, 2012b; Deaton & Cartwright, 2018; Hammersley, 2005). Malgré ces réticences, des chercheurs travaillent pour adapter au mieux son usage lorsqu'on sort du cadre de la médecine clinique, dans la lignée de ceux persuadés de son importance pour fournir des preuves quant à l'efficacité des dispositifs (Boutron et al., 2017; Montgomery et al., 2018).

2. La logique contrefactuelle et les plans d'étude d'évaluation d'impact

Si les propriétés statistiques des ECR ont contribué à les considérer comme référence pour évaluer l'impact d'un dispositif selon une logique contrefactuelle, d'autres méthodes sont apparues au fur et à mesure pour remettre en question ce statut de méthode de référence. En effet, la nécessité d'accumuler les preuves scientifiques ainsi que l'impossibilité de recourir

¹⁰ <https://www.ich.org/>

systématiquement aux ECR ont imposé l'usage d'autres méthodes en pratique (Cartwright & Hardie, 2012b; Deaton & Cartwright, 2018; Petticrew et al., 2005). Ces méthodes dans lesquelles l'allocation des dispositifs ne se fait pas aléatoirement sont appelées méthodes quasi-expérimentales (Reichardt, 2019g; Rubin, 1972) lorsqu'elles servent néanmoins à estimer l'impact des dispositifs (les méthodes expérimentales désignant dans ce contexte celles où l'allocation des dispositifs se fait de manière aléatoire, donc les ECR). Tout comme les ECR, la possibilité d'effectuer une analyse causale dépend de deux éléments : l'existence d'un ensemble d'hypothèses scientifiques viables permettant de définir une relation causale basée sur un raisonnement contrefactuel et la possibilité de quantifier cette relation causale à l'aide de théories issues de la probabilité et de la statistique. Les méthodes quasi-expérimentales ne facilitent pas la mise en pratique de ces hypothèses. Elles permettent néanmoins de satisfaire ces conditions pourvu qu'on dispose d'assez d'informations à propos du mécanisme qui détermine la constitution des groupes bénéficiant et ne bénéficiant pas des dispositifs à évaluer (Rosenbaum & Rubin, 1983; Rubin, 1972, 1977).

a. La causalité contrefactuelle n'est pas que statistique

La capacité des procédures de randomisation à faciliter l'hypothèse contrefactuelle et fournir des estimations d'impacts assorties d'estimations de leur précision ont rendu les méthodes expérimentales (les ECR) plus séduisantes que les méthodes quasi-expérimentales. Les difficultés pratiques et éthiques liées à la nature de l'expérimentation (par exemple des dispositifs sanitaires et sociaux, parfois déjà déployés), la nature des données et des impacts à mesurer (lorsqu'ils nécessitent une taille d'échantillon conséquente ou une longue période d'observation par exemple), ou l'ampleur des ressources à mobiliser (lorsque les moyens humains, financiers ou même les infrastructures sont insuffisants ou inadaptés) ne permettent pas d'en user systématiquement (Petticrew et al., 2005; Reichardt, 2019c). Ces difficultés ont

amené des chercheurs, dont Ronald Fisher lui-même (Armitage, 2003; Fisher & Bartlett, 1931), à trouver des solutions permettant d'effectuer des inférences causales dans le cas où les groupes comparés n'ont pas été constitués par randomisation. Si les travaux de Fisher s'imposent comme étant les prémices de l'évaluation d'impact tels que connus actuellement dans le paradigme des politiques basées sur des données probantes, le cadre théorique utilisé pour quantifier l'impact dans une logique causale est surtout liée aux travaux de Ronald Rubin, inspirés par les travaux de Ronald Fisher et Jerzy Neyman (Fisher, 1925, 1971; Neyman et al., 1990; Rubin, 2005). Le modèle causal de Neyman-Rubin ou modèle à résultats potentiels (terme préféré à contrefactuel par Rubin lui-même (Rubin, 2005)) développe de manière plus explicite les conditions pour lesquelles la randomisation permet d'estimer l'impact d'une intervention selon une logique contrefactuelle et comment ces conditions s'appliquent également aux méthodes quasi-expérimentales. Le modèle à résultats potentiels est avant tout un modèle théorique d'analyses causales dans le cadre d'estimations statistiques. Ce modèle distingue les solutions scientifiques, c'est-à-dire les hypothèses scientifiques viables¹¹ qui permettent de définir ce qu'est un effet d'un dispositif, et les solutions statistiques, qui proposent une quantification de l'effet moyen du dispositif au sein de la population à qui il est destiné (Holland, 1986; Rubin, 2005). Les solutions scientifiques proposées par ce cadre théorique résident principalement dans un ensemble d'hypothèses appelé *Stable Unit Treatment Value Assumption* (SUTVA) (Holland, 1986; Reichardt, 2019a; Rubin, 2005). Ce faisceau d'hypothèses stipule qu'il existe un effet de l'intervention :

- Spécifique à chaque individu qu'il bénéficie ou non du dispositif. C'est un principe de non interférence : l'effet est spécifique de manière à ce que l'estimation obtenue ne

¹¹ Holland parle de *scientific solution* en opposition avec *statistical solution*

dépende pas du mécanisme de constitution des groupes à comparer peu importe qui reçoit finalement le dispositif lors de l'expérimentation.

- Spécifique à chaque individu peu importe dans quelles conditions il bénéficie du dispositif. L'effet est invariable (ou très peu variable) de manière à ce qu'il ne puisse pas être fondamentalement modifié ou défini par les conditions expérimentales.

Ces conditions illustrent les termes de résultats potentiels (Figure 1) : pour chaque individu, chaque échantillon, il existe un effet potentiel du traitement et un effet potentiel de l'absence de traitement uniques, l'estimation causale reviendrait à une comparaison quantitative des effets potentiels appliquée à chaque unité.

Figure 1 : Illustration du modèle à résultats potentiels de Rubin

Unité considérée	Effets potentiels		Effet causal au niveau individuel	Effet causal sur l'ensemble
	Traitement	Contrôle		
1	$Y_1(1)$	$Y_1(0)$	$Y_1(1) \text{ vs } Y_1(0)$	<i>Comparaison de l'ensemble des effets</i> $Y_i(1) \text{ vs } Y_i(0)$
.	.	.	.	
.	.	.	.	
l	$Y_l(1)$	$Y_l(0)$	$Y_l(1) \text{ vs } Y_l(0)$	
.	.	.	.	
.	.	.	.	
N	$Y_N(1)$	$Y_N(0)$	$Y_N(1) \text{ vs } Y_N(0)$	

Source : Adapté de (Rubin, 2005), *Lecture* : la colonne « Unité considérée » représente les individus, la colonne « Traitement » représente l'effet potentiel du bénéfice du dispositif, la colonne « Contrôle » représente l'effet potentiel en absence du dispositif, la colonne « Effet causal au niveau individuel » représente l'impact du dispositif au niveau individuel, la colonne « Effet causal sur l'ensemble » représente l'impact moyen du dispositif sur l'ensemble des individus.

Une conséquence directe du SUTVA est la possibilité de quantifier un estimateur de l'effet moyen du dispositif (Équation 1). En effet, si les effets potentiels sont invariables et indépendants du mécanisme d'assignation du dispositif et des conditions expérimentales, il en résulte que la différence entre les résultats potentiels moyens en présence et en absence de traitement équivaut à la différence moyenne entre les résultats potentiels en présence et en absence de traitement. Cette égalité informe sur le fait que les informations issues de groupes

bénéficiant, d'une part, et d'autre part de groupes ne bénéficiant pas du dispositif peuvent être mises à contribution pour estimer son impact sur une population (Holland, 1986; Rubin, 1972).

Équation 1 : Illustration de l'estimation de l'effet d'un dispositif dans le cadre du modèle à effets potentiels

$$\bar{Y}_1 - \bar{Y}_0 = \sum_{i=1}^N \frac{Y_i(1) - Y_i(0)}{N}.$$

Source : (Rubin, 2005), Lecture : \bar{Y}_1 et \bar{Y}_0 représentent respectivement les effets potentiels moyens du bénéfice et de l'absence du dispositif, $\sum_{i=1}^N \frac{Y_i(1) - Y_i(0)}{N}$ représente l'impact moyen du dispositif sur l'ensemble des individus

L'application du SUTVA dans le cadre de la randomisation est aisée car celle-ci présente des propriétés favorables : elle permet d'obtenir un estimateur sans biais de l'effet moyen du dispositif et associer à celui-ci la probabilité que le hasard en soit responsable (Rubin, 1972). Cependant, en regardant de plus près les hypothèses sous-jacentes au SUTVA et en y ajoutant une connaissance plus fine de l'attribution des dispositifs, il est possible d'effectuer le même type d'analyse causale en utilisant des méthodes quasi-expérimentales (Rosenbaum, 1984; Rosenbaum & Rubin, 1983; Rubin, 1972, 1977, 2005). En premier lieu, peu importe la procédure expérimentale adoptée, les liens causaux sont définis de la même manière (Holland, 1986; Rubin, 2005) : les liens causaux existent, sont stables et invariables et ne dépendent ni du mécanisme d'assignation du dispositif, ni des conditions expérimentales. En second lieu, les estimations dans le cadre de méthodes expérimentales comme quasi-expérimentales sont exposées à de potentiels facteurs de confusion pouvant aboutir à une variation des résultats selon les échantillons constitués lors d'une expérimentation (Rubin, 1972). Ainsi, au-delà de l'application des hypothèses de SUTVA, une connaissance du mécanisme par lequel les dispositifs sont alloués et la possibilité pour chaque individu d'en bénéficier (c'est à dire une propension non nulle à bénéficier ou ne pas bénéficier du dispositif appelée hypothèse forte d'ignorabilité¹²) (Rosenbaum, 1984; Rosenbaum & Rubin, 1983) permettent d'estimer l'effet

¹² Il s'agit ici d'une traduction libre de la notion de « *strongly ignorable treatment assignment* »

moyen d'un dispositif dans n'importe quelle expérimentation. Des estimations sans biais de l'effet moyen du dispositif peuvent être obtenues en tenant compte du mécanisme d'assignation du dispositif et éventuellement de la distribution des facteurs de confusion dans les groupes de comparaison constitués (Rosenbaum, 1984; Rosenbaum & Rubin, 1983; Rubin, 1972, 1977). Il est également possible d'en déduire le niveau de significativité statistique correspondant. Des informations concernant les mécanismes d'assignation des dispositifs sont donc nécessaires et peuvent s'avérer essentielles pour estimer correctement l'impact d'un dispositif peu importe si ces mécanismes sont complètement aléatoires ou non lors de la réalisation d'une expérimentation (Pearl, 2000, 2009; Pocock et al., 2002).

b. Les plans quasi-expérimentaux

La partie précédente concernant le modèle à résultats potentiels nous rappelle que peu importe le mécanisme d'attribution des dispositifs, randomisé ou non, l'interprétation causale ne relève pas que de méthodes quantitatives mais également d'hypothèses scientifiques viables. Il existe des configurations et des raisons pour lesquelles il est nécessaire d'évaluer l'efficacité d'un dispositif alors qu'il n'est ni possible ni souhaitable de recourir aux ECR (Petticrew et al., 2005; Reichardt, 2019c). Dans ces configurations, les données disponibles relèvent d'une configuration quasi-expérimentale. En fonction des cas :

- L'évaluation des dispositifs concerne un seul groupe de personnes ayant bénéficié du dispositif à évaluer, l'évaluation étant effectuée en considérant la période qui suit la mise en œuvre du dispositif (étude après mise en œuvre) ou les périodes précédant et succédant cette mise en œuvre (étude avant-après)
- L'évaluation des dispositifs compare des groupes de personnes parmi lesquels certains ont bénéficié du dispositif et d'autres non (groupes non équivalents).

- L'évaluation des dispositifs combine l'avantage des deux précédents, c'est-à-dire : la comparaison des groupes bénéficiant et ne bénéficiant pas du dispositif avant et après sa mise en œuvre.

La première configuration (étude après mise en œuvre) où on ne dispose de données que pour les personnes ayant bénéficié du dispositif est sans doute celle considérée moins fiable (Reichardt, 2019e). Elle est parfois considérée comme un devis pré-expérimental et non quasi-expérimental étant donné la difficulté évidente à appliquer un raisonnement contrefactuel dans l'analyse des données. En effet, on ne dispose même pas d'information concernant l'effet potentiel réciproque de celui qui est observé. Cette configuration est néanmoins considérée dans cette thèse étant donné qu'elle est parfois mobilisée dans les analyses de modélisation mathématique pour fournir des estimations contrefactuelles. Des exemples de ce type d'études sont fréquents concernant les programmes de lutte contre le VIH (Ravalihasy et al., 2023) en mobilisant des données d'études observationnelles ou interventionnelles à l'instar du projet *Yerelon* au Burkina Faso (Traore et al., 2015) ou le projet Sida 1/2/3 au Bénin (Williams et al., 2014).

Lorsqu'on ne dispose de données que sur les personnes ayant bénéficié de l'intervention, une meilleure alternative à la configuration précédente consiste à disposer de données avant et après la mise en œuvre d'un dispositif (Reichardt, 2019f). Cette procédure a le mérite de comparer l'usage du dispositif au sein du même groupe de personnes en absence d'une procédure de randomisation dans son attribution. Cependant, en pratique il est difficile de s'assurer que les estimations respectent les hypothèses du SUTVA étant donné que les conditions expérimentales avant et après l'utilisation du dispositif peuvent varier. Une extension de ce type d'étude pour limiter ces désavantages consiste à appliquer plusieurs mesures dans le temps avant et après traitement. La multiplication des points de comparaison

contribue à homogénéiser les conditions pré et post intervention (les différences observées sous plusieurs conditions ne sont peut-être plus dues au hasard et la faiblesse du plan d'étude mais bien attribuable au traitement) même si elle ne corrige pas tous les biais (Reichardt, 2019b, 2019e, 2019i). Elle a également d'autres avantages comme la possibilité d'investiguer les tendances temporelles de l'effet d'un dispositif ou améliorer les procédures éthiques de sélection des bénéficiaires d'un dispositif (Reichardt, 2019i). Des données d'études observationnelles ou interventionnelles, éventuellement issues de cohortes, peuvent être mises à profit dans ce type d'évaluation à l'exemple du projet *DREAMS* (Birdthistle et al., 2021) ou du projet *UALE* (Sabido et al., 2009).

L'évaluation basée sur la comparaison de groupes recevant et ne recevant pas le bénéfice du dispositif est sans doute la plus utilisée dans le domaine des programmes sociaux et comportementaux (Reichardt, 2019d). Lorsque l'attribution des dispositifs ne se fait pas par une procédure randomisée, les hypothèses de SUTVA sont difficilement vérifiables. Bien que diverses méthodes statistiques soient disponibles pour faire l'effort de corriger les biais éventuels de cette attribution non homogène des dispositifs (Pearl, 2000; Rosenbaum, 1984; Rosenbaum & Rubin, 1983; Rubin, 1972, 1977, 2005), leur utilisation est tributaire de la capacité des procédures lors des expérimentations à fournir une connaissance et des mesures fines concernant les mécanismes d'attribution des dispositifs. Et même lorsque ces méthodes sont utilisées, il est toujours permis d'évoquer l'existence éventuelle de biais non mesurés. Dans certains cas, ce type d'évaluation est réalisé en fonction des données disponibles à l'exemple des résultats issus du projet *AVAHAN* (Thilakavathi et al., 2011) ou du projet *MEMA kwa Vijana* (Doyle et al., 2010).

Les limites évoquées concernant les méthodes quasi-expérimentales peuvent être contournées en combinant les avantages des devis avant-après et des groupes non équivalents pour garantir au mieux la robustesse des résultats d'évaluation lorsque c'est possible. Par exemple, la

combinaison d'une série temporelle avec une comparaison de groupes non équivalents permet de mieux considérer les conditions expérimentales avant et après le bénéfice du dispositif tout en disposant de groupes contrefactuels tout le long. De nombreux exemples de ce type d'évaluation existent, parmi lesquels les programmes *LALS* (Peak et al., 1995), *AVAHAN* (Boily et al., 2013), ou *China GF-5* (Ye et al., 2012).

3. Les enjeux de validité des résultats au-delà des enjeux statistiques

Les enjeux méthodologiques et statistiques autour de la causalité dans les études d'évaluation d'impact tels que décrits précédemment découlent d'enjeux plus larges liés à l'évaluation des programmes : la pertinence et la prise de décision quant à l'utilisation des dispositifs qui sont l'objet même de l'évaluation (Gertler et al., 2016b). Les conditions dans lesquelles les expérimentations et les évaluations ont lieu amènent logiquement à questionner la pertinence des résultats dans le contexte de ces expérimentations et au-delà, puisque ces dispositifs ont pour vocation à être utilisés dans des conditions réelles. Campbell et Stanley (Campbell & Stanley, 1966) ont proposé les notions de validités interne et externe des résultats de l'évaluation d'impact en guise de réponse à cette question. Les notions de validité renvoient à la pertinence des inférences statistiques effectuées à partir des estimations obtenues. La notion de validité interne fait référence à la crédibilité des résultats pour isoler l'effet d'un dispositif sur ses bénéficiaires. En d'autres termes, la validité interne décrit dans quelle mesure l'estimation obtenue est nette de toute interférence extérieure et permet de conclure qu'elle est intrinsèque uniquement au dispositif qui est seul responsable (cause) du changement au niveau des bénéficiaires (Reichardt, 2019j). La validité interne est spécifique au contexte de l'expérimentation et demande d'ailleurs un degré de contrôle assez conséquent notamment

concernant huit paramètres¹³ liés à l'environnement et aux dispositifs d'application de cette expérimentation (Campbell & Stanley, 1966). La validité externe, quant à elle, détermine dans quelles mesures les résultats obtenus sont valables dans d'autres contextes lorsqu'ils sont soumis aux interférences extérieures liées à ces contextes. En d'autres termes, la validité externe permet de décrire le maintien de l'effet des dispositifs dans les conditions qui intéressent le déploiement du dispositif. Un effet ayant une bonne validité externe est alors généralisable à la population à qui est destiné le dispositif (Reichardt, 2019j). Pour assurer la validité externe, Campbell et Stanley préconisent de porter l'attention sur quatre paramètres¹⁴ (Campbell & Stanley, 1966) qui nuisent la représentativité de l'échantillon et des effets du dispositif en dehors du contexte expérimental. Ces deux types de validité sont complémentaires pour fournir des éléments d'inférences garantissant à la fois un lien de causalité entre dispositifs et bénéficiaires tout en assurant que les effets soient transposables dans la réalité quotidienne des usagers. Ils sont également interdépendants dans le sens où l'optimisation de l'un lèse l'autre, et vice versa.

Les travaux pour assurer la crédibilité de l'inférence faite à partir des résultats ont amené à identifier deux autres types de validité, la validité de construit et la validité statistique des

¹³ Ces huit paramètres sont : 1) l'historique : des événements pouvant se produire durant l'expérimentation et interférer sur les effets à mesurer, 2) la maturation : les effets du temps qui agissent sur l'individu au cours de l'expérimentation et qui influencent indirectement sa réaction vis-à-vis du dispositif, 3) l'effet de test : l'adaptation des réponses/réactions de l'individu au dispositif expérimental par la répétition des mesures effectuées durant l'évaluation, 4) les instruments : biais de mesure liés à la précision des instruments/indicateurs utilisés ou des intervenants lors de leur utilisation, 5) la régression : les groupes bénéficiant et ne bénéficiant pas du dispositif sont différents en fonction de leurs caractéristiques initiales, l'impact du dispositif pouvant être confondu par ces caractéristiques, 6) la sélection : les groupes à comparer sont constitués différemment par un effet de sélection pouvant être lié au contexte de mise en œuvre de l'expérimentation, 7) l'attrition : on observe des perdus de vue dans l'échantillon au cours de l'expérimentation, l'attrition pouvant également être lié au contexte de mise en œuvre de l'expérimentation, 8) sélection-maturation interaction : la mise en œuvre de l'expérimentation au cours du temps introduit un effet différentiel de maturation au sein des groupes ou entre les groupes destinés à être comparés lors de l'évaluation

¹⁴ Ces quatre paramètres sont : 1) l'interaction avec les tests : la sensibilité des individus au dispositif est modifiée par leur interaction avec celui-ci tout au long de l'expérimentation, l'effet mesuré ne pouvant alors pas être généralisé à la population cible, 2) la sélection : l'échantillon expérimental est différent de l'échantillon cible par un effet de sélection pouvant être lié au contexte de mise en œuvre de l'expérimentation, 3) la réaction expérimentale : l'effet mesuré est spécifique au contexte expérimental de mise en œuvre du dispositif et difficilement transposable en situation réelle, 4) l'interférence d'autres dispositifs : les effets mesurés sont confondus par l'effet d'autres dispositifs déployés au cours de l'expérimentation

résultats (Cook & Campbell, 1979; Reichardt, 2019j; Shadish et al., 2001). La validité de construit questionne la crédibilité des procédures et outils mis en place pour effectuer l'évaluation. Leur capacité à rendre compte des relations entre les dispositifs et les changements au niveau des bénéficiaires est également essentiel dans l'évaluation de l'impact du dispositif. Selon les auteurs, la notion de validité du construit est proche de celle de validité externe dans le sens où elle décrit des variations externes (tributaires des procédures et outils de mesure) de l'effet du dispositif (Shadish et al., 2001) ou proche de celle de validité interne dans le sens où elle contribue à isoler les effets nets des dispositifs (Reichardt, 2019j). La validité statistique des résultats concerne tout élément qui peut influencer la précision de l'estimation d'impact du dispositif, qu'elle concerne les données ou les méthodes utilisées pour effectuer le calcul. Elle permet de s'assurer que l'estimation obtenue reflète bien la quantité d'effet des dispositifs chez les récipiendaires.

Dans la démarche campbellienne d'analyse de la validité de l'inférence, optimiser la validité interne est essentielle dans sa capacité à garantir une estimation de l'impact cohérente avec la définition de la causalité selon une logique contrefactuelle. Dans ce cas, la validité externe expose ces estimations à de nombreuses variations contextuelles limitant l'attribution de l'effet aux dispositifs. Tout en mettant en avant l'importance de la validité interne, la nécessité de conserver l'effet des dispositifs au-delà du contexte expérimental, du moins au niveau de l'ensemble des personnes concernées par le dispositif en question a vu naître l'approche dite « top down » dans la conduite de l'évaluation (Flay, 1986; Flay et al., 2005) pour garantir les différents types de validité de l'inférence. Ainsi, la conduite idéale de l'évaluation d'impact des programmes de santé passerait par une succession d'étapes commençant par des phases pilotes pour tester l'applicabilité des dispositifs, suivi d'une phase d'expérimentation en conditions optimales garantissant la validité interne des résultats, suivi d'une expérimentation

en conditions réelles pour garantir la généralisabilité des résultats aux bénéficiaires cibles et enfin un déploiement des dispositifs à l'échelle des populations.

4. D'autres modèles causaux pour l'évaluation des interventions complexes ?

La démarche campbellienne conjuguée avec l'approche top down de l'évaluation d'impact a permis de proposer une démarche d'évaluation afin de mieux effectuer des inférences pour identifier les dispositifs les plus à même d'être déployés en population générale. La large adoption de cette démarche réside dans sa capacité à proposer des méthodes et procédures concrètes et rigoureuses pour évaluer les résultats d'un dispositif (Chen, 2015a). Cela lui a valu de contribuer à l'instauration d'une forme de hiérarchie parmi les méthodes utilisées pour fournir des preuves d'efficacité d'un dispositif dans la recherche médicale (Aguayo-Albasini et al., 2014; Cartwright & Hardie, 2012a) où les ECR se hissent au plus haut rang, comme en témoigne son utilisation en tant que bonne pratique par la *Food and Drug Administration* (Food and Drug Administration, 1992). Sa large diffusion et son adaptation dans d'autres domaines, notamment la promotion de la santé (Chen, 2015a), n'a cependant pas trouvé le même écho au point d'être considérée comme démarche de référence dans l'évaluation d'impact (de Salazar & Hall, 2009; Potvin & McQueen, 2009; Rootman et al., 2001).

L'approche top-down campbellienne n'est pas exempte de limites. Le premier reproche formulé à son égard réside dans le fait de mettre en avant la validité interne au détriment de la validité externe, donc de privilégier les expérimentations en conditions optimales plutôt qu'en conditions réelles (Cronbach & Shapiro, 1982). Bien que la possibilité d'attribuer les estimations obtenues uniquement à l'effet d'un dispositif soit séduisante, cette démarche limite la portée et l'applicabilité des résultats en situation réelle. En premier lieu, négliger la validité externe des résultats revient potentiellement à proposer des interventions qui ne sont efficaces qu'en condition expérimentale, si les expérimentations en conditions optimales ne sont pas

suivies, comme c'est souvent le cas (Sussman et al., 2006; Teutsch, 1992), d'une évaluation en conditions réelles. Les conditions expérimentales de déploiement des dispositifs, lorsqu'elles visent à garantir la validité interne, sont souvent difficilement répliquables au quotidien. Quand bien même elles permettent d'isoler l'effet des dispositifs, leur réplcation dans une certaine mesure pourrait être une des conditions *sine qua non* de la transposition des résultats et de la mise en œuvre des dispositifs en dehors du contexte expérimental, rendant cette mise en œuvre non viable. Deuxièmement, la réplcation des conditions expérimentales peut se heurter à la valeur pragmatique et opérationnelle des dispositifs sur le terrain auprès des intervenants qui la mettent en œuvre. En dernier lieu, la dissémination de dispositifs dans le monde réel rencontre des enjeux qui dépassent ceux ayant attiré à des expérimentations (Chen, 2010; Habicht et al., 1999; Joint United Nations Programme on HIV/AIDS, 2010). Ces enjeux concernent non seulement les infrastructures techniques liées à ces dispositifs mais également les moyens à disposition dans la mise en œuvre, l'adaptabilité et la cohérence des dispositifs auprès de toutes les parties impliquées dans leur déploiement.

Le deuxième reproche quant à l'adoption d'une démarche top-down campbellienne réside dans la conception de la causalité elle-même qui est dite « successionniste » (Berger, 2010; Pawson & Tilley, 1997). Cette démarche offre uniquement la possibilité de constater, par une succession régulière d'événements, si des dispositifs produisent ou non des effets sans pouvoir expliquer pourquoi et comment ces effets se réalisent. La validité interne tient une place importante par rapport à la validité externe en portant la promesse d'invariabilité et régularité de la preuve, et la rendant par la même occasion intrinsèque au dispositif seul. Cette place est importante à un point tel que l'importance du contexte et de la mise en œuvre dans le déclenchement des effets se trouve effacée, laissant peu de place à la description du mécanisme ayant abouti à l'adoption du dispositif par les bénéficiaires lors de sa mise en œuvre. Pour répondre à ces limites, certains auteurs préconisent d'ancrer l'usage des méthodes

quantitatives dans une évaluation mettant en avant les conditions réelles. Par exemple, une approche dite intégrée de l'évaluation (Chen, 2015b, 2015e; Alterman et al., 1984) permet de mettre à contribution les intervenants qui sont les premiers concernés par la mise en œuvre des dispositifs, que ça ait lieu dans un cadre expérimental ou non. Une telle approche permettrait d'intégrer des enjeux pragmatiques de mise en œuvre de l'intervention garantissant l'efficacité des dispositifs dans d'autres contextes. En effet, la participation des intervenants voire des bénéficiaires permettrait de créer un environnement propice à l'usage et l'efficacité des dispositifs que le contexte soit expérimental ou non. Chen propose même d'inverser la séquence d'évaluation proposée dans les approches top-down campbelliennes en préconisant la réalisation d'expérimentations en conditions réelles avant de tester la régularité des effets en conditions optimales. Il introduit par la même occasion la notion de viabilité (Chen, 2010) de l'intervention parmi les critères de validité d'inférence. Tout en plaçant la validité externe des interventions comme critère de validité à mettre en avant, la notion de viabilité permet de prendre en compte la valeur pragmatique de la mise en œuvre de l'intervention au regard des activités à réaliser ou des moyens à disposition. Elle se distingue des autres types de validité par l'accent qu'elle met sur la plausibilité d'intégrer des activités issues d'un contexte expérimental dans le cadre d'un dispositif ordinaire mis en œuvre par des praticiens et intervenants travaillant dans des structures ordinaires, comme une clinique communautaire ou un établissement relevant d'organisations non gouvernementales (Chen, 2015a). En ce sens, un dispositif est viable si les activités qui en relèvent sont réalisables au sein d'une organisation donnée, économiquement abordables à mettre en œuvre, économiquement attractives pour les bénéficiaires, et surtout convaincantes auprès des intervenants et des bénéficiaires.

Figure 2 : Illustration d'une approche orientée par la théorie de l'intervention basée sur l'approche dite « intervening mechanism approach »



Source : Adapté de (Chen, 2015f), Lecture : cette approche illustre la place intermédiaire des déterminants de l'efficacité de l'intervention tels que décrits dans la théorie de l'intervention

Une autre proposition consiste à adopter une approche orientée par la théorie de l'intervention (Chen, 2015f; Van Belle et al., 2010). Cette proposition étend la précédente en incluant les éléments de recontextualisation et de mise en œuvre dans l'analyse causale des effets de l'intervention. Plus précisément, cette approche intègre dans ses éléments d'évaluation les hypothèses sur lesquelles se basent la chaîne causale garantissant l'efficacité du dispositif via ses composantes (théorie du changement) et les hypothèses qui garantissent que les activités prévues sont réalisables sur le terrain étant donné le contexte de mise en œuvre (théorie de la mise en œuvre de l'intervention¹⁵) (Chen, 2015c). En plus d'estimer l'impact du dispositif sur la situation des bénéficiaires, cette approche permet de distinguer les effets propres au dispositif et ses composantes de ceux qui relèvent de l'interaction des bénéficiaires avec les intervenants et le contexte de déploiement du dispositif. En fonction des objectifs de l'évaluation, celle-ci donne plus de place à la théorie du changement¹⁶ (Figure 2), à la théorie de la mise en œuvre¹⁷ (Figure 3), ou considère les deux¹⁸ (Figure 4). Ces approches nécessitent néanmoins l'usage de méthodes supplémentaires, telles que des méthodes qualitatives, pour dépasser la conception successiviste de la causalité telle que traitée dans l'approche top-down campbellienne (Chen, 2015f).

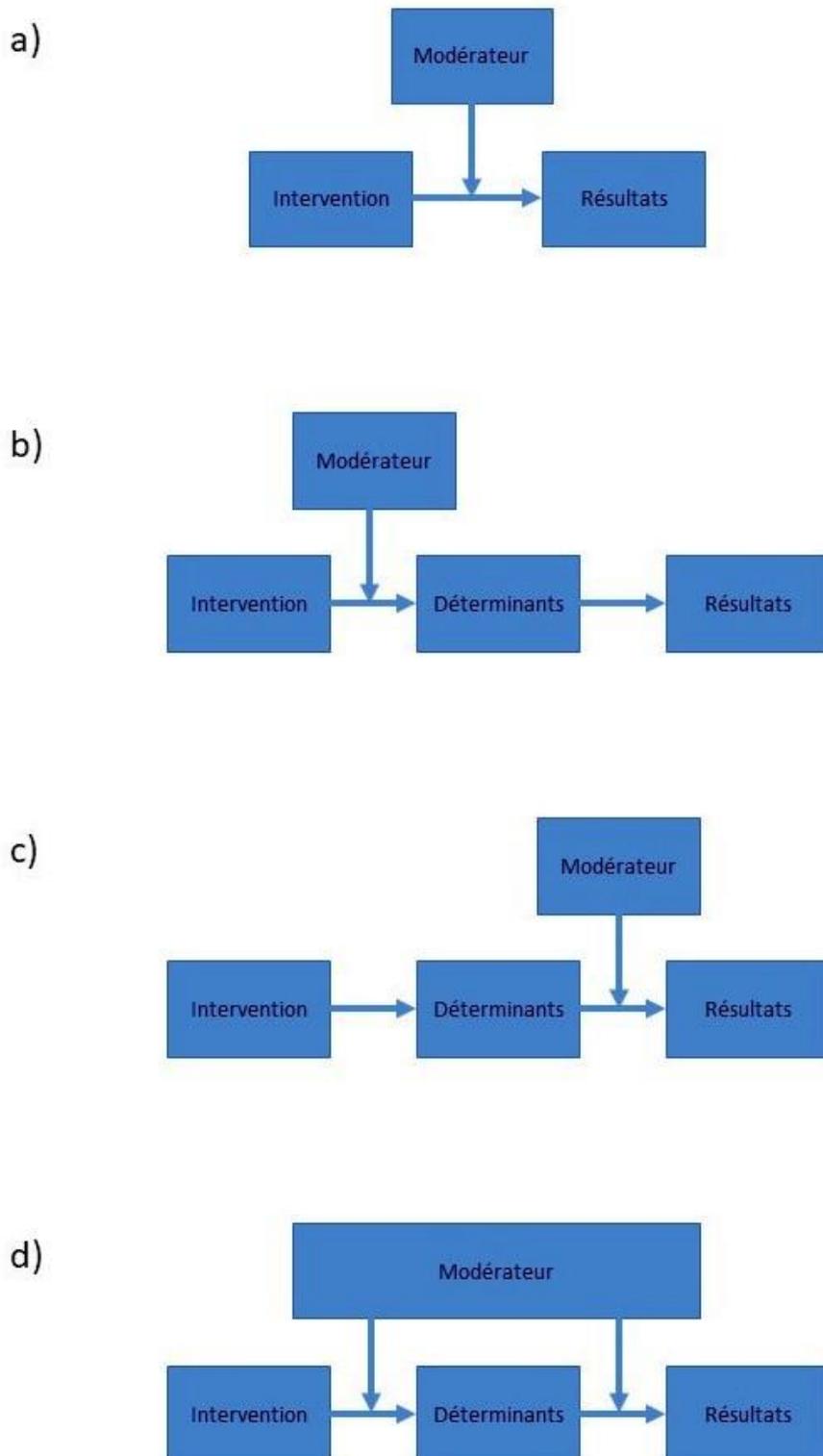
¹⁵ Nous utilisons le terme « théorie du changement » pour traduire le terme « theory of change » et « théorie de la mise en œuvre » pour traduire le terme « intervention theory ». Le terme « théorie de l'intervention » renvoie ici à la conception globale de l'intervention et peut concerner les deux précédents.

¹⁶ L'approche est alors connue sous le terme de « intervening mechanism approach »

¹⁷ L'approche est alors connue sous le terme de « moderating mechanism approach »

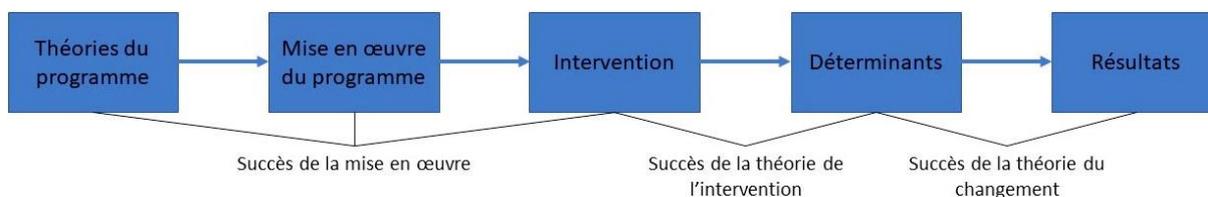
¹⁸ L'approche est alors connue sous le terme de « Integrative Process/Outcome Evaluation »

Figure 3 : Illustration d'une approche orientée par la théorie de l'intervention basée sur l'approche dite « moderating mechanism approach »



Source : Adapté de (Chen, 2015f), *Lecture* : cette approche illustre à quels moments et dans quelles mesures la mise en œuvre de l'intervention peut agir pour conditionner son efficacité. Le terme « modérateur » renvoie à l'interaction des bénéficiaires avec les intervenants et le contexte de déploiement du dispositif tels qu'une relation de confiance ou même des paramètres sociodémographiques qui expliquent l'attractivité du dispositif (Chen, 2015c).

Figure 4 : Illustration d'une approche orientée par la théorie de l'intervention basée sur l'approche dite « Integrative Process/Outcome Evaluation »



Source : Adapté de (Chen, 2015f). Lecture : cette approche prend en compte à la fois les éléments de la mise en œuvre et des composantes du programme pour illustrer la chaîne causale qui en détermine les résultats

Les variations méthodologiques citées précédemment mettent en avant la nécessité de mieux prendre en compte la capacité des programmes à donner des effets en conditions réelles pour avoir des dispositifs sensibles et adaptables au contexte de mise en œuvre. Elles préservent cependant la possibilité d’user des méthodes quantitatives empruntées aux démarches campbelliennes pour constater l’ampleur des effets en plus d’apporter des informations expliquant ces effets. Une position plus radicale consisterait à postuler que constater uniquement les effets serait futile et donner une valeur compréhensive et explicative à la causalité est la seule chose essentielle dans l’évaluation. Cette position est celle proposée dans l’évaluation dite réaliste (Pawson & Tilley, 1997). Cette approche considère la causalité dans une perspective dite « générativiste » (Berger, 2010; Pawson & Tilley, 1997) : l’expérimentation, en introduisant un nouveau dispositif, modifie l’environnement des individus de manière à favoriser les changements escomptés chez les bénéficiaires. L’expérimentation s’insère donc dans un système social plus large et doit être analysée en perspective avec l’organisation de ce système qui contribue à activer ou désactiver les effets attendus d’un dispositif. Dans ce cadre, l’analyse de la causalité consiste à expliquer comment l’interaction entre le contexte de mise en œuvre et les éléments du dispositif permettent de changer la situation des bénéficiaires. Cette interaction concerne trois éléments obligatoires ; le contexte, les mécanismes et les résultats ; et donne sens à l’effet du dispositif ainsi qu’à l’inférence causale. La notion de mécanisme renvoie ici aux processus qui ont provoqué la réaction des bénéficiaires de manière à ce qu’ils se saisissent des ressources et opportunités

proposées par le dispositif à évaluer. Elle ne se résume donc pas aux composantes du dispositif telles que décrites par la théorie du changement ou la théorie de la mise en œuvre mais intègre aussi des processus en lien avec le système social dans lequel les bénéficiaires évoluent et les interactions qui y ont lieu, et qui motivent leurs décisions et actions. La notion de contexte rappelle que les dispositifs sont déployés dans des conditions et des configurations spécifiques. Le contexte renvoie à l'identification de ces conditions et configurations dans lesquelles les mécanismes ont pu être activés. Les effets des dispositifs sont les résultats réguliers de la conjonction des mécanismes et des contextes. Selon une perspective réaliste, un dispositif produit donc des résultats lorsque les ressources et opportunités (mécanisme) proposées sont cohérentes avec les conditions sociales et culturelles (contexte) des bénéficiaires. L'usage de méthodes quantitatives dans le cadre de ce type d'évaluation ne peut prendre sens que si la causalité est considérée d'un point de vue réaliste générativiste. De ce fait, les stratégies d'analyse imposent l'usage de méthodes permettant d'intégrer l'enjeu des contextes et des mécanismes dans les inférences, telles que les approches orientées par la théorie¹⁹ ou celles permettant d'envisager la généralisation des mécanismes dans d'autres contextes (Devaux-Spatarakis, 2014; Yin, 2010, 2018).

Ces différentes approches de l'évaluation sont à l'origine de débats enflammés qui rappellent que la précaution est impérative dans la considération des liens causaux concernant l'impact des dispositifs avant de pouvoir distinguer l'usage potentiel des méthodes quantitatives en dehors des courants qui les ont vu naître et appliqués, à l'instar du cas des ECR (Bonell et al., 2016; Jamal et al., 2015; Van Belle et al., 2016). A la lueur de ces débats, on oublierait presque que la crédibilité statistique n'est qu'un critère insuffisant mais nécessaire parmi tant d'autres

¹⁹ Il s'agit ici d'une traduction de la notion de « theory-based evaluation », la notion de théorie dans ce contexte dépasserait les notions de théorie du changement et théorie de la mise en œuvre des interventions pour intégrer des éléments plus généraux relevant du système social dans lequel évolue les individus (Louart et al., 2023)

pour établir des liens causaux, y compris dans la recherche biomédicale précédant l'approche campbellienne (Hill, 1965).

5. Conclusion

L'approche expérimentale de l'évaluation trouve ses racines dans les travaux de Fisher et Neyman, mais c'est surtout Campbell qui est connu comme celui qui l'a introduit dans les sciences de l'évaluation (Christie & Alkin, 2013; Shadish & Luellen, 2013). L'apport des notions de validité externe et surtout interne ont contribué à hisser les ECR au sommet d'une hiérarchie des méthodes quantitatives mais aussi à les placer au cœur de polémiques concernant leur pertinence pour évaluer un lien de causalité entre les effets d'un dispositif et la situation des bénéficiaires (Cartwright & Hardie, 2012a; Deaton & Cartwright, 2018; Reichardt, 2019c). Les prémices et succès de leur usage dans le domaine biomédical ont rayonné de manière à convaincre cet usage au-delà de ce domaine, y compris dans les sciences sociales et dans le champ de l'évaluation. Leur usage à ce niveau se retrouve pourtant confronté aux critiques des évaluateurs y officiant, qui eux-mêmes sont confrontés aux réalités de la prise de décision et la mise en œuvre des dispositifs évalués. La vocation de l'évaluation à orienter ou justifier l'usage de dispositifs en conditions réelles a alors amené des débats quant à la bonne manière d'organiser l'évaluation et utiliser les méthodes, allant jusqu'à remettre en cause l'usage qui est fait des méthodes expérimentales et quasi-expérimentales. Paradoxalement, ces débats et les méthodes d'évaluation en général poursuivent un même objectif : fournir les preuves de la pertinence d'un dispositif pour son usage en situation réelle. La controverse est d'une ampleur telle que les méthodes se retrouvent prisonnières des approches et mouvements évaluatifs alors qu'elles pourraient bien apporter des informations pertinentes au-delà des courants qui leur sont attribués, peut-être à tort, à l'instar des méthodes randomisées (Cartwright & Hardie, 2012b; Deaton & Cartwright, 2018). En définitive, il est

important de dissocier les enjeux statistiques (et plus largement ceux inhérents aux méthodes quantitatives) des enjeux paradigmatiques et épistémologiques de l'évaluation pour tirer parti des atouts de ces méthodes selon le choix d'évaluation et le contexte dans lequel il est effectué. Il est important que les méthodes puissent être mises à profit avec précaution et de manière cohérente à partir du moment où des objectifs clairs sont énoncés, qu'ils soient intrinsèques aux enjeux de mise en œuvre ou aux enjeux scientifiques.

Chapitre – 2 : Les enjeux d'évaluation dans la lutte contre l'épidémie de VIH

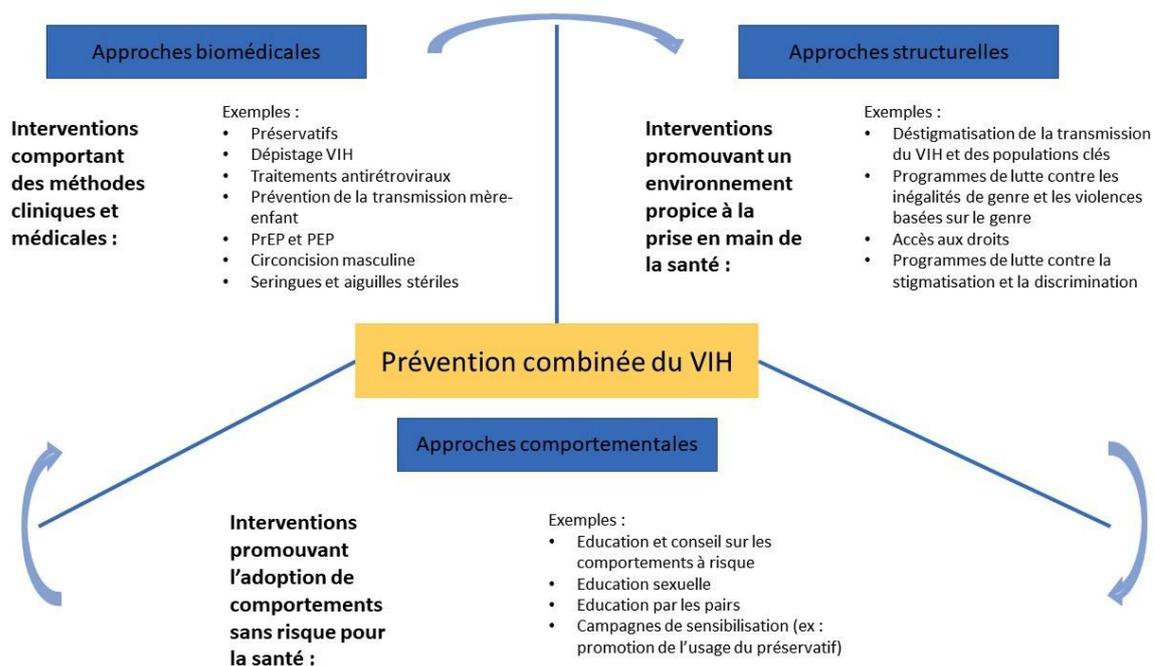
1. Vers des programmes de plus en plus complexes

Depuis l'identification du virus de l'immunodéficience humaine (VIH) responsable du syndrome d'immunodéficience acquise (SIDA) en 1983 (Barré-Sinoussi et al., 1983), les travaux scientifiques pour trouver des outils biomédicaux permettant de contrôler sa propagation se sont multipliés (Padian et al., 2010; Piot et al., 2008). Les efforts ayant amené à proposer ces outils et les efforts de prévention pour changer les comportements sexuels n'ont pas suffi pour mettre fin à cette pandémie. En effet, l'existence des outils biomédicaux, théoriquement efficaces, ne suffit pas à éradiquer le virus, témoignant de la complexité de la lutte contre la pandémie dont les enjeux échappent aux solutions purement médicales. Les dynamiques sociales et structurelles inhérentes à la lutte contre la propagation du virus ont poussé à rechercher des solutions qui prennent en compte la difficulté à mettre en œuvre les programmes. L'efficacité des programmes de lutte contre le VIH relève de leur capacité à soutenir un système ou un environnement de manière à favoriser l'acceptabilité et l'adoption de comportements limitant la transmission du virus à travers les pratiques individuelles et sociales, ainsi que l'usage des outils biomédicaux. Bien que de tels efforts aient été préconisés depuis longtemps (Joint United Nations Programme on HIV/AIDS, 2007, 2015; Kurth et al., 2011; Piot et al., 2008), les effets de la pandémie de COVID sur la lutte contre le VIH rappellent que ces objectifs n'ont pas été réalisés et appellent à réitérer les efforts, en mettant notamment l'accent sur les inégalités socio structurelles (Joint United Nations Programme on HIV/AIDS, 2021). Encore aujourd'hui, le fardeau du VIH est tel qu'en 2022, 39 millions de personnes vivent avec le VIH dans le monde, 5.5 millions ne savaient pas qu'elles étaient

infectées, et 630000 personnes sont décédées des conséquences de l'infection (Joint United Nations Programme on HIV/AIDS, 2023). Nous vivons donc une époque où cette épidémie demeure un réel enjeu de santé publique sur le plan mondial : les solutions quant à son éradication ne peuvent être que communes à bien des égards. C'est pourquoi, la prévention combinée du VIH (Figure 5) est la voie la plus à même d'aboutir à une maîtrise de l'épidémie en combinant les avantages des outils biomédicaux, des approches comportementales et des approches structurelles. La mise en place de tels programmes est cependant un défi car ils ne consistent pas uniquement à une simple addition de ces différentes composantes mais surtout leur synergie, leur gestion, leur déploiement de manière cohérente au contexte et adaptée aux populations. Ces programmes de promotion de la santé (Piot et al., 2008) visent à adopter une approche dynamique, fondée sur les droits, pour fournir la bonne combinaison d'interventions biomédicales, comportementales et structurelles visant à fournir un effort soutenu et durable pour réduire les nouvelles infections par le VIH²⁰ (Joint United Nations Programme on HIV/AIDS, 2007, 2010, 2015).

²⁰ Traduit de l'anglais où la définition est la suivante « a dynamic, rights-based approach to providing the right mix of biomedical, behavioral and structural interventions aiming to have the greatest, sustained effort on reducing new HIV infections »

Figure 5 : Description des approches combinées de la prévention du VIH



Source : Adapté de <https://www.paho.org/en/topics/combo-hiv-prevention>

a. Les approches biomédicales de la prévention combinée

De nos jours, les outils biomédicaux de la prévention combinée offrent plusieurs options permettant d'identifier et traiter les personnes infectées (tests de diagnostic VIH), de diminuer la transmission du VIH lors de rapports sexuels (préservatifs, circoncision) ou lors de l'utilisation de drogues injectables (matériels d'injection), ou permettant de freiner sa prolifération dans l'organisme (traitements antirétroviraux, prophylaxie pré exposition ou PrEP, prophylaxie post exposition ou PEP). Bien que cette diversité semble de bon augure pour la lutte contre le VIH, le déploiement de ces outils au sein des populations qui peuvent et devraient s'en approprier demeure un défi. L'efficacité théorique prometteuse de ces outils ne se retranscrivent pas en pratique (Padian et al., 2008; Piot et al., 2008). La dissémination de ces outils au sein des populations rencontre différents niveaux de barrières non propices à leur utilisation. Lorsqu'il ne s'agit pas du manque de connaissances ; l'acceptabilité et l'adoption de ces outils par les concernés ne rencontrent pas les objectifs formulés par les programmes

(Joint United Nations Programme on HIV/AIDS, 2007, 2010, 2015). Et quand bien même ils sont utilisés, on remarque une utilisation suboptimale et une adhérence insuffisante (Padian et al., 2008). L'usage des outils biomédicaux présente donc des enjeux sociaux et individuels qui mettent au défi les intervenants quant à leur déploiement, encourageant l'adjonction d'approches comportementales et structurelles.

b. Les approches comportementales de la prévention combinée

Les approches comportementales ont le potentiel de bloquer par différentes configurations la chaîne causale amenant à la propagation du virus. Ces approches reposent sur l'éducation et le conseil de manière à permettre aux personnes d'éviter les comportements à risque et élaborer des stratégies personnelles ou communautaires pour faire face aux vulnérabilités individuelles liées au VIH (Bonell & Imrie, 2001; Coates et al., 2008a). Ces approches sont loin d'être faciles à mettre en œuvre en plus d'être insuffisantes (Coates et al., 2008b) et se heurtent aux dynamiques socioculturelles du contexte de leur déploiement. Pour pouvoir bénéficier de leur potentiel, il est pourtant nécessaire que les changements soient francs. Ainsi, les changements escomptés ne concernent pas que les comportements à risque ou les situations de vulnérabilité mais se doivent d'avoir un impact sur l'usage des outils biomédicaux et des infrastructures de soins. Peu de programmes basés sur ce type d'approche ont démontré une efficacité de manière à pouvoir promouvoir leur usage. Il n'en demeure pas moins que leur usage est recommandé parmi l'arsenal de stratégies de lutte contre le VIH ne serait-ce que pour améliorer l'adoption des outils biomédicaux.

c. Les approches structurelles de la prévention combinée

Les approches structurelles sont utilisées pour améliorer en termes social, économique, et politique l'environnement dans lequel les personnes évoluent pour faciliter la prise en charge et la prise en main de leur santé (Blankenship et al., 2006; Parkhurst, 2014). Ce type

d'approche, moins utilisée que les deux précédentes (Globerman et al., 2017; Iskarpatyoti et al., 2018; Joint United Nations Programme on HIV/AIDS, 2007, 2010), est pourtant pertinente et essentielle pour prendre en compte la réalité des contextes de mise en œuvre. Elles permettent de prendre en compte le contexte et les dynamiques propres au contexte en question en agissant sur le système social, politique et économique qui conditionne les décisions des individus dans leur environnement. L'ambition de ce type d'approche la rend difficile à concevoir, à mettre en œuvre et à évaluer car elle suppose l'implication d'acteurs issus de divers secteurs pour agir à différents niveaux dans le système ou l'environnement de mise en œuvre des programmes qui en résultent.

d. Combiner les approches biomédicales, comportementales et structurelles

Chacune des composantes de la prévention combinée du VIH présente des enjeux propres quant à leur dissémination. La combinaison de ces approches pour lutter contre l'épidémie imbriquent ces différents enjeux tout en en faisant naître de nouveaux qui sont liés à la nécessité d'une synergie entre ces approches. En effet, penser l'approche combinée de la prévention du VIH comme une somme de ces composantes est insuffisante. Une telle approche implique une véritable gestion commune de ces composantes de manière à tirer profit des avantages qu'elles procurent potentiellement en faisant attention à ce que leurs limites ne se nourrissent pas entre elles de manière à entraver la mise en œuvre et l'efficacité des programmes de lutte contre le VIH. Il s'agit là d'un défi complexe tant sur la conception et la planification, que la mise en œuvre, ou le suivi et l'évaluation. En ce qui concerne l'évaluation, on observe de plus en plus d'exemples de succès de mise en place de stratégies combinées (Joint United Nations Programme on HIV/AIDS, 2007, 2010), mais en nombre encore insuffisant étant donné le besoin d'informations à ce sujet pour les acteurs impliqués dans la lutte contre cette pandémie. L'évaluation et l'accumulation d'exemples de PPCV

effectifs est plus que nécessaire pour orienter ces acteurs dans la conception, le choix de mise en œuvre et l'adaptation de stratégies potentiellement efficaces pour maîtriser la propagation du virus dans leur contexte.

2. Recommandations en termes d'évaluation

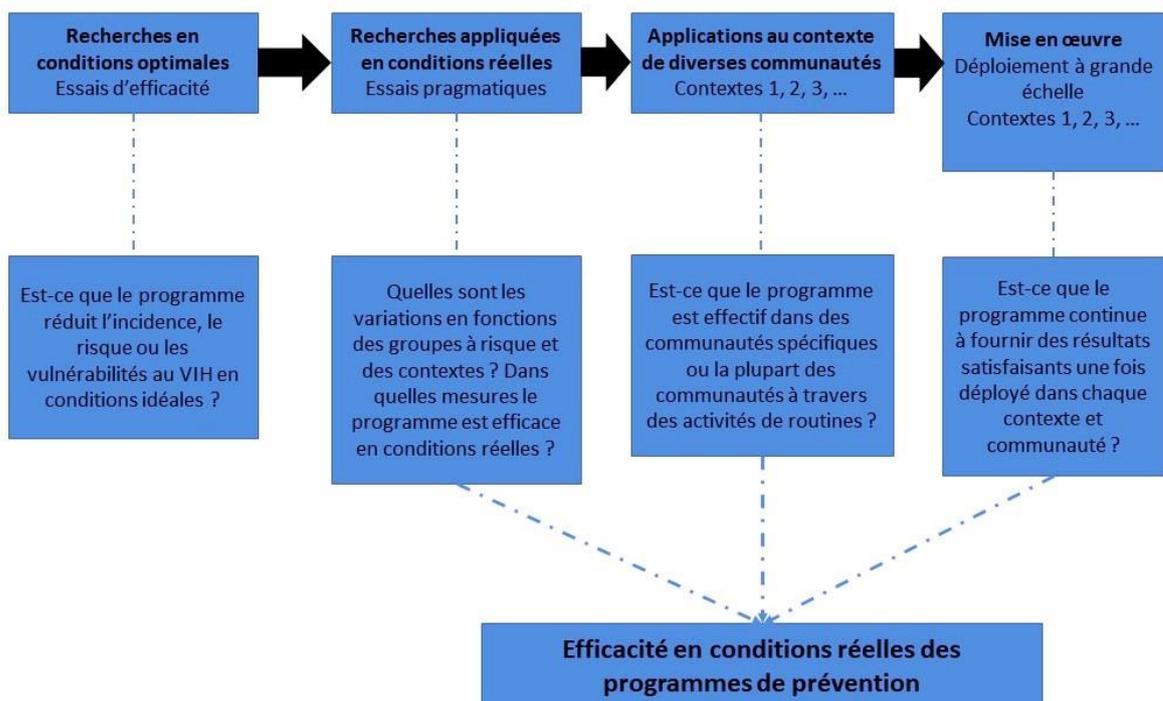
La complexité des PPCV s'avère être un défi pour leur évaluation. Si l'évaluation a pour objectif de justifier une stratégie au détriment d'une autre, elle doit pouvoir mettre en exergue la pertinence des différentes composantes des interventions pour diminuer le risque d'infection. La multiplicité des composantes et leur interaction compliquent l'évaluation de ces programmes en conditions expérimentales, rendant encore plus difficiles l'inférence et l'adaptation des résultats en conditions réelles. En effet, la transposition de la spécificité de chaque composante d'un programme ainsi que leur synergie en dehors du contexte d'expérimentation ou à l'échelle de la population bénéficiaire ne peut être garantie si l'évaluation n'est pas menée correctement. Le besoin de résultats opérationnels et transposables, au moins au niveau de la population à qui est destinée un programme, nourrit les critiques quant à l'usage de méthodes classiques telles que les ECR, et prône l'utilisation de méthodes permettant une meilleure compréhension du fonctionnement du programme.

S'il est nécessaire de vérifier la crédibilité des interventions de leur conception à leur mise à l'échelle, la démarche top-down campbellienne utilisée également dans l'évaluation des programmes de prévention, y compris du VIH (Figure 6), peut être remise en question ou au moins complétée par des informations permettant de comprendre l'efficacité de ces programmes auprès des populations dans leur contexte. Les recommandations actuelles n'orientent donc pas vers une méthode ou un paradigme particulier de l'évaluation ; elles orientent plutôt vers les questions auxquelles il faut apporter des réponses pour que les résultats d'une évaluation soient exploitables :

- Quelle est la chaîne causale proposée dans l'intervention ?
- Est-ce qu'une évaluation est faisable ?
- Quels indicateurs et méthodes choisir ?
- Comment évaluer non seulement l'impact mais aussi les déterminants de la mise en œuvre ?
- Comment mettre à profit les résultats de l'évaluation pour contribuer aux décisions ?

Au final, le choix des méthodes doit être orientée en fonction des ressources, des acteurs impliqués, du contexte et des objectifs d'évaluation, aucune méthode ne pouvant être supérieure à une autre qu'en fonction de ces contextes et ces objectifs (Joint United Nations Programme on HIV/AIDS, 2010; Julnes & Rog, 2007).

Figure 6 : Illustration de l'utilisation de l'approche campbellienne dans les programmes de prévention



Source : Adapté de (Joint United Nations Programme on HIV/AIDS, 2010)

3. Le projet MAKASI : un exemple concret de programme de prévention combinée du VIH

Les paragraphes précédents présentent de manière théorique le défi de combiner approches biomédicales, comportementales et structurelles dans la lutte contre le VIH. Le cas du projet MAKASI permet d'illustrer l'opérationnalisation de ces approches dans le cadre d'une recherche interventionnelle et d'en présenter les défis, notamment ceux portant sur l'évaluation d'impact. Cette étude propose une intervention de prévention du VIH destinée aux immigrés originaires d'Afrique subsaharienne vivant en France. A travers une démarche d'*empowerment*, elle propose d'intégrer les approches biomédicales, comportementales et structurelles en orientant les bénéficiaires de manière à ce qu'ils puissent se saisir des ressources et dispositifs existant dans le système de santé et social. Plus précisément, des médiatrices de santé sont chargées de réaliser un entretien personnalisé avec les bénéficiaires pour les aider à hiérarchiser et répondre à leurs besoins, que ces réponses soient d'ordre biomédical, comportemental ou structurel. Cette prise en charge globale vise à leur fournir les éléments nécessaires pour se protéger dans les interactions sexuelles et relationnelles. Pour mettre en œuvre l'intervention, le projet MAKASI tire parti de la collaboration entre des chercheurs et des acteurs de terrains provenant d'associations communautaires et non gouvernementales à travers une recherche participative. Les défis propres aux PPCV sont ainsi présents dans le cas du projet MAKASI qu'ils soient relatifs à la conception de l'intervention ou à sa mise en œuvre, il en va de même pour l'évaluation de ses effets au niveau des bénéficiaires.

4. Hypothèses et objectifs de la recherche doctorale

Ce travail de thèse part de l'hypothèse selon laquelle l'usage des méthodes doit être adaptée en fonction des besoins des programmes de santé. Au-delà des débats sur les différents

courants de l'évaluation, il existe une réelle question méthodologique quant à la mobilisation des méthodes quantitatives. Dans quelles mesures peuvent-elles être mobilisées de manière à répondre aux interrogations et objectifs propres à l'évaluation dans son contexte ? Cette question est particulièrement importante dans l'évaluation d'impact des programmes en promotion de la santé. Les principes et enjeux des programmes élaborés dans ce champ sont en demande de méthodes d'évaluation permettant de prendre en compte le contexte, les besoins et la participation des acteurs, y compris les communautés premières concernées par ces programmes.

Cette thèse a pour objectif de comprendre dans quelles mesures les méthodes quantitatives telles qu'elles sont utilisées actuellement peuvent être mises à profit pour mieux prendre en compte les contextes des interventions dans un objectif opérationnel de prise de décision. Nous posons l'hypothèse que la mobilisation de ces méthodes répond à des enjeux propres à la méthodologie statistique et mathématique, et aux objectifs des programmes évalués. Nous choisissons le cas de l'évaluation des PPCV pour explorer ces hypothèses car la lutte contre le VIH et les recommandations concernant les programmes à mettre en place et leur évaluation sont cohérentes et propices à l'analyse de ces problématiques. Plus précisément, nous effectuons une étude de cas à partir d'un PPCV, l'intervention MAKASI. Cette thèse repose sur 4 étapes :

- Etape 1 : recenser, décrire et discuter les méthodes actuellement mobilisées pour évaluer l'impact des programmes de prévention combinée contre le VIH
- Etape 2 : effectuer une étude de cas concernant la mobilisation de méthodes quantitatives dans l'évaluation d'un programme de prévention combinée contre le VIH (l'intervention MAKASI)

- Etape.3 : discuter les enjeux de mobilisation des méthodes dans l'évaluation de l'intervention MAKASI
- Etape 4 : discuter l'ensemble des résultats au regard de la littérature sur l'évaluation, des méthodes mobilisées et des résultats de l'étude de cas

Les chapitres qui suivent sont dédiés à la présentation des résultats de ces étapes.

Partie – II : Revue critique des méthodes utilisées pour évaluer l'impact des programmes de prévention combinée sur la propagation du VIH

« The vital point is that this discipline provides tools, patterns of thought, and habits of heart that will allow you to deal with data with integrity. At its core statistics is not about cleverness and technique, but rather about honesty. Its real contribution to society is primarily moral, not technical. It is about doing the right thing when interpreting empirical information. »

(Le point essentiel est que cette discipline fournit des outils, des modes de pensée, et des habitudes de cœur qui vous permettront de traiter les données avec intégrité. Au fond les statistiques ne relèvent pas de l'ingéniosité ou de la technique, mais plutôt de l'honnêteté. Leur réelle contribution pour la société est avant tout morale, non technique. Il s'agit de faire les choses correctement pour interpréter les informations empiriques. [Traduction libre])

Stephen Bruce Vaderman et Max Morris, *Statistics and Ethics: Some Advice for Young Statisticians*, 2003

Chapitre – 3 : La difficulté et la nécessité de comparer les différentes méthodes quantitatives mobilisées pour évaluer l’impact des programmes de prévention combinée du VIH

Ce chapitre décrit les procédures qui ont été mises en place pour pouvoir recenser et comparer les méthodes quantitatives utilisées pour évaluer l’impact des PPCV. Ce travail a fait l’objet d’une publication dans la revue *Systematic Reviews* en 2022. La section suivante présente le résumé de ce travail, l’article est présenté en Annexe 1.

Résumé de l'article 1

Afin de comprendre dans quelles mesures les méthodes quantitatives peuvent être mises à profit pour prendre en compte les contextes des interventions, il est nécessaire de savoir quelles sont ces méthodes et quelles sont leurs spécificités. Recenser, décrire et discuter les méthodes actuellement mobilisées pour évaluer l'impact des PPCV suppose que ces méthodes puissent être comparées. Cela signifie qu'il existe un certain nombre de critères :

- Permettant d'identifier les études d'évaluation d'impact des PPCV sur la transmission du VIH auprès des populations concernées et utilisant des méthodes quantitatives
- Communs à ces méthodes et permettant de les comparer pour faire sortir les avantages et inconvénients spécifiques à chacune

Nous avons recensé et choisi ces critères à travers l'élaboration d'un protocole pour une revue systématique des méthodes quantitatives utilisées pour mesurer l'impact des PPCV sur la transmission du VIH. Trois étapes ont été nécessaires pour réaliser ce protocole. La première consistait à identifier les bases de données et les équations spécifiques à appliquer sur ces bases pour recenser de manière exhaustive les études qui nous intéressent. Grâce au concours de documentalistes, nous avons pu établir ces équations de recherche considérant trois bases de données suffisantes pour effectuer nos recherches : *Web of Science*, *Scopus* et *PubMed*.

La deuxième consistait à identifier les guides et bonnes pratiques publiés dans des articles scientifiques concernant les méthodes quantitatives mobilisées pour réaliser une évaluation d'impact de programmes similaires aux PPCV. Pour ce faire, des recherches bibliographiques ont été effectuées en utilisant les bases de données de *Web of Science*, *Scopus* et du réseau EQUATOR²¹. A l'issue de ces recherches, nous avons identifié trois types de critères : ceux

²¹ <https://www.equator-network.org/>

qui permettent de caractériser les plans d'étude et les plans expérimentaux, ceux qui permettraient de caractériser les méthodes d'estimation statistiques, et ceux qui permettraient de caractériser les méthodes de simulation et d'estimation mathématiques.

La troisième étape consistait à élaborer une grille d'analyse permettant d'agréger les critères identifiés pour recenser et comparer les méthodes utilisées pour évaluer l'impact des PPCV.

A l'issue de cette phase, nous avons pu créer une grille de lecture commune permettant d'effectuer une analyse critique des méthodes quantitatives d'évaluation d'impact des programmes de santé.

Cette grille est celle utilisée pour la réalisation de la revue systématique des méthodes quantitatives utilisées pour mesurer l'impact des PPCV telle que présentée dans le chapitre suivant.

Chapitre – 4 : Quelles méthodes quantitatives pour évaluer l’impact des interventions dans la lutte contre le VIH ?

Ce chapitre décrit et compare les méthodes quantitatives utilisées actuellement pour évaluer l’impact des PPCV. Ce travail a fait l’objet d’une publication dans la revue *AIDS and Behavior* et d’une communication orale à la conférence *AIDSImpact* en 2023. La section suivante présente le résumé de ce travail, l’article est présenté en Annexe 2.

Résumé de l'article 2

Les recommandations encourageant l'usage de la prévention combinée pour combattre l'épidémie de VIH voit naître des programmes de santé de plus en plus complexes. La complexité inhérente à la conception de ces programmes se retranscrit également dans l'évaluation de leur impact sur la transmission du VIH. Des recommandations ont été formulées quant à l'évaluation de l'impact de ces programmes sur la transmission du VIH, cependant une revue systématique permettrait d'analyser dans quelles mesures elles sont respectées. Ce type de revue de la littérature permet également d'analyser comment ces méthodes ont été mobilisées et comment leur usage peut être amélioré.

Grâce à la stratégie de recherche que nous avons présentée dans le chapitre précédent, nous avons mené une revue de la littérature concernant les méthodes quantitatives utilisées pour l'évaluation d'impact des PPCV. Dans la revue que nous avons menée, nous avons identifié 58 articles scientifiques concernant l'évaluation quantitative d'impact des PPCV sur la transmission du VIH. Ces articles concernent l'évaluation de 46 PPCV. La plupart des études concernaient les travailleuses de sexe (39.7%), ont été publiées à partir de l'année 2001 (94.8%), et ont été effectuées en Asie (36.2%) et en Afrique subsaharienne (44.8%). Moins de la moitié des études ont conclu à une diminution significative de la transmission du VIH (49.1% dans les études où la transmission du VIH était un critère de jugement principal et 40.0% dans les études où la transmission du VIH n'était pas un critère de jugement principal). Les études menées en situation réelle ont plus souvent conclu à une réduction significative de la transmission du VIH (58% comparé à 31% dans les études en conditions optimales, c'est-à-dire des conditions facilitant la démonstration de l'efficacité des dispositifs). La plupart des études utilisaient un plan quasi-expérimental (62.1%) et des méthodes statistiques pour évaluer l'impact des programmes (87.9%). Les méthodes statistiques étaient décrites de manière plus

précise dans les études expérimentales randomisées. Seules des études utilisant un devis quasi-expérimental (15.5%) présentaient des résultats de méthodes de simulation mathématique.

A la lueur de nos résultats, il est possible d'affirmer que l'évaluation d'impact des PPCV met à profit diverses méthodes tel que recommandé par le programme commun des Nations Unies sur le VIH/Sida. Dans cette diversité, certaines études mobilisent des méthodes basées sur la théorie du changement des interventions ou utilisaient des variables décrivant la mise en œuvre de programmes, donnant ainsi une portée explicative à l'évaluation d'impact. Ces résultats mettent également en exergue le fait qu'un nombre important d'études d'évaluation d'impact de PPCV ne présentaient pas les résultats de manière assez précise : seules respectivement 1.7%, 5.9%, et 22.2% des études donnaient des informations précises concernant les plans d'études, la mobilisation des méthodes d'analyse statistique et la mobilisation des méthodes d'analyse mathématique. L'adoption de bonnes pratiques de présentation de ces résultats par la communauté scientifique est donc une nécessité. Enfin, afin d'améliorer la réalisation d'études d'impact ayant une meilleure portée explicative des mécanismes d'action des programmes, la réalisation de ces études doivent tenir compte des bénéficiaires et des objectifs de ces programmes, prendre en compte la réalité des mécanismes par lesquelles les données ont été produites, et prendre en compte des paramètres qui décrivent la mise en œuvre des PPCV, en particulier la fidélité de la mise en œuvre vis-à-vis des théories utilisées pour concevoir ces programmes.

Partie – III : Étude de cas de la mise en œuvre de l'évaluation dans le cas de l'intervention MAKASI

« Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise. »

(Il est bien préférable d'avoir une réponse approximative à la bonne question, qui est souvent vague, qu'une réponse exacte à la mauvaise question, qui peut toujours être plus précise. [Traduction libre])

John Wilder Tukey, *The Future of Data Analysis*, 1962.

Chapitre – 5 : L'étude MAKASI

1. Contexte de l'étude

La situation épidémiologique de l'infection au VIH en France identifie la population immigrée originaire d'Afrique subsaharienne comme étant le deuxième groupe le plus touché par l'infection VIH et en fait une population clé dans la lutte contre cette épidémie (Santé Publique France, 2017). L'étude ANRS Parcours réalisée entre 2012 et 2013 en Île de France a montré que 35% des personnes nées en Afrique subsaharienne et suivies pour une infection VIH en Île-de-France ont été infectées après leur arrivée (Desgrées-du-Loû et al., 2015). Cette étude a également permis de mettre en évidence des liens significatifs entre les situations de précarité vécues après l'arrivée en France et l'exposition au risque de l'infection par le VIH (Desgrées-du-Lou et al., 2016). Le projet de recherche MAKASI s'inscrit dans la continuité de l'étude ANRS Parcours en proposant une intervention sociale dans le but de réduire les expositions à l'infection au VIH chez les IASS en Île-de-France vivant des situations de précarité. L'intervention proposée entend agir sur la capacité des IASS vivant en Île-de-France à agir et se saisir des ressources disponibles via le système de soins pour la prévention de l'infection. L'intervention se base sur le concept d'*empowerment* qui peut être défini ici comme un processus psychosocial promouvant la participation et la capacité d'agir des personnes pour améliorer le contrôle de leur santé (Wallerstein, 1992).

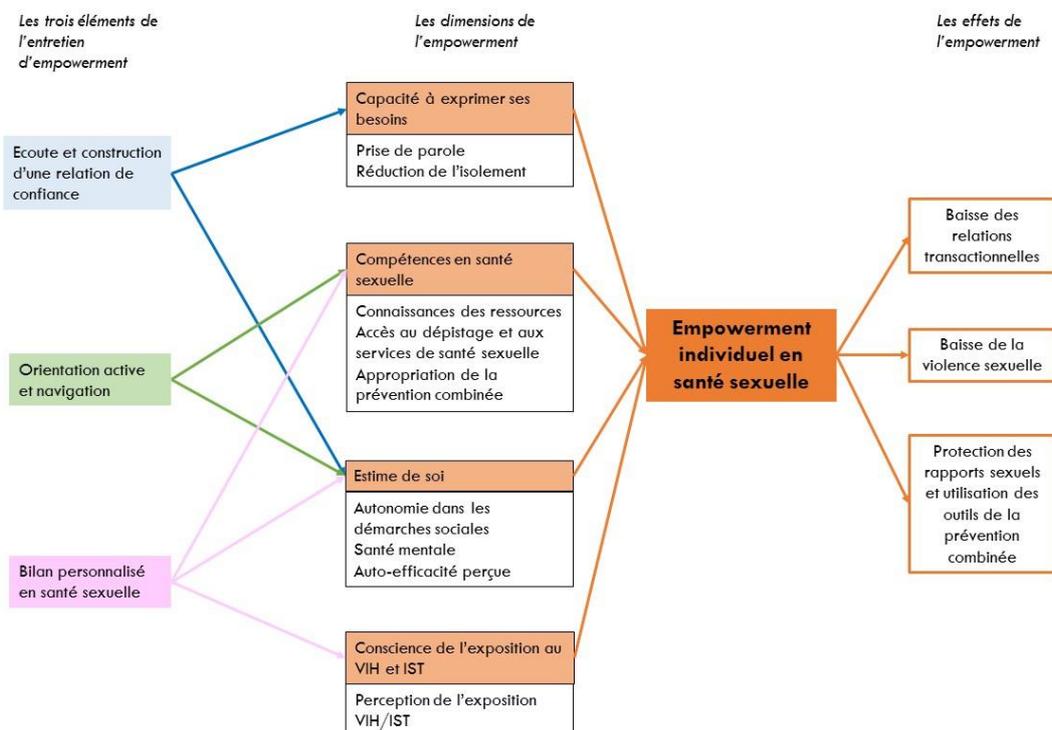
2. Conception et mise en œuvre de l'intervention

a. Théorie du changement : les composantes actives de l'intervention

L'intervention MAKASI (Gosselin et al., 2019) propose d'améliorer l'*empowerment* individuel en santé sexuelle (EISS) par un entretien motivationnel (Miller & Rollnick, 2013)

unique. Cet entretien permet d’orienter les immigrés pour améliorer leur autonomie par rapport aux outils et dispositifs existant dans le système de soins et leur capacité de protection dans les interactions sexuelles et relationnelles. La réalisation de cet entretien est censé améliorer la situation des individus en agissant sur les quatre dimensions de l’*empowerment* individuel tels que proposés dans le cadre théorique de William Ninacs (Ninacs, 2003) : i) la participation (capacité à exprimer ses besoins), ii) les compétences et les aptitudes (compétences en santé sexuelle), iii) l’estime de soi et iv) la conscience critique (conscience de l’exposition au VIH et aux infections sexuellement transmissibles (IST)). L’effet attendu de l’*empowerment* au niveau individuel réside dans une meilleure prise en main de la santé sexuelle au point d’observer une diminution de l’exposition aux risques sexuels d’infection par le VIH (Figure 7).

Figure 7 : Théorie du changement de l'intervention MAKASI

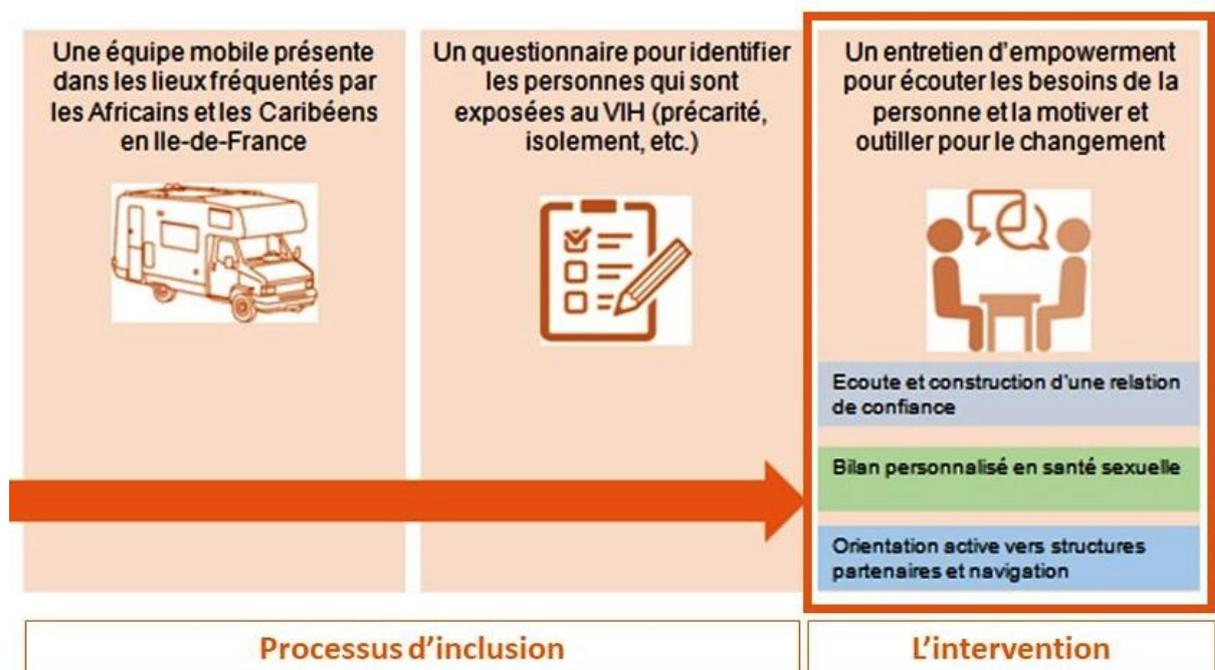


Source : projet MAKASI

b. Théorie de la mise en œuvre de l'intervention

L'intervention proposée par le programme de recherche MAKASI s'insère dans une action de sensibilisation à la santé sexuelle et d'offre de test rapide d'orientation diagnostique du VIH menée par l'équipe mobile d'une association, Afrique Avenir, sur des lieux stratégiques de passage (marchés, places, gares de RER...). Les intervenants de l'association Afrique Avenir repèrent à l'aide d'un questionnaire les personnes à qui l'intervention est destinée et leur propose un entretien personnalisé avec une médiatrice de santé, afin de les aider à hiérarchiser leurs besoins, et à les orienter dans le système de santé et d'aide sociale (Figure 8).

Figure 8 : Théorie de la mise en œuvre de l'intervention MAKASI



Source : projet MAKASI

3. Matériels et méthodes

a. Population d'étude

L'intervention est destinée aux personnes âgées de 18 ans et plus, nées dans un pays d'Afrique subsaharienne et de nationalité de l'un de ces pays à la naissance. Pour être éligibles, elles ne

vivent pas avec le VIH au moment de l'enquête et sont concernées par au moins une des situations de vulnérabilité suivantes : l'absence de logement personnel stable, l'absence d'activité rémunératrice, l'isolement social, l'irrégularité de la situation administrative, l'absence de protection sociale en santé, le fait de ne pas accéder à une consultation médicale, le fait de subir des violences.

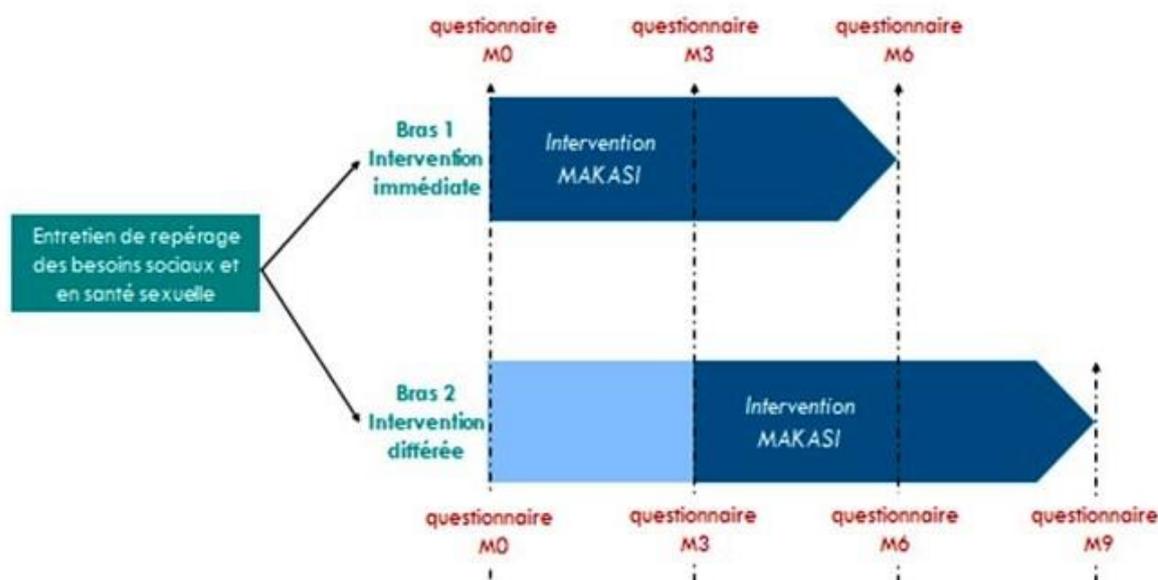
b. Plan d'étude d'évaluation d'impact

Le schéma expérimental adopté est une randomisation avec permutation séquentielle (Hemming et al., 2015, 2018). Les personnes éligibles acceptant de participer à l'expérimentation sont réparties aléatoirement selon le jour de recrutement dans un groupe intervention et un groupe contrôle. Les personnes recrutées dans le groupe intervention bénéficient immédiatement de l'entretien motivationnel tandis que celles du groupe contrôle en bénéficient de manière différée 3 mois après le recrutement. Toutes les personnes incluses dans le programme de recherche sont suivies pendant 6 mois à partir de l'entretien motivationnel (Figure 9). Certains indicateurs sont observés lors de l'entretien motivationnel, et ensuite 3 mois et 6 mois après l'entretien motivationnel. La comparaison entre le groupe intervention bénéficiant de l'entretien motivationnel et le groupe contrôle n'en bénéficiant pas immédiatement permettra d'évaluer l'impact immédiat de l'intervention grâce aux indicateurs récoltés dans les trois premiers mois. Les indicateurs récoltés lors du suivi des participants à l'étude permettront de voir l'évolution à moyen terme de l'impact de l'intervention.

Les critères principaux permettant d'évaluer l'impact de l'intervention sont : la réduction des rapports sexuels occasionnels non protégés dans les trois derniers mois, la réduction des relations transactionnelles dans les trois derniers mois et la réduction des rapports sexuels forcés dans les trois derniers mois. Les quatre dimensions de l'*empowerment* constituent un critère secondaire de l'impact de l'intervention et sont évalués grâce à des indicateurs basés

sur la perception des participants (possibilité d'exprimer ses besoins en présence d'un professionnel, isolement, perception du risque d'exposition au VIH) et des indicateurs objectifs (échelle de mesure de la littératie en santé, échelle de mesure de la dépression, score d'efficacité personnelle en santé sexuelle, proportion de personnes ayant des droits ouverts à la couverture maladie, proportion de personnes ayant un logement stable ou ayant entamé des démarches, proportion de personnes ayant un titre de séjour ou ayant déposé une demande, couverture des besoins en santé sexuelle).

Figure 9 : Schéma expérimental du projet MAKASI

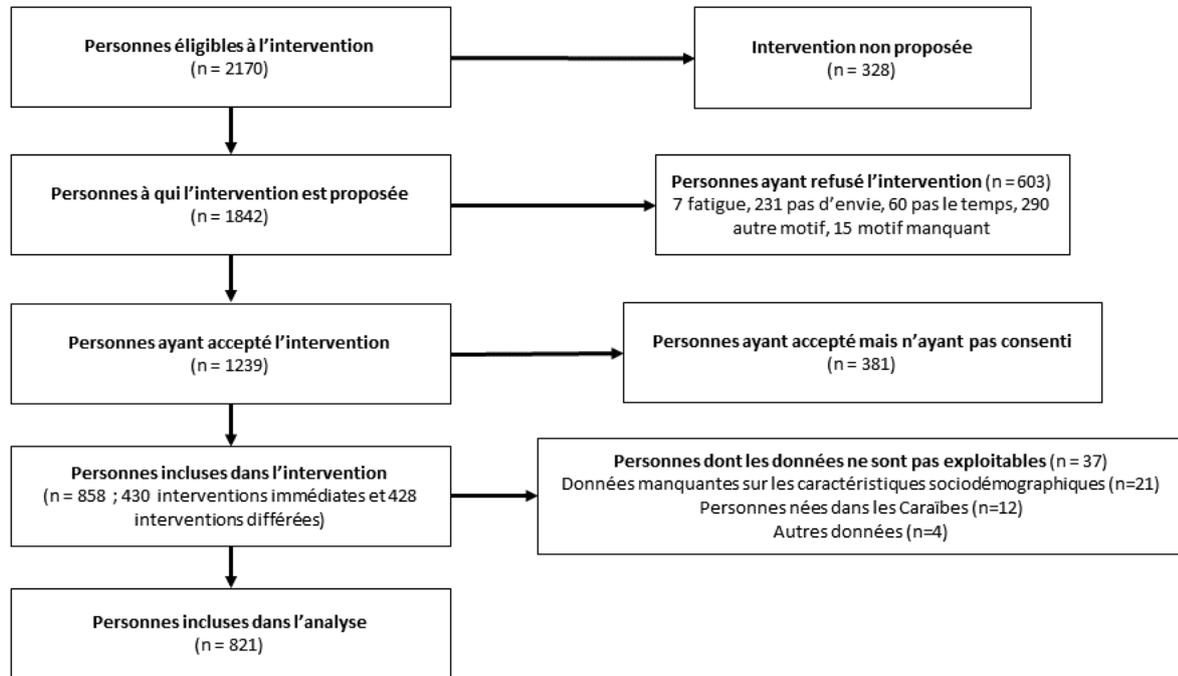


Source : projet MAKASI

Un échantillon d'étude composé de 1200 personnes (600 dans le groupe intervention et 600 dans le groupe contrôle) permettrait de mettre en évidence une réduction de 15% de rapports occasionnels non protégés dans le groupe intervention pour un risque de première espèce à 5% et une puissance à 80%. A la fin de la phase de constitution de l'échantillon, parmi 2170 personnes éligibles à l'étude, 821 (406 dans le groupe intervention et 415 dans le groupe contrôle) parmi 858 (430 dans le groupe intervention et 428 dans le groupe contrôle) personnes

ayant accepté de participer à l'étude ont fourni des données complètes permettant d'analyser l'intervention (Figure 10).

Figure 10 : Diagramme de flux de la participation à l'étude MAKASI



Source : projet MAKASI

c. Données récoltées

Des informations concernant les caractéristiques sociodémographiques et les critères d'inclusion (âge, lieu de naissance, exposition aux risques sexuels, précarité et isolement social, accès aux services de santé et de prévention) sont récoltées sur tous les éligibles repérés dans le cadre des actions menées par l'équipe mobile de l'association Afrique Avenir. Des données sur les pratiques contraceptives et la sexualité, les connaissances des ressources sociales et de santé, les compétences en santé sexuelle, *l'empowerment* et le recours aux soins durant les trois derniers mois sont récoltées sur les participants au moment du recrutement dans le programme de recherche. Des données sur la situation sociale et de santé sont récoltées pendant le suivi des participants au moment de l'entretien motivationnel, et ensuite 3 mois et 6 mois après l'entretien motivationnel.

d. Analyses réalisées dans le cadre de la thèse

A travers cette étude de cas, nous mobilisons des méthodes quantitatives dans le cadre de l'évaluation de l'impact de l'intervention MAKASI en tenant compte des objectifs de l'intervention. Plus précisément, nous mobilisons ces méthodes de manière à prendre en compte la notion d'*empowerment* tel que défini dans l'intervention et les vulnérabilités sociales et structurelles tels que considérés dans le contexte de l'intervention. Nous abordons ces questions à travers deux analyses illustrant deux enjeux distincts de l'analyse d'impact. La première analyse inhérente à l'enjeu des indicateurs de mesure d'impact concerne le développement et la validation d'une échelle de mesure de l'*empowerment* en santé sexuelle dans le cadre de l'évaluation d'impact de l'intervention. En l'absence de données cliniques permettant de calculer l'incidence du VIH, la deuxième analyse inhérente à l'enjeu des méthodes d'estimation concerne la modélisation mathématique de l'impact de l'intervention sur la transmission du VIH, non seulement au niveau des participants de l'étude mais de manière étendue aux personnes éligibles. Nous mettons ensuite en perspective la mise en œuvre de ces méthodes et les résultats obtenus à l'aide d'une démarche réflexive pour évaluer comment les enjeux de l'évaluation d'impact de l'intervention sont imbriqués avec les enjeux statistiques, les enjeux de la mise en œuvre de l'intervention, les enjeux liés à la position des différents participants à l'évaluation et les enjeux de la position de l'évaluateur dans l'intervention.

Chapitre – 6 : Les enjeux des indicateurs de mesure de l’impact de l’intervention

Ce chapitre illustre les enjeux liés aux indicateurs de mesure d’impact dans le cadre d’interventions complexes, notamment basées sur la notion d’*empowerment*. Le développement de l’échelle de mesure de l’EISS dans le projet MAKASI a fait l’objet d’une publication dans la revue *American Journal of Health Education*. L’ensemble de ce chapitre a également fait l’objet d’un chapitre de livre, de deux communications orales au séminaire du département HEALTH de l’Institut Convergence Migrations en 2021 et à la 24^è Conférence mondiale de l’UIPES sur la Promotion de la Santé en 2022, et de deux communications affichées aux 10^è et 11^è éditions de la conférence internationale francophone VIH/Hépatites (Afravih). La section suivante présente le résumé de ce travail, l’article est présenté en Annexe 3.

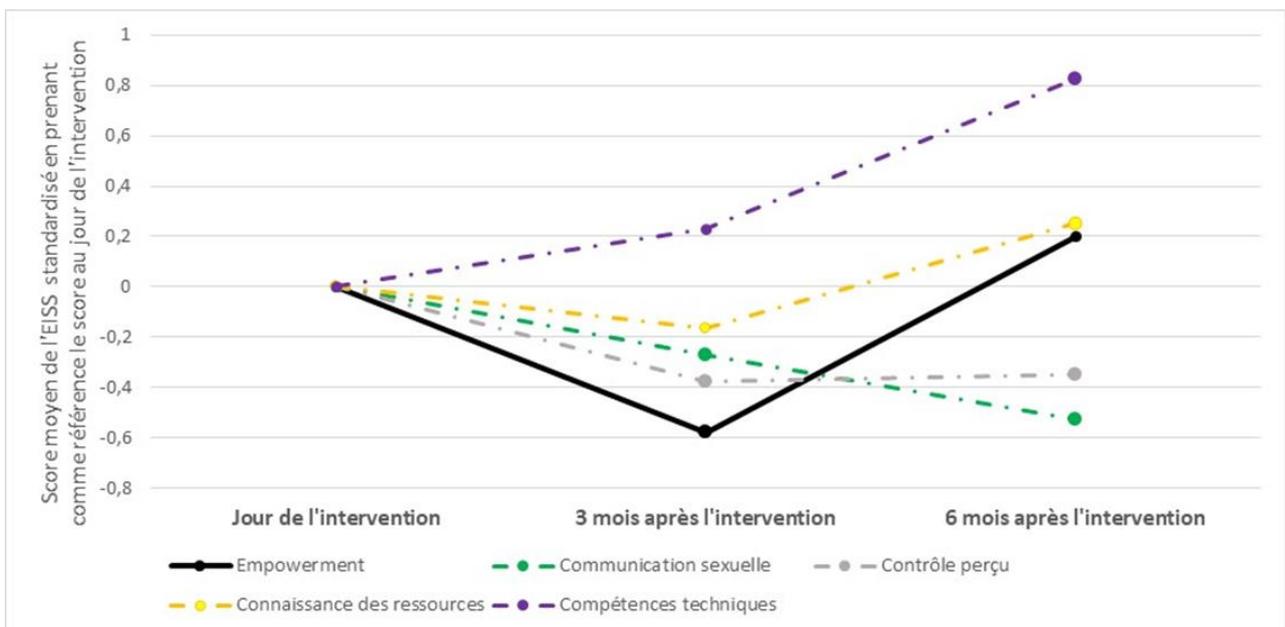
Résumé de l'article 3

L'*empowerment* en santé sexuelle, bien que considéré comme un indicateur secondaire de l'impact de l'intervention dans le devis de son évaluation, tient une place centrale dans sa conception. C'est sur cette notion que repose les hypothèses expliquant la diminution de l'exposition aux risques sexuels donc la diminution du risque de transmission du VIH dans le cadre du projet MAKASI. Mesurer l'*empowerment* individuel en santé sexuelle (EISS) dans le cadre de l'intervention revient donc à mesurer un déterminant intermédiaire et non secondaire de l'impact de l'intervention. Disposer d'indicateurs cohérents avec la conception, la mise en œuvre et les objectifs de l'intervention est primordial pour en évaluer l'impact.

Dans l'analyse suivante, nous évaluons dans quelles mesures les indicateurs de communication sexuelle, les indicateurs de perception de contrôle sur les décisions, les indicateurs de connaissances vis-à-vis des ressources sociales et outils médicamenteux de la prévention du VIH, et les indicateurs de littératie en santé collectés dans le cadre de l'intervention permettent de mesurer l'EISS de manière cohérente avec l'intervention. Les résultats montrent que ces indicateurs provenant d'autres études et d'autres contextes sont assez hétérogènes pour mesurer différents aspects de l'EISS mais assez homogènes pour être utilisés conjointement. Leur mise en commun permet de développer une échelle de mesure dont les scores varient en fonction de la capacité individuelle à se saisir des ressources existantes pour améliorer la situation sociale et de santé, de l'habileté à communiquer des informations sur la situation sociale et de santé à des professionnels, du sentiment d'appartenance à un groupe social, du renoncement aux soins, de l'utilisation de contraceptifs et de la réalisation de tests de diagnostic du VIH. Le développement de cette échelle montre qu'il est possible de disposer d'indicateurs d'impact cohérents avec le contexte d'une intervention mais nécessite au moins l'intégration du choix ou du développement de ces

indicateurs en amont de la conception de l'intervention ainsi qu'une concession entre les aspects théoriques (tels que l'usage d'un cadre théorique de l'EISS) et pratiques (tels qu'un environnement propice à la participation des communautés concernées à chaque activité de l'intervention). Cet exercice a également montré que la nature multidimensionnelle de l'EISS nécessite d'anticiper comment et à quels moments l'échelle peut être utilisée dans le cadre de l'évaluation de l'impact de l'intervention. Les différents processus de l'EISS imbriqués et mesurés dans le cadre de l'intervention connaissent des évolutions différentes durant le suivi des participants (Figure 11). Ces évolutions différentes montrent que ces processus ne sont pas linéaires et nécessitent une caution supplémentaire dans l'utilisation d'échelles de mesure : la notion d'impact ne constitue ici qu'une mesure à un temps donné et doit être mise en perspective avec ce qui est connu des différents processus. Il est donc nécessaire non seulement de disposer d'indicateurs mais de les utiliser à bon escient pour réaliser au mieux les évaluations d'impact en respectant non seulement les contraintes statistiques mais aussi les contraintes conceptuelles liées à la mise en œuvre des interventions.

Figure 11 : Evolution du score moyen global et par dimension d'EISS chez les participants au cours du suivi dans l'intervention



Source : projet MAKASI, Lecture : le score moyen d'EISS (trait plein de couleur noire) augmente de 0.2 point 6 mois après l'intervention, le score moyen correspondant aux compétences techniques (trait discontinu de couleur violette) augmente de 0.8 point 6 mois après l'intervention

Chapitre – 7 : Les enjeux des méthodes d'estimation d'impact de l'intervention

Ce chapitre illustre les enjeux liés aux méthodes d'estimation d'impact dans le cadre d'interventions complexes. Il concerne les résultats d'une modélisation mathématique du risque de transmission du VIH chez les personnes bénéficiant de l'intervention MAKASI. Ce travail a fait l'objet d'une communication affichée à la conférence AIDSImpact en 2023 et fera l'objet d'une publication dans une revue à comité de lecture. La section suivante présente le résumé de ce travail, un document de travail présentant l'ensemble des méthodes et des résultats de cette modélisation est présenté en Annexe 4.

Résumé de l'article 4

L'intervention MAKASI base sa stratégie de prévention du VIH sur l'*empowerment* en santé sexuelle. Dans le contexte de cette intervention, cette notion signifie présenter aux individus les ressources de soin et les ressources sociales existantes dans leur environnement et les accompagner à s'en saisir selon leur choix afin qu'ils puissent améliorer leur situation globale de manière à améliorer par la même occasion leur santé sexuelle. Selon l'hypothèse de l'intervention, les bénéficiaires adopteraient des comportements qui leur permettrait de diminuer leur exposition sexuelle au risque de transmission du VIH à travers le processus d'amélioration de la santé sexuelle.

Le dispositif de recherche et d'évaluation mis en place ne permet cependant pas de mesurer l'impact de l'intervention sur la transmission du VIH : le temps de suivi des participants est court et aucune procédure permettant de mesurer l'incidence du VIH chez ces derniers n'a été mise en place. Dans ce cas de figure, il est courant d'avoir recours à des méthodes de modélisation mathématique de la transmission du VIH. Ces méthodes utilisent le langage mathématique pour simuler dans quelles mesures un individu est susceptible de s'infecter lors de relations sexuelles étant donné comment il se protège et étant donné les partenaires qu'il est susceptible de rencontrer. Ces simulations permettent de considérer un certain nombre de configurations éventuelles concernant les rencontres sexuelles et leur déroulement pour un seul et même individu. L'estimation du risque de transmission chez un individu consiste en un consensus, c'est-à-dire une moyenne ou une médiane, de l'ensemble des scénarios envisagés. La pertinence de l'estimation obtenue repose sur l'expression (mathématique) utilisée et les informations (données de l'intervention, d'autres interventions ou de la littérature) qui sont utilisées dans cette expression.

Dans le cas de l'intervention MAKASI, nous utilisons une de ces méthodes, la méthode de simulation de Monte Carlo, pour exprimer le bénéfice potentiel de l'intervention à diminuer le risque d'infection au VIH chez les participants. L'expression du risque de transmission s'inspire d'un modèle qui a largement fait ses preuves pour évaluer la transmission du VIH et utilise des informations issues de l'intervention MAKASI, d'enquêtes représentatives de la population des IASS en France et de la population générale française, et des données de la littérature concernant la sexualité des participants ainsi que la dynamique de transmission de la maladie. Afin de discuter de la pertinence de l'intervention, dans un premier temps nous opposons un scénario de déploiement réussi de cette intervention avec un scénario d'absence d'intervention et des scénarios de déploiement réussi d'outils biomédicaux de la prévention combinée du VIH (PrEP, PEP). Procéder de cette manière nous permet :

- d'une part de discuter le gain éventuel à implémenter ce type d'intervention dans un environnement où il n'y a pas d'offre de prévention offrant un accompagnement pour une prise en charge globale des problématiques sociales et structurelles pouvant impacter la santé sexuelle,
- et d'autre part de discuter le gain éventuel à mettre en œuvre ce type d'intervention comparé aux efforts menés pour baser la prévention uniquement sur le déploiement d'outils biomédicaux.

Ces résultats préliminaires montrent que donner le choix et accompagner les individus pour qu'ils puissent se saisir des outils qui leur conviennent est une stratégie raisonnable que ce soit pour faciliter la mise en œuvre de la prévention ou que ce soit pour limiter le nombre de cas incidents d'infection au VIH dans cette population.

Dans un second temps, nous mettrons en contribution les différents acteurs de l'intervention pour discuter et affiner les hypothèses testées jusque-là. Nous nous basons sur les principes de

la méthode de narration quantitative²² (Saltelli & Giampietro, 2017) pour formuler des hypothèses faisables, viables et désirables. Nous mettrons également à contribution ces acteurs pour améliorer les procédures choisies pour analyser la sensibilité des résultats de manière à prendre en compte non seulement l'aspect technique de la modélisation mais également leur sensibilité au contexte de mise en œuvre de l'intervention (Saltelli et al., 2013). Enfin, nous mettrons nos résultats en regard avec les résultats des études sur l'acceptabilité et l'adoption des outils biomédicaux de la prévention du VIH pour saisir les enjeux de leur déploiement.

Dans son état actuel, cet exercice de modélisation montre que la prise en compte des dynamiques sociales de l'épidémie de VIH est tout aussi importante que la prise en compte des dynamiques biologiques dans les études de modélisation de la transmission du VIH au sein des populations. Il démontre la nécessité d'identifier et de produire des données pour permettre d'étudier l'impact de ces enjeux sur la propagation du virus. Enfin, les outils de modélisation mathématique sont des outils narratifs et illustratifs pertinents pour démontrer le potentiel des stratégies de prévention basées sur les problématiques socio structurelles dans la lutte contre les épidémies.

²² Il s'agit ici d'une traduction libre de la notion de « quantitative story telling »

Chapitre – 8 : Analyse réflexive de la mise en œuvre des méthodes quantitatives dans le cadre du projet MAKASI

Ce chapitre concerne une analyse réflexive de l'usage des méthodes quantitatives dans l'évaluation d'impact d'interventions complexes en prenant pour cas le projet MAKASI. Cette analyse permet de rediscuter la mise en œuvre de ces méthodes étant donné la position de l'évaluateur vis-à-vis du contexte de l'intervention. Ce travail fait l'objet d'une publication actuellement en révision dans une revue à comité de lecture. La section suivante présente le résumé de ce travail, la dernière version du manuscrit soumis est présentée en Annexe 5.

Résumé de l'article 5

Les études présentées dans les chapitres 6 et 7 reflètent comment les méthodes quantitatives ont été mobilisées pour évaluer l'impact de l'intervention MAKASI. Ces exemples montrent que les objectifs de l'intervention sont pris en compte, malgré certaines limites, sur l'ensemble des activités prévues dans le projet de la conception de l'intervention à la mise en œuvre et l'évaluation. Les résultats obtenus sont la résultante de choix concertés à travers un dispositif participatif impliquant chercheurs, intervenants de terrain et pairs communautaires. La réalisation de ces études est donc influencée par les dynamiques engendrées par l'interaction entre ces différents acteurs ainsi que les normes et pratiques qui accompagnent ces derniers. Une analyse réflexive de la mise en œuvre des méthodes quantitatives dans ce contexte permet de mettre en perspective ces dynamiques avec la validité statistique et mathématique des estimations de l'impact des interventions de santé publique.

Dans ce chapitre nous effectuons une analyse réflexive sur l'usage des méthodes quantitatives dans le cas de l'évaluation d'impact de l'intervention MAKASI. Les données illustrant le processus de mise en œuvre des méthodes quantitatives sont générées de manière inductive et analysées de manière déductive à l'aide de cadres théoriques de la réflexivité. Nous utilisons deux cadres théoriques proposant une typologie et une conception des niveaux de la réflexivité en promotion de la santé. Notre analyse montre que le choix des méthodes à mettre en œuvre et le processus de mise en œuvre sont influencés par les objectifs fixés dans le cadre des interventions ainsi que les normes et les disciplines scientifiques véhiculées par les acteurs impliqués dans l'évaluation. La création d'espaces de concertations impliquant les acteurs de l'intervention est propice pour concilier les intérêts de chaque acteur, y compris les participants à l'intervention, sans mettre en péril la validité des analyses quantitatives. Notre analyse montre également que la place des évaluateurs est centrale pour créer une dynamique favorable à la réalisation des objectifs des interventions de santé publique. En

particulier, leur identité peut être déterminante dans l'interaction avec les acteurs issus de la recherche, du terrain ou des communautés, et surtout vis-à-vis des attentes de ces derniers.

Ces résultats illustrent :

- l'importance du choix des méthodes les plus cohérentes avec le contexte des interventions et ne remettant pas en question les rigueurs d'application des méthodes quantitatives,
- l'importance de la création d'espaces de concertation dialogiques pour améliorer les pratiques évaluatives auprès des acteurs,
- et l'importance des approches intersectionnelles pour identifier les acteurs et les communautés impliquées dans la mise en œuvre des interventions en promotion de la santé.

Dans le cas précis de l'évaluation d'impact de l'intervention MAKASI, la démarche réflexive révèle le potentiel de l'évaluation et de l'évaluateur comme leviers de l'*empowerment* au niveau communautaire.

Partie – IV : Discussion générale

*« Personne n'éduque autrui, personne ne s'éduque seul, les hommes
s'éduquent ensemble par l'intermédiaire du monde. »*

Paulo Freire, *Pédagogie des opprimés*, 1974.

Les travaux dans la littérature identifiés au cours de ce travail et les études de cas effectuées ont mis en lumière la diversité des méthodes disponibles et valides pour évaluer l'impact des CHPI, et de manière plus large les programmes complexes en santé publique. Cette diversité implique une diversité dans le choix et la mobilisation de ces méthodes quantitatives d'évaluation d'impact. L'orientation donnée à l'évaluation concernant les méthodes quantitatives est dépendante des dynamiques engendrées par le système qui régit l'organisation des activités liées à la mise en place des programmes en santé publique. Cette orientation dépend donc également de l'interaction entre les différents acteurs engagés sans que cela nuise forcément à la validité statistique ou mathématique des méthodes telles que mobilisées. En effet, les dynamiques inhérentes aux activités et acteurs impliqués dans ces programmes déterminent les ressources organisationnelles ainsi que méthodologiques et garantissent l'usage de ces ressources vis-à-vis des objectifs de l'intervention, tandis que les contraintes statistiques et mathématiques garantissent la cohérence entre les données collectées, les méthodes utilisées et les inférences effectuées. Ces enjeux bien distincts sont inextricables en réalité car déterminent le bien fondé et la nature des produits de l'évaluation. Ce travail de thèse a permis d'identifier ces enjeux, d'analyser leur spécificité et d'en déduire quelques recommandations quant à l'usage des méthodes quantitatives dans l'évaluation d'impact des programmes complexes en promotion de la santé. En particulier, nous avons identifié l'importance de l'identité de l'évaluateur, des communautés constituées autour de la mise en œuvre des programmes et des pratiques et savoirs intrinsèques aux disciplines mobilisées lors de l'évaluation de ces programmes.

Apports identitaires des évaluateurs

L'identité des évaluateurs impliqués dans la mise en œuvre des programmes de santé publique est une question complexe qui dépasse la simple dichotomie entre les acteurs provenant et ne provenant pas des communautés ou populations récipiendaires de ces programmes. Les analyses que nous avons effectuées ont mis en évidence des enjeux intersectionnels concernant l'identité des évaluateurs. Cette réflexion peut être étendue aux questions d'intersection entre les communautés auxquelles les acteurs et participants des programmes de santé appartiennent et s'identifient, et aux questions de compétences culturelles en ce qui concerne l'évaluation de ces programmes.

L'implication des acteurs communautaires engagés dans l'évaluation ne peut se penser uniquement au regard de l'ampleur des épidémies et problématiques sanitaires à partir desquelles on identifie les populations à qui les programmes sont destinés. La complexité de leurs identités amène à considérer les communautés auxquelles ils appartiennent sous différents points de vue. Il est possible, par exemple, de considérer l'appartenance communautaire en fonction des origines géographiques, des situations vécues, des intérêts, des aspirations ou des professions (Millington, 2010). Chacun de ces points de vue contribue aux atouts dont ils disposent pour évoluer dans le système qui régit les programmes de santé et négocier avec ou au détriment d'autres acteurs les décisions inhérentes à leur fonction dans ce système. A titre d'exemple, l'analyse réflexive que nous avons menée démontre que l'imbrication entre l'origine géographique de l'évaluateur, les situations vécues et sa profession le distingue des bénéficiaires de l'intervention tels qu'ils ont été définis. Bien que son origine géographique s'apparente à celle des bénéficiaires de l'intervention, amenant parfois à certaines situations communément vécues avec ces derniers, son statut professionnel au sein de l'intervention lui confère une marge de manœuvre différente des bénéficiaires dans

le système des interventions de santé publique. Les questions communautaires, surtout lorsqu'elles portent des attentes liées à la mise en œuvre des interventions, doivent être pensées en amont dès leur conception. Elles déterminent en partie l'orientation que peut prendre les processus d'évaluation et les produits qui en résultent.

Si les questions communautaires telles qu'évoquées dans le paragraphe précédent apportent un début de réponse concernant les attentes vis-à-vis des acteurs engagés dans les interventions de santé publique, elles sont souvent orientées par les attentes provenant des commanditaires de ces programmes. Repenser l'appartenance communautaire en lien avec les compétences culturelles permet de concilier les attentes de l'ensemble des acteurs engagés autour de ces interventions (Gautier, 2022; Gill et al., 2016). Cela implique d'identifier comment mettre à profit les atouts des acteurs de l'intervention selon les communautés auxquelles ils appartiennent ou comment leur positionnement vis-à-vis de ces communautés est pertinent vis-à-vis du contexte de l'évaluation pour répondre aux objectifs des programmes de santé. Conjuguer les informations concernant les appartenances communautaires avec les notions de compétence culturelle peut s'avérer utile pour ancrer l'évaluation dans la recherche de solutions pragmatiques pour les bénéficiaires.

Apports intercommunautaires

Tandis que l'identité de chaque acteur ne peut être défini sans considérer les intersections entre les appartenances communautaires et les dynamiques de pouvoir inhérentes, la mise en œuvre des interventions de santé publique imposent l'appartenance à une communauté de pratique (Millington, 2010) définie par les fonctions des acteurs et les différentes activités. Dans notre cas on observe que la participation des communautés de chercheurs impliqués dans la conception des programmes, des associations communautaires et des organisations non

gouvernementales impliqués dans la mise en œuvre sur le terrain, et des pairs communautaires porteurs de savoirs expérientiels a été pensée en amont (Gosselin et al., 2019). La participation de ces communautés et la délimitation des fonctions de chacune de ces communautés dans la mise en œuvre des interventions constituent une opportunité pour confronter les connaissances scientifiques aux connaissances expérientielles et contextuelles afin de produire des informations utiles en pratique. Alors que l'identité des acteurs apporte des informations spécifiques dans le processus d'évaluation, les communautés de pratique (Millington, 2010) apportent non seulement des informations et des connaissances mais reflètent aussi les intérêts et marge de manœuvre de chaque communauté dans le cadre de l'intervention. A la différence de l'identité individuelle, l'identité communautaire liée aux fonctions dans l'intervention sert des intérêts communs spécifiques restreints par ces fonctions. Confronter ces intérêts dans un processus participatif peut être une opportunité en faveur de l'évaluation : la participation de chaque communauté permet d'équilibrer les rapports de pouvoirs et les intérêts de chaque partie. Dans notre cas, les produits de l'évaluation d'impact de l'intervention MAKASI ne se sont pas limités à la production scientifique, la participation des communautés a vu la présentation des résultats sous différentes formes adaptées à l'ensemble des acteurs, telles que des notes de recherche. Cela permet une meilleure participation et une appropriation des produits de l'évaluation par l'ensemble des acteurs engagés dans ce type de programme.

Apports interdisciplinaires

Les acteurs impliqués dans la mise en œuvre des programmes de santé proviennent souvent de disciplines et de systèmes de pratiques différents. La participation de ces acteurs voit alors la rencontre entre ces pratiques et ces disciplines ainsi que les savoirs et les procédés qui en découlent. Les produits de l'évaluation sont donc définis par les savoirs produits dans le cadre

de ces systèmes. Les méthodes véhiculées par ces systèmes de connaissance sont assez nombreuses pour fournir des possibilités de choix pertinents pour évaluer l'impact des interventions. Néanmoins, les marges de manœuvre de chaque acteur, qui peut être différent selon sa position dans le système des programmes de santé, peuvent biaiser le choix des méthodes si une précaution n'est pas mise en place pour remettre en question ces choix de manière à en garantir la pertinence. En effet, si les méthodes choisies résultent uniquement d'une position liée aux disciplines et cautionnée par un rapport de pouvoir déséquilibré dans l'organisation des programmes, la pertinence des produits de l'évaluation risque de ne pas être garantie vis-à-vis du contexte de l'intervention. Il ne s'agit pas ici simplement d'un biais méthodologique statistique ou mathématique, mais peut refléter un biais lié aux inégalités entre les acteurs dans la mise en œuvre des interventions. Par contre la confrontation de ces connaissances, de leurs avantages et de leurs inconvénients dans le cadre d'un espace méthodologique dédié peut s'avérer être une opportunité. Reconsidérer ces méthodes comme de simples outils aux caractéristiques bien précises et amenant à des possibilités d'inférence limitées permet de recentrer le choix uniquement sur les caractéristiques statistiques et mathématiques et l'adéquation de ces propriétés vis-à-vis des données, tendant à effacer les dynamiques de pouvoir entre les acteurs. Grâce à ce processus, il devient possible de faire le choix des méthodes en fonction du contexte de l'intervention pour dépasser les intérêts des acteurs et les positions hiérarchiques. Ces méthodes ont vu le jour au sein de ces disciplines pour permettre de prendre en compte des défis méthodologiques d'analyses des données et non liés à la notoriété de ces disciplines.

Place des enjeux des méthodes dans les courants de l'évaluation

Nous avons pu voir tout au long de cette thèse que l'importance donnée au contexte de l'intervention dans l'évaluation dépend autant de la qualité statistique ou mathématique des méthodes mobilisées que de l'interaction entre les acteurs de l'évaluation de l'intervention. La portée dans d'autres contextes des informations et inférences faites à partir des méthodes quantitatives se limite aux données considérées dans la mise en œuvre des méthodes. Que ce soit la validité interne (isolement de l'effet de l'intervention), la validité externe (maintien de l'effet de l'intervention au-delà du contexte de l'intervention ou même la viabilité (prise en compte des enjeux de terrain des intervenants) ; elles dépendent surtout des informations traitées lors de la mise en œuvre des méthodes quantitatives, l'inférence ne pouvant aller au-delà. Nous avons pu voir que la démarche participative a permis de faire des choix concertés sur ces informations et leurs traitements statistique et mathématique. Donner l'opportunité aux différents acteurs d'interagir et réagir sur leur usage agit non seulement sur les trois enjeux de validité précédents, mais aussi sur la validité de construit des données et des résultats de l'évaluation, c'est-à-dire leur crédibilité et leur sensibilité vis-à-vis du contexte de l'intervention. Plusieurs exemples, depuis le choix des indicateurs à considérer pour mesurer l'*empowerment* jusqu'au ressenti des intervenants sur la sensibilité de ces indicateurs au contexte d'administration des questionnaires, en témoignent.

Il existe néanmoins des conditions à respecter qui ne dépendent que des méthodes elles-mêmes et non de l'interaction entre les acteurs engagés dans l'intervention. Ces méthodes possèdent des propriétés statistiques et mathématiques sur lesquelles il est impossible de faire l'impasse au risque d'obtenir des estimations dénuées de sens. Les conditions de validité mathématique et statistique de ces méthodes sont spécifiques aux théories utilisées pour élaborer et dans lesquelles mesures les données satisfont les hypothèses sous-jacentes. Le fait de satisfaire ces

conditions de validité garantit dans quelles mesures la distribution des variables utilisées et la distribution des erreurs de mesures permettent de mesurer précisément l'effet des programmes et estimer à quel point les effets observés sont juste un phénomène aléatoire ou attribuables au dispositif testé en cohérence avec la méthode utilisée.

La pertinence des méthodes pour évaluer les interventions en santé publique, au-delà des courants de l'évaluation, dépend d'un équilibre reposant sur les dynamiques entre les acteurs (garantissant l'usage des informations en faveur des objectifs de l'intervention) et sur les conditions de validité statistique/mathématique de ces méthodes (garantissant que les outils soient utilisés correctement pour exploiter ces informations). Ainsi, le respect des conditions de validité garantit que la précision des estimations et la prise en compte des erreurs liées à ces estimations : le choix et la mise en œuvre des méthodes retranscrit la lecture des données vis-à-vis du dispositif de recherche. Les résultats obtenus à partir d'un échantillon constitué à partir de la même population et traitée dans les mêmes conditions sont censés être similaires lorsque les méthodes utilisées sont mises en œuvre correctement. En ce sens, la notion de généralisabilité statistique consiste en la reproductibilité des effets aux mêmes conditions expérimentales et d'administration des programmes en fonction des informations et méthodes utilisées. La généralisabilité des effets, non pas au sens mathématique mais pratique, dépend quant à elle de la reproductibilité de l'administration de l'intervention. C'est pourquoi nous avons insisté sur la notion de transférabilité de l'intervention dans cette thèse. Dans notre travail, effectuer une inférence sur l'usage de l'intervention au niveau de la population dépend des informations disponibles pour inférer sur la mise en œuvre. Ainsi, il est nécessaire de considérer les déterminants de la mise en œuvre (Proctor et al., 2011) dans la réflexion sur l'évaluation. Bien que les recherches sur les déterminants de la mise en œuvre existent, au cours de l'intervention les dynamiques entre acteurs détermineront ceux qui seront considérés dans les informations exploitées concernant l'évaluation. Donner la possibilité aux

communautés de participer tout au long permet de faire ressurgir ceux qui ont plus de sens sur le terrain.

En définitive, il revient donc aux acteurs engagés dans l'évaluation de recueillir et exploiter les informations nécessaires pour prendre en compte le contexte des programmes de santé. La question de l'usage qui est fait des méthodes quantitatives se distingue donc des débats ayant lieu à l'intersection des différents courants de l'évaluation. Si les questions concernant la manière de définir la causalité sont bien pertinentes, la nature des méthodes quantitatives n'en fait pas forcément des outils incompatibles avec une vision générativiste de la causalité ou restreints à une vision successionniste de celle-ci.

Le cas particulier de la modélisation prospective

Dans cette thèse, nous avons étudié le cas particulier de la modélisation prospective pour étudier l'impact de l'intervention MAKASI. Ce type de méthode, courant dans l'évaluation des stratégies de lutte contre le VIH, présente des limites propres qui nécessitent qu'on y porte attention. L'utilité de ce type de modélisation est bien identifiée dans le domaine des politiques basées sur les données probantes, cependant les informations qui sont produites à l'aide de ces modèles pourraient erroner l'appréciation d'un dispositif, voire même les prises de décisions y afférant (Saltelli & Funtowicz, 2014; Saltelli & Giampietro, 2017). En effet, la difficulté à simuler correctement les systèmes complexes que ces modèles ont pour ambition de mimer résulte en une simplification des hypothèses et informations modélisées. Cette simplification contribue parfois à omettre de manière intentionnelle ou non, sous couvert de technicité, le caractère hypothétique de ces estimations et leur écart avec la réalité de mise en œuvre des programmes. Pour que les décisions portant sur le déploiement de dispositifs soient basées

sur les preuves au lieu d'orienter une construction erronée de ces dernières²³ (Marmot, 2004; Saltelli & Giampietro, 2017), certains principes déjà identifiés doivent être adoptés lors de l'usage de ces modèles, et des méthodes quantitatives en général (Saltelli et al., 2013; Saltelli & Giampietro, 2017).

L'adoption de ces principes en pratique n'est pas aisée mais néanmoins envisageable. Dans le cas du projet MAKASI, nous avons tenu à être vigilants sur deux points précis dans l'exercice de modélisation prospective : la construction des hypothèses à simuler et la participation des différents acteurs dans la construction de ces hypothèses. Le premier point concernant la formulation des hypothèses à tester est essentiel pour garantir la pertinence des estimations obtenues à l'aide des modèles mathématiques. La pandémie de coronavirus rappelle à quel point ces estimations sont sensibles à la formulation et surtout à la simplification de ces hypothèses, au risque de compromettre les prises de décision (Pawson, 2020). Il n'en demeure pas moins que la modélisation prospective reste un outil d'aide à la décision utile si les hypothèses formulées tiennent compte de réalités plus complexes malgré leur simplification. Pour ce faire, Saltelli *et al.* (Saltelli & Giampietro, 2017) proposent des principes de faisabilité (l'existence d'infrastructures suffisantes pour déployer le dispositif), de viabilité (l'existence de ressources pour mettre en œuvre le dispositif) et de désirabilité (la pertinence du dispositif dans son contexte pour répondre aux besoins des usagers). Le second point concernant la participation des acteurs dans la formulation des hypothèses à tester est essentiel pour identifier les informations permettant d'affiner ces hypothèses et garantir leur crédibilité. La réalisation d'une démarche participative dans le cas d'une modélisation prospective nécessitera une vulgarisation de ces méthodes permettant au moins à chaque acteur d'en saisir les enjeux, dont techniques, pour pouvoir apporter des informations

²³ Nous faisons ici référence aux notions opposées d' « evidence based policy » et de « policy based evidence »

pertinentes dans leur mise en œuvre, notamment pour évaluer leur sensibilité et la plausibilité des résultats (Saltelli et al., 2013). Ainsi, en adoptant ces principes, les modèles demeurent des outils utiles d'aide à la décision, ne serait-ce que pour illustrer les conséquences possibles du déploiement de dispositifs lorsque les informations à disposition sont insuffisantes et lorsque la pertinence des hypothèses a été pensée en amont, même si la pertinence des estimations obtenues est conditionnelle aux hypothèses testées.

Application à la promotion de la santé

L'usage des méthodes quantitatives n'est un frein à l'évaluation des interventions en promotion de la santé que si on efface la voix des communautés et l'importance des contextes. Le travail que nous avons effectué montre qu'il est possible à la fois de collaborer avec les communautés dans l'usage des méthodes quantitatives d'évaluation et de considérer le contexte dans l'usage de ces méthodes. Cela nécessite cependant l'adoption des principes d'équité, de justice sociale, de participation et de pouvoir d'agir des participants propres à la promotion de la santé en amont (Organisation mondiale de la Santé, 1986). Pour ce faire, il est nécessaire d'adopter une démarche participative, ceci étant dit, il ne s'agit pas seulement de compter les communautés dans les rangs : il faut créer un environnement propice à leur participation. Dans cet environnement, des partages d'informations, de compétence et des transferts de connaissance sont en jeu. La pertinence de l'usage des méthodes quantitatives pour l'évaluation des programmes en promotion de la santé ne peut être garantie qu'à travers une mutualisation réussies des compétences des acteurs dans cet environnement. Bien que l'usage des méthodes quantitatives nécessitent l'acquisition de compétences via un parcours académique, un effort de transfert de connaissance en cours de la mise en œuvre de l'évaluation des programmes de santé est possible et salutaire pour une participation effective des communautés dans l'évaluation.

En ce qui concerne la pertinence des méthodes quantitatives, deux questions se posent souvent en filigrane pour en critiquer l'usage :

- Les méthodes quantitatives sont-elles toujours non pertinentes ? : la réponse est non : les propriétés de ces méthodes ne sont pas un frein à leur usage en promotion de la santé.
- L'usage qui est fait de ces méthodes est-il toujours pertinent ? : la réponse est non : pour cela il faut adopter les principes compatibles avec la promotion de la santé dans leur usage

La vraie question à poser est sans doute dans quelles mesures le processus d'évaluation, donc d'usage des méthodes quantitatives dans ce cadre le cas échéant, respecte-t-il les principes de la promotion de la santé. La pertinence des outils n'est pas à remettre en question, il s'agirait plutôt de remettre en question leur usage. Une démarche participative respectant et mettant à profit l'individualité de chacun, l'individualité des communautés et les compétences propres à chacun permet de faire de l'évaluation utilisant les méthodes quantitatives en promotion de la santé.

Conclusion

Cette thèse avait pour ambition de comprendre dans quelles mesures les méthodes quantitatives peuvent être mises à profit pour mieux prendre en compte les contextes des interventions de promotion de la santé lors des études d'évaluation d'impact. Nos résultats, si elles ne permettent pas de répondre à cette question de manière précise, montrent que la mobilisation de ces méthodes est régie par des dynamiques complexes qui concernent les disciplines et pratiques mobilisées, les contextes de mise en œuvre des programmes, et l'interaction entre les acteurs impliqués dans ces évaluations. De ces résultats, il est possible d'affirmer que l'usage des méthodes quantitatives dans l'évaluation des interventions en promotion de la santé génère une rencontre entre plusieurs courants et systèmes de pensée (paradigmes) influençant le processus d'évaluation d'impact, qu'ils relèvent des sciences quantitatives, des sciences de l'évaluation ou des pratiques des acteurs de la promotion de la santé hors du champ académique. Les choix liés aux méthodes utilisées résultent donc de dynamiques complexes qui reflètent parfois les situations des acteurs dans le contexte de ces interventions. Il semble à la lueur de nos analyses que ces choix peuvent à la fois garantir la pertinence mathématique des méthodes utilisées et les principes de la promotion de la santé, tout en tenant compte, ne serait-ce que partiellement, des contextes des interventions. Il est cependant nécessaire que les acteurs et les évaluateurs adoptent des principes qui y soient favorables. Le positionnement de l'évaluateur vis-à-vis du dispositif à évaluer peut influencer le processus d'évaluation. Cependant, sa marge de manœuvre dépend de la position des autres acteurs engagés dans le dispositif qu'ils appartiennent au monde académique ou non, mais aussi les commanditaires de l'évaluation, y compris ceux qui ne sont pas évoqués ici tels que les financeurs, jouant un rôle non négligeable dans l'orientation de l'évaluation et ses produits. En définitive, si l'usage des méthodes quantitatives dans l'évaluation des programmes de

promotion de la santé ne peut être réduite à une démarcation entre les différents paradigmes évoqués, sa pertinence dépend du consensus entre les différents acteurs impliqués et leur choix d'adopter ou non une démarche cohérente avec ces programmes en ce qui concerne les évaluations, leurs produits et l'utilisation de ces produits.

Bibliographie

- Aguayo-Albasini, J. L., Flores-Pastor, B., & Soria-Aledo, V. (2014). GRADE System : Classification of Quality of Evidence and Strength of Recommendation. *Cirugía Española (English Edition)*, 92(2), 82-88. <https://doi.org/10.1016/j.cireng.2013.08.002>
- Alkin, M. C., & Taut, S. M. (2021). *Démêler les usages de l'évaluation*. <https://scienceetbiencommun.pressbooks.pub/evaluationanthologie/chapter/alkin-marvin-c-sandy-m-taut-2003-demeler-les-usages-de-levaluation/>
- Alterman, R., Carmon, N., & Hill, M. (1984). Integrated evaluation : A synthesis of approaches to the evaluation of broad-aim social programs. *Socio-Economic Planning Sciences*, 18(6), 381-389. [https://doi.org/10.1016/0038-0121\(84\)90047-8](https://doi.org/10.1016/0038-0121(84)90047-8)
- Armitage, P. (2003). Fisher, Bradford Hill, and randomization. *International Journal of Epidemiology*, 32(6), 925-928. <https://doi.org/10.1093/ije/dyg286>
- Baron, J. (2018). A Brief History of Evidence-Based Policy. *The ANNALS of the American Academy of Political and Social Science*, 678(1), 40-50. <https://doi.org/10.1177/0002716218763128>
- Barré-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vézinet-Brun, F., Rouzioux, C., Rozenbaum, W., & Montagnier, L. (1983). Isolation of a T-Lymphotropic Retrovirus from a Patient at Risk for Acquired Immune Deficiency Syndrome (AIDS). *Science*, 220(4599), 868-871. <https://doi.org/10.1126/science.6189183>
- Beckfield, J., Olafsdottir, S., & Bakhtiari, E. (2013). Health Inequalities in Global Context. *American Behavioral Scientist*, 57(8), 1014-1039. <https://doi.org/10.1177/0002764213487343>
- Begg, C., Cho, M., Eastwood, S., Horton, R., Moher, D., Olkin, I., Pitkin, R., Rennie, D., Schulz, K. F., Simel, D., & Stroup, D. F. (1996). Improving the Quality of Reporting of Randomized Controlled Trials : The CONSORT Statement. *JAMA*, 276(8), 637-639. <https://doi.org/10.1001/jama.1996.03540080059030>
- Berger, N. (2010). Sociologie analytique, mécanismes et causalité : Histoire d'une relation complexe. *L'Année sociologique*, 60(2), 419-443. <https://doi.org/10.3917/anso.102.0419>

- Birdthistle, I., Kwaro, D., Shahmanesh, M., Baisley, K., Khagayi, S., Chimbindi, N., Kamire, V., Mthiyane, N., Gourlay, A., Dreyer, J., Phillips-Howard, P., Glynn, J., & Floyd, S. (2021). Evaluating the impact of DREAMS on HIV incidence among adolescent girls and young women : A population-based cohort study in Kenya and South Africa. *PLOS Medicine*, 18(10), e1003837. <https://doi.org/10.1371/journal.pmed.1003837>
- Blankenship, K. M., Friedman, S. R., Dworkin, S., & Mantell, J. E. (2006). Structural Interventions : Concepts, Challenges and Opportunities for Research. *Journal of Urban Health*, 83(1), 59-72. <https://doi.org/10.1007/s11524-005-9007-4>
- Bleich, S. N., Jarlenski, M. P., Bell, C. N., & LaVeist, T. A. (2012). Health Inequalities : Trends, Progress, and Policy. *Annual Review of Public Health*, 33(1), 7-40. <https://doi.org/10.1146/annurev-publhealth-031811-124658>
- Boily, M.-C., Pickles, M., Lowndes, C. M., Ramesh, B. M., Washington, R., Moses, S., Deering, K. N., Mitchell, K. M., Reza-Paul, S., Blanchard, J., Vassall, A., Alary, M., & Vickerman, P. (2013). Positive impact of a large-scale HIV prevention programme among female sex workers and clients in South India. *AIDS*, 27(9), 1449-1460. <https://doi.org/10.1097/QAD.0b013e32835fba81>
- Bonell, C., & Imrie, J. (2001). Behavioural interventions to prevent HIV infection : Rapid evolution, increasing rigour, moderate success. *British Medical Bulletin*, 58(1), 155-170. <https://doi.org/10.1093/bmb/58.1.155>
- Bonell, C., Warren, E., Fletcher, A., & Viner, R. (2016). Realist trials and the testing of context-mechanism-outcome configurations : A response to Van Belle et al. *Trials*, 17(1), 478. <https://doi.org/10.1186/s13063-016-1613-9>
- Boutron, I., Altman, D. G., Moher, D., Schulz, K. F., & Ravaud, P. (2017). CONSORT Statement for Randomized Trials of Nonpharmacologic Treatments : A 2017 Update and a CONSORT Extension for Nonpharmacologic Trial Abstracts. *Annals of Internal Medicine*, 167(1), 40-47. <https://doi.org/10.7326/M17-0046>
- Cambon, L., & Alla, F. (2019). Current challenges in population health intervention research. *J Epidemiol Community Health*, 73(11), 990-992. <https://doi.org/10.1136/jech-2019-212225>
- Campbell, D. T. (1971). Reforms as Experiments. *Urban Affairs Quarterly*, 7(2), 133-171. <https://doi.org/10.1177/107808747100700202>
- Campbell, D. T., & Stanley, J. C. (1966). *Experimental and Quasi-experimental Designs for Research*. R. McNally.
- Cartwright, N., & Hardie, J. (2012a). Evidence-Ranking Schemes, Advice Guides, and Choosing Effective Policies. In *Evidence-Based Policy : A Practical Guide to Doing*

- It Better* (p. 135-143). Oxford University Press.
<https://oxford.universitypressscholarship.com/view/10.1093/acprof:osobl/9780199841608.001.0001/acprof-9780199841608>
- Cartwright, N., & Hardie, J. (2012b). What are RCTs good for? In *Evidence-Based Policy : A Practical Guide to Doing It Better* (p. 122-134). Oxford University Press.
<https://oxford.universitypressscholarship.com/view/10.1093/acprof:osobl/9780199841608.001.0001/acprof-9780199841608>
- Chalmers, I. (2003). Trying to do more Good than Harm in Policy and Practice : The Role of Rigorous, Transparent, Up-to-Date Evaluations. *The ANNALS of the American Academy of Political and Social Science*, 589(1), 22-40.
<https://doi.org/10.1177/0002716203254762>
- Chen, H.-T. (2010). The bottom-up approach to integrative validity : A new perspective for program evaluation. *Evaluation and Program Planning*, 33(3), 205-214.
<https://doi.org/10.1016/j.evalprogplan.2009.10.002>
- Chen, H.-T. (2015a). Evaluation and Dissemination Top-Down Approach Versus Bottom-Up Approach. In *Practical program evaluation : Theory-driven evaluation and the integrated evaluation perspective* (Second edition, p. 388-411). SAGE Publications.
- Chen, H.-T. (2015b). Fundamentals of Program evaluation. In *Practical program evaluation : Theory-driven evaluation and the integrated evaluation perspective* (Second edition, p. 3-34). SAGE Publications.
- Chen, H.-T. (2015c). Logic Models and the Action Model/ Change Model Schema (Program Theory). In *Practical program evaluation : Theory-driven evaluation and the integrated evaluation perspective* (Second edition, p. 58-93). SAGE Publications.
- Chen, H.-T. (2015d). *Practical program evaluation : Theory-driven evaluation and the integrated evaluation perspective* (Second edition). SAGE Publications.
- Chen, H.-T. (2015e). The Holistic Effectuality Evaluation Approach to Outcome Evaluation. In *Practical program evaluation : Theory-driven evaluation and the integrated evaluation perspective* (Second edition, p. 260-304). SAGE Publications.
- Chen, H.-T. (2015f). The Theory-Driven Approach to Outcome Evaluation. In *Practical program evaluation : Theory-driven evaluation and the integrated evaluation perspective* (Second edition, p. 304-339). SAGE Publications.
- Christie, C. A., & Alkin, M. C. (2013). An evaluation theory tree. In M. C. Alkin (Éd.), *Evaluation roots : A wider perspective of theorists' views and influences* (2nd ed, p. 20-64). SAGE Publications.

- Coates, T. J., Richter, L., & Caceres, C. (2008a). Behavioural strategies to reduce HIV transmission: How to make them work better. *The Lancet*, 372(9639), 669-684. [https://doi.org/10.1016/S0140-6736\(08\)60886-7](https://doi.org/10.1016/S0140-6736(08)60886-7)
- Coates, T. J., Richter, L., & Caceres, C. (2008b). Behavioural strategies to reduce HIV transmission: How to make them work better. *The Lancet*, 372(9639), 669-684. [https://doi.org/10.1016/S0140-6736\(08\)60886-7](https://doi.org/10.1016/S0140-6736(08)60886-7)
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation : Design & Analysis Issues for Field Settings*. Houghton Mifflin.
- Cronbach, L. J., & Shapiro, K. (1982). *Designing evaluations of educational and social programs* (1st ed). Jossey-Bass.
- Dahlgren, G., & Whitehead, M. (1991). Policies and strategies to promote social equity in health. Background document to WHO - Strategy paper for Europe. *Arbetsrapport*, Article 2007:14. https://ideas.repec.org/p/hhs/ifswps/2007_014.html
- Dahlgren, G., & Whitehead, M. (2021). The Dahlgren-Whitehead model of health determinants: 30 years on and still chasing rainbows. *Public Health*, 199, 20-24. <https://doi.org/10.1016/j.puhe.2021.08.009>
- Deaton, A., & Cartwright, N. (2018). Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine*, 210, 2-21. <https://doi.org/10.1016/j.socscimed.2017.12.005>
- de Salazar, L., & Hall, M. (2009). Developing Evaluation Questions : Beyond the Technical Issues. In L. Potvin, D. V. McQueen, M. Hall, L. de Salazar, L. M. Anderson, & Z. M. A. Hartz (Éds.), *Health Promotion Evaluation Practices in the Americas : Values and Research* (p. 49-62). Springer. https://doi.org/10.1007/978-0-387-79733-5_4
- Desgrées-du-Loû, A., Pannetier, J., Ravalihasy, A., Gosselin, A., Supervie, V., Panjo, H., Bajos, N., Lert, F., Lydié, N., Dray-Spira, R., & Group5, T. P. S. (2015). Sub-Saharan African migrants living with HIV acquired after migration, France, ANRS PARCOURS study, 2012 to 2013. *Eurosurveillance*, 20(46), 30065. <https://doi.org/10.2807/1560-7917.ES.2015.20.46.30065>
- Desgrées-du-Lou, A., Pannetier, J., Ravalihasy, A., Le Guen, M., Gosselin, A., Panjo, H., Bajos, N., Lydie, N., Lert, F., Dray-Spira, R., & PARCOURS Study Group. (2016). Is hardship during migration a determinant of HIV infection? Results from the ANRS PARCOURS study of sub-Saharan African migrants in France. *AIDS (London, England)*, 30(4), 645-656. <https://doi.org/10.1097/QAD.0000000000000957>
- Devaux-Spatarakis, A. (2014). L'évaluation « basée sur la théorie », entre rigueur scientifique et contexte politique. *Politiques et management public*, Vol 31/1, Article Vol 31/1.

- Devaux-Spatarakis, A., Delahais, T., & Ridde, A. R. et V. (2021). *Introduction : À quoi sert l'évaluation?*
<https://scienceetbiencommun.pressbooks.pub/evaluationanthologie/chapter/introduction-a-quoi-sert-levaluation/>
- Diderichsen, F., Evans, T., & Whitehead, M. (2001). The Social Basis of Disparities in Health. In T. Evans, M. Whitehead, F. Diderichsen, A. Bhuiya, & M. Wirth (Éds.), *Challenging Inequities in Health : From Ethics to Action* (p. 0). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195137408.003.0002>
- Diderichsen, F., Hallqvist, J., & Whitehead, M. (2019). Differential vulnerability and susceptibility : How to make use of recent development in our understanding of mediation and interaction to tackle health inequalities. *International Journal of Epidemiology*, 48(1), 268-274. <https://doi.org/10.1093/ije/dyy167>
- Doll, R. (1998). Controlled trials : The 1948 watershed. *BMJ*, 317(7167), 1217-1220. <https://doi.org/10.1136/bmj.317.7167.1217>
- Doyle, A. M., Ross, D. A., Maganja, K., Baisley, K., Masesa, C., Andreasen, A., Plummer, M. L., Obasi, A. I. N., Weiss, H. A., Kapiga, S., Watson-Jones, D., Changalucha, J., Hayes, R. J., & Grp, M. K. V. T. S. (2010). Long-Term Biological and Behavioural Impact of an Adolescent Sexual Health Intervention in Tanzania : Follow-up Survey of the Community-Based MEMA kwa Vijana Trial. *PLOS MEDICINE*, 7(6). <https://doi.org/10.1371/journal.pmed.1000287>
- Fisher, R. A. (1925). *Statistical Methods for Research Workers*. Oliver and Boyd.
- Fisher, R. A. (1971). *The Design of Experiments*.
- Fisher, R. A., & Bartlett, M. S. (1931). Pasteurised and Raw Milk. *Nature*, 127(3207), Article 3207. <https://doi.org/10.1038/127591c0>
- Flay, B. R. (1986). Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Preventive Medicine*, 15(5), 451-474. [https://doi.org/10.1016/0091-7435\(86\)90024-1](https://doi.org/10.1016/0091-7435(86)90024-1)
- Flay, B. R., Biglan, A., Boruch, R. F., Castro, F. G., Gottfredson, D., Kellam, S., Mościcki, E. K., Schinke, S., Valentine, J. C., & Ji, P. (2005). Standards of Evidence : Criteria for Efficacy, Effectiveness and Dissemination. *Prevention Science*, 6(3), 151-175. <https://doi.org/10.1007/s11121-005-5553-y>
- Food and Drug Administration. (1992). Guidelines for the clinical evaluation of analgesic drugs. *Pharmakopsychiatrie, Neuro-Psychopharmakologie*, 13(6), 344-350.

- Francis Jr, T., Korns, R. F., Voight, R. B., Boisen, M., Hemphill, F. M., Napier, J. A., & Tolchinsky, E. (1955). An evaluation of the 1954 poliomyelitis vaccine trials. *American journal of public health and the nation's health*, 45(5 Pt 2), 1-63.
- Francis-Oliviero, F., Cambon, L., Wittwer, J., Marmot, M., & Alla, F. (2020). Theoretical and practical challenges of proportionate universalism : A review. *Revista Panamericana de Salud Pública*, 44, e110. <https://doi.org/10.26633/RPSP.2020.110>
- Gautier, L. (2022). Towards CR² evaluation : Culturally -reflexive and -responsive evaluation in crises times and beyond. *The International Journal of Health Planning and Management*, hpm.3579. <https://doi.org/10.1002/hpm.3579>
- Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. J. (2016a). *Impact Evaluation in Practice, Second Edition*. Washington, DC: Inter-American Development Bank and World Bank. <https://doi.org/10.1596/978-1-4648-0779-4>
- Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. J. (2016b). Why evaluate? In *Impact Evaluation in Practice, Second Edition*. Washington, DC: Inter-American Development Bank and World Bank. <http://hdl.handle.net/10986/25030>
- Gill, S., Kuwahara, R., & Wilce, M. (2016). Through a Culturally Competent Lens : Why the Program Evaluation Standards Matter. *Health Promotion Practice*, 17(1), 5-8.
- Globerman, J., Mitra, S., Gogolishvili, D., Rueda, S., Schoffel, L., Gangbar, K., Shi, Q., & Rourke, S. B. (2017). HIV/STI Prevention Interventions : A Systematic Review and Meta-analysis. *Open Medicine*, 12, 450-467. <https://doi.org/10.1515/med-2017-0064>
- Gosselin, A., Carillon, S., Coulibaly, K., Ridde, V., Taéron, C., Kohou, V., Zouménou, I., Mbiribindi, R., Derche, N., Desgrées du Loû, A., Gubert, F., Melchior, M., Delbe, A., Ebongue, J., El Khoury, F., Gaywahali, C., Lert, F., Lutonadio, B., Plenel, E., ... for the MAKASI Study Group. (2019). Participatory development and pilot testing of the Makasi intervention : A community-based outreach intervention to improve sub-Saharan and Caribbean immigrants' empowerment in sexual health. *BMC Public Health*, 19(1), 1646. <https://doi.org/10.1186/s12889-019-7943-2>
- Gwatkin, D. R. (2017). Trends in health inequalities in developing countries. *The Lancet Global Health*, 5(4), e371-e372. [https://doi.org/10.1016/S2214-109X\(17\)30080-3](https://doi.org/10.1016/S2214-109X(17)30080-3)
- Habicht, J. P., Victora, C. G., & Vaughan, J. P. (1999). Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *International Journal of Epidemiology*, 28(1), 10-18. <https://doi.org/10.1093/ije/28.1.10>

- Hall, N. S. (2007). R. A. Fisher and his advocacy of randomization. *Journal of the History of Biology*, 40(2), 295-325. <https://doi.org/10.1007/s10739-006-9119-z>
- Hammersley, M. (2005). Is the evidence-based practice movement doing more good than harm? Reflections on Iain Chalmers' case for research-based policy making and practice. *Evidence & Policy*, 1(1), 85-100. <https://doi.org/10.1332/1744264052703203>
- Hemming, K., Haines, T. P., Chilton, P. J., Girling, A. J., & Lilford, R. J. (2015). The stepped wedge cluster randomised trial : Rationale, design, analysis, and reporting. *BMJ*, 350, h391. <https://doi.org/10.1136/bmj.h391>
- Hemming, K., Taljaard, M., McKenzie, J. E., Hooper, R., Copas, A., Thompson, J. A., Dixon-Woods, M., Aldcroft, A., Doussau, A., Grayling, M., Kristunas, C., Goldstein, C. E., Campbell, M. K., Girling, A., Eldridge, S., Campbell, M. J., Lilford, R. J., Weijer, C., Forbes, A. B., & Grimshaw, J. M. (2018). Reporting of stepped wedge cluster randomised trials : Extension of the CONSORT 2010 statement with explanation and elaboration. *BMJ*, 363, k1614. <https://doi.org/10.1136/bmj.k1614>
- Hill, A. B. (1965). The Environment and Disease : Association or Causation? *Proceedings of the Royal Society of Medicine*, 58(5), 295-300.
- Holland, P. W. (1986). Statistics and Causal Inference. *Journal of the American Statistical Association*, 81(396), 945-960. <https://doi.org/10.2307/2289064>
- Iskarpatyoti, B. S., Lebov, J., Hart, L., Thomas, J., & Mandal, M. (2018). Evaluations of Structural Interventions for HIV Prevention : A Review of Approaches and Methods. *AIDS and Behavior*, 22(4), 1253-1264. <https://doi.org/10.1007/s10461-017-1997-x>
- Jamal, F., Fletcher, A., Shackleton, N., Elbourne, D., Viner, R., & Bonell, C. (2015). The three stages of building and testing mid-level theories in a realist RCT : A theoretical and methodological case-example. *Trials*, 16(1), 466. <https://doi.org/10.1186/s13063-015-0980-y>
- Joint United Nations Programme on HIV/AIDS. (2007). *Combination HIV Prevention : Tailoring and Coordinating Biomedical, Behavioural and Structural Strategies to Reduce New HIV Infections*. https://www.unaids.org/sites/default/files/media_asset/JC2007_Combination_Prevention_paper_en_0.pdf
- Joint United Nations Programme on HIV/AIDS. (2010). *Strategic Guidance for Evaluating HIV Prevention Programmes*. UNAIDS, Geneva. https://www.unaids.org/sites/default/files/sub_landing/files/12_7_MERG_Guidance_Evaluating%20HIV_PreventionProgrammes.pdf

- Joint United Nations Programme on HIV/AIDS. (2015). *Fast-tracking combination prevention*. UNAIDS, Geneva.
https://www.unaids.org/sites/default/files/media_asset/20151019_JC2766_Fast-tracking_combination_prevention.pdf
- Joint United Nations Programme on HIV/AIDS. (2021). *Unequal, unprepared, under threat : Why bold action against inequalities is needed to end AIDS, stop COVID-19 and prepare for future pandemics*.
https://www.unaids.org/sites/default/files/media_asset/2021_WAD_report_en.pdf
- Joint United Nations Programme on HIV/AIDS. (2023). *Global HIV & AIDS statistics—Fact sheet*. <https://www.unaids.org/en/resources/fact-sheet>
- Julnes, G., & Rog, D. J. (2007). Editors' notes : Current federal policies and controversies over methodology in evaluation. *New Directions for Evaluation*, 2007(113), 1-12.
<https://doi.org/10.1002/ev.212>
- Junod, S. (2008). FDA and Clinical Drug Trials : A Short History. *FDLI Update*, 2008, 55.
- Kurth, A. E., Celum, C., Baeten, J. M., Vermund, S. H., & Wasserheit, J. N. (2011). Combination HIV Prevention : Significance, Challenges, and Opportunities. *Current HIV/AIDS reports*, 8(1), 62-72. <https://doi.org/10.1007/s11904-010-0063-3>
- Linfield, K. J., & Posavac, E. J. (2018). *Program evaluation : Methods and case studies* (9th edition). Routledge.
- Louart, S., Baldé, H., & Ridde, É. R. et V. (2023). *Évaluation réaliste*. <https://scienceetbiencommun.pressbooks.pub/evaluationpolpub/chapter/evaluation-realiste/>
- Marmot, M. (2004). Evidence based policy or policy based evidence? *BMJ*, 328(7445), 906-907. <https://doi.org/10.1136/bmj.328.7445.906>
- Marmot, M., & Bell, R. (2012). Fair society, healthy lives. *Public Health*, 126, S4-S10.
<https://doi.org/10.1016/j.puhe.2012.05.014>
- McQueen, D. V., & Anderson, L. M. (2001). What counts as evidence : Issues and debates. In *Evaluation in Health Promotion : Principles and Perspectives* (WHO Regional Publications, p. 63-81). WHO Regional Office Europe.
- Medical Research Council. (1948). Streptomycin Treatment of Pulmonary Tuberculosis. *British Medical Journal*, 2(4582), 769-782.
- Miller, W. R., & Rollnick, S. (2013). *L'entretien motivationnel - 2e éd. - Aider la personne à engager le changement : Aider la personne à engager le changement* (2e édition). InterEditions.

- Millington, R. (2010, novembre 23). Different types of communities. *FeverBee*.
<https://www.feverbee.com/different-types-of-communities/>
- Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J., Elbourne, D., Egger, M., & Altman, D. G. (2010). CONSORT 2010 Explanation and Elaboration : Updated guidelines for reporting parallel group randomised trials. *BMJ*, *340*. <https://doi.org/10.1136/bmj.c869>
- Moher, D., Schulz, K. F., & Altman, D. G. (2001). The CONSORT statement : Revised recommendations for improving the quality of reports of parallel-group randomised trials. *The Lancet*, *357*(9263), 1191-1194. [https://doi.org/10.1016/S0140-6736\(00\)04337-3](https://doi.org/10.1016/S0140-6736(00)04337-3)
- Montgomery, P., Grant, S., Mayo-Wilson, E., Macdonald, G., Michie, S., Hopewell, S., Moher, D., Lawrence Aber, J., Altman, D., Bhui, K., Booth, A., Clark, D., Craig, P., Eisner, M., Fraser, M. W., Gardner, F., Grant, S., Hedges, L., Hollon, S., ... on behalf of the CONSORT-SPI Group. (2018). Reporting randomised trials of social and psychological interventions : The CONSORT-SPI 2018 Extension. *Trials*, *19*(1), 407. <https://doi.org/10.1186/s13063-018-2733-1>
- Newcomer, K. E., Hatry, H. P., & Wholey, J. S. (Éds.). (2015). *Handbook of Practical Program Evaluation : Newcomer/Handbook*. John Wiley & Sons, Inc. <https://doi.org/10.1002/9781119171386>
- Neyman, J., Dabrowska, D. M., & Speed, T. P. (1990). On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9. *Statistical Science*, *5*(4), 465-472.
- Ninacs, W. A. (2003). Empowerment : Cadre conceptuel et outil d'évaluation de l'intervention sociale et communautaire. *La clé: La coopérative de consultation en développement*, 26.
- O'Neill, M. (2003). Pourquoi se préoccupe-t-on tant des données probantes en promotion de la santé? *Sozial- und Präventivmedizin/Social and Preventive Medicine*, *48*(5), 317-326. <https://doi.org/10.1007/s00038-003-2036-4>
- Organisation mondiale de la Santé. (1946). *Constitution de l'Organisation Mondiale de la Santé*. Conférence internationale de la santé, New York.
- Organisation mondiale de la Santé. (1978). *Déclaration d'Alma-Ata* (WHO/EURO:1978-3938-43697-61473). Organisation mondiale de la Santé. <https://apps.who.int/iris/handle/10665/347880>

- Organisation mondiale de la Santé. (1986). *Promotion de la santé : Charte d'Ottawa* (WHO/EURO:1986-4044-43803-61666). Organisation mondiale de la Santé. <https://apps.who.int/iris/handle/10665/349653>
- Padian, N. S., Buvé, A., Balkus, J., Serwadda, D., & Cates, W. (2008). Biomedical interventions to prevent HIV infection : Evidence, challenges, and way forward. *The Lancet*, *372*(9638), 585-599. [https://doi.org/10.1016/S0140-6736\(08\)60885-5](https://doi.org/10.1016/S0140-6736(08)60885-5)
- Padian, N. S., McCoy, S. I., Balkus, J. E., & Wasserheit, J. N. (2010). Weighing the gold in the gold standard : Challenges in HIV prevention research. *AIDS*, *24*(5), 621-635. <https://doi.org/10.1097/QAD.0b013e328337798a>
- Parkhurst, J. O. (2014). Structural approaches for prevention of sexually transmitted HIV in general populations : Definitions and an operational approach. *Journal of the International AIDS Society*, *17*(1), 19052. <https://doi.org/10.7448/IAS.17.1.19052>
- Pawson, R. (2020). The coronavirus response : Boxed in by models: *Evaluation*. <https://doi.org/10.1177/1356389020968579>
- Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. Sage.
- Peak, A., Rana, S., Maharjan, S. H., Jolley, D., & Crofts, N. (1995). Declining risk for HIV among injecting drug users in Kathmandu, Nepal : The impact of a harm-reduction programme. *AIDS*, *9*(9), 1067-1070. <https://doi.org/10.1097/00002030-199509000-00013>
- Pearl, J. (2000). *Causality : Models, reasoning, and inference*. Cambridge University Press.
- Pearl, J. (2009). Causal inference in statistics : An overview. *Statistics Surveys*, *3*(none), 96-146. <https://doi.org/10.1214/09-SS057>
- Petticrew, M. (2011). When are complex interventions 'complex'? When are simple interventions 'simple'? *European Journal of Public Health*, *21*(4), 397-398. <https://doi.org/10.1093/eurpub/ckr084>
- Petticrew, M., Cummins, S., Ferrell, C., Findlay, A., Higgins, C., Hoy, C., Kearns, A., & Sparks, L. (2005). Natural experiments : An underused tool for public health? *Public Health*, *119*(9), 751-757. <https://doi.org/10.1016/j.puhe.2004.11.008>
- Piot, P., Bartos, M., Larson, H., Zewdie, D., & Mane, P. (2008). Coming to terms with complexity : A call to action for HIV prevention. *The Lancet*, *372*(9641), 845-859. [https://doi.org/10.1016/S0140-6736\(08\)60888-0](https://doi.org/10.1016/S0140-6736(08)60888-0)
- Pocock, S. J., Assmann, S. E., Enos, L. E., & Kasten, L. E. (2002). Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting : Current practice and problems. *Statistics in Medicine*, *21*(19), 2917-2930. <https://doi.org/10.1002/sim.1296>

- Potvin, L., & McQueen, D. V. (2009). Practical Dilemmas for Health Promotion Evaluation. In L. Potvin, D. V. McQueen, M. Hall, L. de Salazar, L. M. Anderson, & Z. M. A. Hartz (Éds.), *Health Promotion Evaluation Practices in the Americas : Values and Research* (p. 25-45). Springer. https://doi.org/10.1007/978-0-387-79733-5_3
- Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., Griffey, R., & Hensley, M. (2011). Outcomes for Implementation Research: Conceptual Distinctions, Measurement Challenges, and Research Agenda. *Administration and Policy in Mental Health and Mental Health Services Research*, 38(2), 65-76. <https://doi.org/10.1007/s10488-010-0319-7>
- Ravalihasy, A., Ante-Testard, P. A., Kardas-Sloma, L., Yazdanpanah, Y., De Allegri, M., & Ridde, V. (2023). Quantitative Methods Used to Evaluate Impact of Combination HIV Prevention Intervention : A Methodological Systematic Review. *AIDS and Behavior*. <https://doi.org/10.1007/s10461-023-04000-8>
- Reichardt, C. S. (2019a). Cause and effect. In *Quasi-experimentation : A guide to design and analysis* (p. 11-25). The Guilford Press.
- Reichardt, C. S. (2019b). Interrupted Time-Series Designs. In *Quasi-experimentation : A guide to design and analysis* (p. 202-246). The Guilford Press.
- Reichardt, C. S. (2019c). Introduction. In *Quasi-experimentation : A guide to design and analysis* (p. 1-10). The Guilford Press.
- Reichardt, C. S. (2019d). Nonequivalent Group Designs. In *Quasi-experimentation : A guide to design and analysis* (p. 112-162). The Guilford Press.
- Reichardt, C. S. (2019e). One-group posttest-only designs. In *Quasi-experimentation : A guide to design and analysis* (p. 94-98). The Guilford Press.
- Reichardt, C. S. (2019f). Pretest–posttest Designs. In *Quasi-experimentation : A guide to design and analysis* (p. 99-111). The Guilford Press.
- Reichardt, C. S. (2019g). *Quasi-experimentation : A guide to design and analysis*. The Guilford Press.
- Reichardt, C. S. (2019h). Randomized experiments. In *Quasi-experimentation : A guide to design and analysis* (p. 45-93). The Guilford Press.
- Reichardt, C. S. (2019i). Regression Discontinuity Designs. In *Quasi-experimentation : A guide to design and analysis* (p. 163-201). The Guilford Press.
- Reichardt, C. S. (2019j). Threats to validity. In *Quasi-experimentation : A guide to design and analysis* (p. 26-44). The Guilford Press.

- Rootman, I., Goodstadt, M., Potvin, L., & Springett, J. (2001). A framework for health promotion evaluation. In *Evaluation in Health Promotion: Principles and Perspectives* (WHO Regional Publications, p. 7-38). WHO Regional Office Europe.
- Rosenbaum, P. R. (1984). From Association to Causation in Observational Studies : The Role of Tests of Strongly Ignorable Treatment Assignment. *Journal of the American Statistical Association*, 79(385), 41-48. <https://doi.org/10.2307/2288332>
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55. <https://doi.org/10.1093/biomet/70.1.41>
- Rossi, P. H., Lipsey, M. W., & Henry, G. T. (2019). *Evaluation : A systematic approach* (Eighth edition). SAGE.
- Rubin, D. B. (1972). Estimating Causal Effects of Treatments in Experimental and Observational Studies. *ETS Research Bulletin Series*, 1972(2), i-31. <https://doi.org/10.1002/j.2333-8504.1972.tb00631.x>
- Rubin, D. B. (1977). Assignment to Treatment Group on the Basis of a Covariate. *Journal of Educational Statistics*, 2(1), 1-26. <https://doi.org/10.2307/1164933>
- Rubin, D. B. (2005). Causal Inference Using Potential Outcomes. *Journal of the American Statistical Association*, 100(469), 322-331. <https://doi.org/10.1198/016214504000001880>
- Sabido, M., Giardina, F., Hernandez, G., Hugo Fernandez, V., Ernesto Monzon, J., Ortiz, R., Montoliu, A., Casabona, J., & Team, U. (2009). The UALE Project : Decline in the Incidence of HIV and Sexually Transmitted Infections and Increase in the Use of Condoms Among Sex Workers in Guatemala. *JAIDS-JOURNAL OF ACQUIRED IMMUNE DEFICIENCY SYNDROMES*, 51(1), S35-S41. <https://doi.org/10.1097/QAI.0b013e3181a2656f>
- Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine : What it is and what it isn't. *BMJ*, 312(7023), 71-72. <https://doi.org/10.1136/bmj.312.7023.71>
- Saltelli, A., & Funtowicz, S. (2014, janvier 1). When All Models Are Wrong. *Issues in Science and Technology*. <https://issues.org/andrea/>
- Saltelli, A., & Giampietro, M. (2017). What is wrong with evidence based policy, and how can it be improved? *Futures*, 91, 62-71. <https://doi.org/10.1016/j.futures.2016.11.012>
- Saltelli, A., Guimarães Pereira, Â., Van der Sluijs, J. P., & Funtowicz, S. (2013). What do I make of your latinorum? Sensitivity auditing of mathematical modelling. *International*

- Journal of Foresight and Innovation Policy*, 9(2-3-4), 213-234.
<https://doi.org/10.1504/IJFIP.2013.058610>
- Santé Publique France. (2017). *Infection par le VIH et les IST bactériennes. Point épidémiologique du 28 novembre 2017*. <http://invs.santepubliquefrance.fr/Dossiers-thematiques/Maladies-infectieuses/VIH-sida-IST/Infection-a-VIH-et-sida/Actualites/Infection-par-le-VIH-et-les-IST-bacteriennes.-Point-epidemiologique-du-28-novembre-2017>
- Schulz, K. F., Altman, D. G., & Moher, D. (2010). CONSORT 2010 Statement : Updated guidelines for reporting parallel group randomised trials. *BMJ*, 340. <https://doi.org/10.1136/bmj.c332>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
- Shadish, W. R., & Luellen, J. K. (2013). Donald Campbell : The Accidental Evaluator. In M. C. Alkin (Éd.), *Evaluation roots : A wider perspective of theorists' views and influences* (2nd ed, p. 20-64). SAGE Publications.
- Sussman, S., Valente, T. W., Rohrbach, L. A., Skara, S., & Ann Pentz, M. (2006). Translation in the Health Professions : Converting Science into Action. *Evaluation & the Health Professions*, 29(1), 7-32. <https://doi.org/10.1177/0163278705284441>
- Teutsch, S. M. (1992). A framework for assessing the effectiveness of disease and injury prevention. *MMWR. Recommendations and Reports: Morbidity and Mortality Weekly Report. Recommendations and Reports*, 41(RR-3), 1-12.
- Thilakavathi, S., Boopathi, K., Girish Kumar, C. P., Santhakumar, A., Senthilkumar, R., Eswaramurthy, C., Ilaya Bharathy, V., Ramakrishnan, L., Thongamba, G., Adhikary, R., & Paranjape, R. (2011). Assessment of the scale, coverage and outcomes of the Avahan HIV prevention program for female sex workers in Tamil Nadu, India : Is there evidence of an effect? *BMC Public Health*, 11 Suppl 6(Suppl 6), S3. <https://doi.org/10.1186/1471-2458-11-S6-S3>
- Traore, I. T., Meda, N., Hema, N. M., Ouedraogo, D., Some, F., Some, R., Niessougou, J., Sanon, A., Konate, I., Van De Perre, P., Mayaud, P., & Nagot, N. (2015). HIV prevention and care services for female sex workers : Efficacy of a targeted community-based intervention in Burkina Faso. *JOURNAL OF THE INTERNATIONAL AIDS SOCIETY*, 18. <https://doi.org/10.7448/IAS.18.1.20088>
- Van Belle, S., Marchal, B., Dubourg, D., & Kegels, G. (2010). How to develop a theory-driven evaluation design? Lessons learned from an adolescent sexual and reproductive health

- programme in West Africa. *BMC Public Health*, 10(1), 741. <https://doi.org/10.1186/1471-2458-10-741>
- Van Belle, S., Wong, G., Westhorp, G., Pearson, M., Emmel, N., Manzano, A., & Marchal, B. (2016). Can “realist” randomised controlled trials be genuinely realist? *Trials*, 17(1), 313. <https://doi.org/10.1186/s13063-016-1407-0>
- Wallerstein, N. (1992). Powerlessness, empowerment, and health : Implications for health promotion programs. *American Journal of Health Promotion: AJHP*, 6(3), 197-205. <https://doi.org/10.4278/0890-1171-6.3.197>
- Williams, J. R., Alary, M., Lowndes, C. M., Behanzin, L., Labbe, A.-C., Anagonou, S., Ndour, M., Minani, I., Ahoussinou, C., Zannou, D. M., & Boily, M.-C. (2014). Positive Impact of Increases in Condom Use among Female Sex Workers and Clients in a Medium HIV Prevalence Epidemic : Modelling Results from Project SIDA1/2/3 in Cotonou, Benin. *PLOS ONE*, 9(7). <https://doi.org/10.1371/journal.pone.0102643>
- Ye, S., Xiao, Y., Jin, C., Cassell, H., Blevins, M., Sun, J., Vermund, S. H., & Qian, H.-Z. (2012). Effectiveness of integrated HIV prevention interventions among Chinese men who have sex with men : Evaluation of a 16-city public health program. *PLoS One*, 7(12), e50873. <https://doi.org/10.1371/journal.pone.0050873>
- Yin, R. K. (2010). Analytic Generalization. In A. J. Mills, G. Durepos, & E. Wiebe (Éds.), *Encyclopedia of case study research*. Sage.
- Yin, R. K. (2018). *Case study research and applications : Design and methods* (Sixth edition). SAGE.

Annexes

Annexe 1

Article 1 : Protocole pour une revue systématique des méthodes quantitatives utilisées pour mesurer l'impact des PPCV sur la transmission du VIH

PROTOCOL

Open Access



Quantitative methods used to evaluate impact of health promotion interventions to prevent HIV infections: a methodological systematic review protocol

Andrainolo Ravalihasy^{1,2,3*} , Lidia Kardaś-Słoma⁴, Yazdan Yazdanpanah⁴ and Valéry Ridde^{1,2,3}

Abstract

Background: Combination prevention is currently considered the best approach to combat HIV epidemic. It is based upon the combination of structural, behavioral, and biomedical interventions. Such interventions are frequently implemented in a health-promoting manner due to their aims, the approach that was adopted, and their complexity. The impact evaluation of these interventions often relies on methods inherited from the biomedical field. However, these methods have limitations and should be adapted to be relevant for these complex interventions. This systematic review aims to map the evidence-based methods used to quantify the impact of these interventions and analyze how these methods are implemented.

Methods: Three databases (Web of Science, Scopus, PubMed) will be used to identify impact evaluation studies of health promotion interventions that aimed at reducing the incidence or prevalence of HIV infection. Only studies based on quantitative design assessing intervention impact on HIV prevalence or incidence will be included. Two reviewers will independently screen studies based on titles and abstracts and then on the full text. The information about study characteristics will be extracted to understand the context in which the interventions are implemented. The information specific to quantitative methods of impact evaluation will be extracted using items from the Mixed Methods Appraisal Tool (MMAT), the guidelines for reporting Statistical Analyses and Methods in the Published Literature (SAMPL), and the guidelines for Strengthening The Reporting of Empirical Simulation Studies (STRESS). This review will be conducted according to the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) statement.

Discussion: The impact evaluation of HIV prevention interventions is a matter of substantial importance given the growing need for evidence of the effectiveness of these interventions, whereas they are increasingly complex. These evaluations allow to identify the most effective strategies to be implemented to fight the epidemic. It is therefore relevant to map the methods to better implement them and adapt them according to the type of intervention to be evaluated.

*Correspondence: andrainolo.ravalihasy@ird.fr

¹ Centre Population et Développement (IRD, Université Paris Cité, Inserm ERL 1244 SAGESUD), 75006 Paris, France
Full list of author information is available at the end of the article



Systematic review registration: PROSPERO [CRD42020210825](https://www.crd42020210825)

Keywords: Evidence-based impact evaluation, Combination HIV prevention, Health promotion, Study design, Statistical methods, Mathematical methods

Background

The importance of health promotion interventions is acknowledged, thanks to their impacts on social determinants of health and the need to fight against health inequities [1, 2]. Despite this recognition, their effectiveness is often debated and criticized due to the lack of available evidence from impact evaluations using quantitative data, especially randomized controlled trials. However, these methods, inherited from the biomedical field, have several methodological and practical limitations for quantifying the impact of this type of intervention and should be adapted to be relevant [3, 4].

In the context of the fight against HIV epidemic, the lack of curative treatment has brought health promotion interventions at the forefront, particularly those aimed at preventing the transmission of the virus. Currently, biomedical interventions alone have shown their limits, and combination prevention is considered the best approach to curb the epidemic [5, 6]. It is based upon the combination of structural, behavioral, and biomedical interventions that address specific prevention needs at the individual and community levels. The interventions which fall within this framework are frequently implemented in a health-promoting manner due to their aims, the approach adopted, and their complexity. Over the past several years, such interventions have been proposed and evaluated for effectiveness in a variety of ways. Among those evaluated using evidence-based quantitative methods, some have been shown to be effective, while others have not. The challenge of impact evaluation of these interventions is paramount in helping to identify and decide which programs and policies should be implemented and supported to address the epidemic. It is essential that these evaluations are conducted in a rigorous manner which is appropriate to the nature and the complexity of health promotion interventions. Hence, we decided to conduct a review to identify and assess the implementation of the evidence-based impact of evaluation methods that are used to assess these interventions.

Objectives

This review is conducted to systematically review quantitative methods for evaluating the impact of interventions that aimed at preventing HIV transmission during

sexual exposure. The specific questions this review attempts to answer are as follows:

- What quantitative methods (statistical or mathematical) are used to quantify intervention impact on HIV incidence or prevalence?
- What are the designs of these studies?
- Are quantitative methods implemented appropriately?

Answering these questions will allow a critical synthesis of the methods used to assess the impact of the interventions concerned by this systematic review and more broadly to identify directions for methodological development to adapt these methods when assessing the impact of interventions designed to address social and behavioral determinants of health.

Methods/design

This systematic review is developed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols (PRISMA-P) statement [7] (Additional file 1). The protocol has been registered in the International Prospective Register of Systematic Reviews (PROSPERO): CRD42020210825.

Eligibility criteria

This review focuses on impact evaluation studies of health promotion interventions designed to prevent HIV transmission during sexual risk exposure. Health promotion is defined as a process of enabling individuals to improve their health and their control over their health [8]. For this reason, the studies to be included involve behavioral and/or structural components that may or may not be supplemented by the use of biomedical prevention tools. More specifically, this review will include studies based solely on existing interventions; studies based on hypothetical interventions and exclusively simulated data will be excluded. Hence, modeling studies may be included as long as the model inputs are based on data from an existing intervention to be evaluated. The outcomes studied in this review are quantitative methods based on strictly quantitative designs (randomized controlled intervention studies, nonrandomized controlled intervention studies, observational studies) that meet the classification proposed in Deeks

et al. [9] where controlled studies are those where the intervention assignment is done by researchers. Studies that do not meet these conditions will be excluded. Finally, the review focuses on interventions designed to reduce HIV transmission; eligible studies are those that assess intervention impact on HIV prevalence or incidence as a primary or secondary outcome.

Literature search

Five databases were considered in order to identify the studies to be included in the review: Web of Science, Scopus, PubMed, CINAHL, and PsycINFO. Finally, a sensitivity analysis of the search equation showed that three databases (Web of Science, Scopus, PubMed) were sufficient to uncover the studies relevant to this review. Moreover, it has been shown that the joint use of these databases allows to ensure an adequate literature search performance [10]. The search strategy consisted of using three groups of terms and two filters:

- i) Terms associated with HIV prevalence and incidence: to restrict the results to studies that aim to reduce HIV transmission
- ii) Terms associated with the notion of impact of interventions: to focus the search on articles referring to intervention impact evaluation
- iii) Terms associated with prevention interventions and sexual risk exposure: to focus the search on interventions designed to reduce the occurrence of new HIV cases and to focus the search on interventions targeting behavioral and structural barriers.
- iv) Restricting the search to articles written in English and French
- v) Restricting the search to scientific articles and not retracted publications

The search strategy is reported in Table 1 and was developed and tailored to each database with the assistance of a trained librarian.

Data collection

Study selection

References management will be done using a bibliographic management software (Zotero 5.0.96.2). Screening and data extraction will be done using a systematic review management software (Covidence). Duplicates will be removed automatically (using Covidence) and also manually. The study selection will be done in two steps: (i) a selection based on titles and abstracts and (ii) a selection based on full texts. At each stage of the selection, two reviewers will independently

select the studies taking into account the eligibility criteria. At the end of each stage, disagreements on which studies to include will be resolved with the help of a third senior reviewer. Only peer-reviewed scientific articles will be considered.

Data extraction

Two types of data will be extracted from the studies: general information about the included studies and specific information about the methods used to evaluate interventions.

Data concerning the characteristics of the studies These information will help to describe the included studies and the contexts in which the interventions are implemented. These information concern the authors, the title of the article, the date of publication, the place where the studies were carried out, the purpose, and the results of the studies.

Data on impact evaluation methods A data extraction grid was developed for the purposes of this review (Table 2). This grid has three sections that allow for the extraction of (i) information about the design of the studies, (ii) information about the statistical methods (when appropriate), and (iii) information about the mathematical methods (when appropriate). The items used in the extraction grid are based on tools and recommendations consistent with each of the abovementioned sections and with the investigated research question in the review [11–13]. We also verified that the information to be extracted is consistent with the standards of reporting according to the study design, the CONSORT statement and the relevant extensions for randomized controlled studies [14–17], the TREND statement for non-randomized controlled studies [18], and the STROBE statement for observational studies [19].

To test the applicability of the grid, a sample of 14 articles selected on the basis of their content was used (Additional file 2): 6 articles concerning randomized controlled studies (2 of which present mathematical impact evaluation), 3 articles concerning nonrandomized controlled studies, and 5 articles concerning observational studies (2 of which present mathematical impact evaluation). These articles were selected independently of the research questions addressed in the review to be conducted. The item development for each section of the extraction grid is presented below.

Study design information Eligible studies are based on strictly quantitative designs, i.e., randomized controlled studies, non-randomized controlled studies, or

Table 1 Search strategy

Databases	Search terms
Web of Science	<p>i) Terms associated with prevalence and incidence of HIV infection TS=(HIV OR "HIV infection" OR "human immunodeficiency virus") AND TS=(prevalence OR incidence)</p> <p>ii) Terms associated with the notion of effectiveness of interventions AND TS=(effect OR effects OR efficacy OR effectiveness OR impact*) AND TS=(intervention* OR program* OR project OR trial*)</p> <p>iii) Terms associated with prevention interventions and sexual risk exposure AND TS=(prevention OR prophylaxis OR "primary prevention") AND TS=("unsafe sex" OR "safe sex" OR "sex work" OR "sex worker" OR "sex workers" OR "unprotected sex" OR "protected sex" OR "high-risk sex" OR "high risk sex" OR "sexual risk")</p> <p>iv) Language AND LA=(English OR French)</p> <p>v) Document type NOT DT=("RETRACTED PUBLICATION" OR BOOK OR "book chapter" OR "PROCEEDINGS PAPER" OR "MEETING ABSTRACT")</p>
PubMed	<p>i) Terms associated with prevalence and incidence of HIV infection (HIV [Title/Abstract] OR "HIV infection"[Title/Abstract] OR "human immunodeficiency virus"[Title/Abstract]) AND (prevalence[Title/Abstract] OR incidence[Title/Abstract])</p> <p>ii) Terms associated with the notion of effectiveness of interventions AND (effect [Title/Abstract] OR effects [Title/Abstract] OR efficacy [Title/Abstract] OR effectiveness [Title/Abstract] OR impact*[Title/Abstract]) AND (intervention*[Title/Abstract] OR program*[Title/Abstract] OR project[Title/Abstract] OR trial*[Title/Abstract])</p> <p>iii) Terms associated with prevention interventions and sexual risk exposure AND (prevention [Title/Abstract] OR prophylaxis[Title/Abstract] OR "primary prevention"[MeSH]) AND ("unsafe sex"[MeSH] OR "safe sex"[MeSH] OR "sex work"[MeSH] OR "unprotected sex"[Title/Abstract] OR "protected sex"[Title/Abstract] OR "sexual risk"[Title/Abstract] OR "high-risk sex"[Title/Abstract] OR "high risk sex"[Title/Abstract])</p> <p>iv) Language AND (english[Filter] OR french[Filter])</p> <p>v) Document type NOT (bookdocs[Filter] OR congress[Filter] OR retractedpublication[Filter])</p>
Scopus	<p>i) Terms associated with prevalence and incidence of HIV infection TITLE-ABS-KEY(hiv OR "HIV infection" OR "human immunodeficiency virus") AND TITLE-ABS-KEY (prevalence OR incidence)</p> <p>ii) Terms associated with the notion of effectiveness of interventions AND TITLE-ABS-KEY(effect OR effects OR efficacy OR effectiveness OR impact*) AND TITLE-ABS-KEY (intervention* OR program* OR project OR trial*)</p> <p>iii) Terms associated with prevention interventions and sexual risk exposure AND TITLE-ABS-KEY (prevention OR prophylaxis OR "primary prevention") AND TITLE-ABS-KEY("unsafe sex" OR "safe sex" OR "sex work" OR "sex worker" OR "sex workers" OR "unprotected sex" OR "protected sex" OR "high-risk sex" OR "high risk sex" OR "sexual risk")</p> <p>iv) Language AND (LIMIT-TO (LANGUAGE, "English") OR LIMIT-TO (LANGUAGE, "French"))</p> <p>v) Document type AND (EXCLUDE(DOCTYPE, "tb") OR EXCLUDE(DOCTYPE, "ch") OR EXCLUDE(DOCTYPE, "cp"))</p>

observational studies. Items to extract study design-specific information originate from the Mixed Methods Appraisal Tool (MMAT). This tool allows to assess the design quality of quantitative studies [12, 20, 21]. Because these questions are specific to study design, the information to be extracted is different according to the study type. These items offer a choice of three possible answers (yes, no, unclear).

Information about statistical methods Items concerning statistical methods are formulated using the guidelines for reporting Statistical Analyses and Methods in

the Published Literature (SAMPL) [11]. The criteria identified in these guidelines are related to the implementation of statistical methods. The developed items allow the extraction of information common to different statistical methods for purpose to compare included studies. These items can be categorized into two types: (i) those that answer questions about the sample size and (ii) those that answer questions about the implementation of statistical methods in the analyses. The formulated items offer a choice of four possible answers (yes, no, imprecise, not concerned) except for one item allowing the classification of statistical methods (seven

Table 2 Quality assessment grid

Section	Items
Study design informations	<p>Randomized controlled trials</p> <p>Is randomization appropriately performed?</p> <p>Are the groups comparable at baseline?</p> <p>Are there complete outcome data?</p> <p>Are outcome assessors blinded to the intervention provided?</p> <p>Did the participants adhere to the assigned intervention?</p> <p>Non-randomized controlled trials/observational studies</p> <p>Are the participants representative of the target population?</p> <p>Are measurements appropriate regarding both the outcome and intervention (or exposure)?</p> <p>Are there complete outcome data?</p> <p>Are the confounders accounted for in the design and analysis?</p> <p>During the study period, is the intervention administered (or exposure occurred) as intended?</p>
Information about statistical methods	<p>Is the minimum difference considered important to prove the intervention efficacy/effectiveness reported?</p> <p>Are the alpha and beta levels that define statistical significance reported and respected?</p> <p>What is the statistical method used to assess the intervention impact on the primary outcome?</p> <p>What is statistical the method used to assess the intervention impact on HIV prevalence/incidence?</p> <p>Did the authors report analyses to confirm that the statistical assumptions of the analysis were met or the goodness of fit of the model?</p> <p>Did the authors report a measure of precision alongside the impact measure?</p> <p>Is the design taken into account in the statistical analyses?</p>
Information about the mathematical methods	<p>Are the parameters to be estimated in the model specified?</p> <p>Are the uncertainties concerning these parameters evaluated?</p> <p>Are the different components of the model and their roles clearly defined?</p> <p>Are the different scenarios/processes/assumptions tested by the model clearly specified?</p> <p>Are the data and data sources specified?</p> <p>Are the states of the various model components clearly specified from data?</p> <p>Are the methods used to compute estimates specified?</p>

categories: correlation analysis, regression analysis, ANOVA/ANCOVA, bayesian analyses, statistical tests, other, not concerned).

Information about mathematical methods Items concerning mathematical methods are formulated using the guidelines for Strengthening The Reporting of Empirical Simulation Studies (STRESS) [13]. Several criteria identified from these guidelines allow the collection of information which is consistent with several types of models. The developed items are based on these criteria and can be categorized into four types: (i) description of the objectives of the model, (ii) description of the assumptions made by the model, (iii) description of the data used in the model, and (iv) description of the implementation of the model. These items offer a choice of four possible answers (yes, no, unclear, not concerned).

Data analysis

This review will be conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [22]. The data synthesis will report all applicable items from this checklist. This synthesis will present four parts concerning the general characteristics of the included studies and the three sections of the developed extraction grid. The first part, which deals with general information about the studies, provides a summary of the context, the scope of the studies, the assumptions made, and the presented results. The second part, concerning the study design, gives insights about the methodological quality of the considered studies. The third and fourth sections allow to identify and assess the implementation of quantitative methods. This synthesis will be presented using descriptive tables and graphs and interpreted accordingly. All analyses will be presented according to study design and regardless of the design when relevant.

Discussion

Intervention impact evaluation is one of many tools that support evidence-based program evaluation. It allows to determine whether an intervention improves significantly the situation of those who benefit from it compared to those who do not in the context of the study. Henceforth, this type of evaluation has been rooted in evaluation practices for the purpose of quantifying intervention effects and has legitimated a form of hierarchy among methods in terms of evidence [23]. Yet, these methods (including randomized controlled trials) are not without limitations, and while useful, their implementation requires better understanding of what kind of evidence they provide as well as how they may be used to inform research and practices [24–26]. In addition to the general criticisms, these methods have specific limitations regarding the field of health promotion and should be tailored before being used [27–29].

In the context of the fight against HIV epidemic, there is an urgent demand for evidence-based health promotion interventions to make available enough prevention strategies to stem the epidemic [30]. This dual demand for health promotion interventions and evidence warrants this review to map the methods commonly used to evaluate the impact of these interventions. This review will not only help to identify these methods but will also discuss the operationalization of these methods and the adaptations that have been made or could be made. Therefore, it is not intended to question the relevance of these studies which attempt to evaluate and make available useful interventions. On the contrary, it responds to the need for knowledge about existing methods and to consider how to use them, improve them, and even design complementary tools to better evaluate interventions, especially in the field of health promotion.

This protocol has limitations. The included studies are expected to quantify the intervention impact on HIV transmission. Such studies rely on confirmatory methods based on quantitative data, thus leading to exclude mixed studies whose aims do not focus on hypothesis or theory testing [31]. Nonetheless, as long as the eligibility criteria are met and the data extraction grid remains applicable, such study will be considered. Similarly, all studies that do not refer to a measure of HIV prevalence or incidence as an outcome are not considered. Thus, we will not be able to identify all health promotion studies used to stem the epidemic. However, this choice is warranted given that the review focuses on quantitative methods, while the selected studies are supposed to assess the extent to which these interventions curb the epidemic.

Despite its limitations, this review will be useful in informing practices of impact evaluation of HIV programs. The current recommendations for evaluating HIV

programs [32] acknowledge the relevance of numerous methods depending on the questions to be answered and the context of these programs. This review will help to identify the gap between these recommendations and what actually is being done, by identifying what methods are implemented and how they are implemented. In addition, the analysis of the implementation of the methods according to the study characteristics will allow to discuss in what circumstances each method applies. Thus, the results of this review will highlight some methodological challenges concerning the impact evaluation of complex interventions in order to guide future methodological developments. Given the need for evidence of the effectiveness of complex interventions, reviewing impact evaluation methods is necessary in order to map and improve these methods and consequently to improve decision-making pertaining to the concerned interventions.

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials; MMAT: Mixed Methods Appraisal Tool; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols; SAMPL: Statistical Analyses and Methods in the Published Literature; STRESS: Strengthening The Reporting of Empirical Simulation Studies; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TREND: Transparent Reporting of Evaluations with Nonrandomized Designs.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13643-022-01970-z>.

Additional file 1. PRISMA-P 2015 Checklist.

Additional file 2. List of references used to test data extraction grid.

Acknowledgements

We would like to thank Elisabeth Adjadj (DESP, Inserm, France) for her assistance in developing the search strategy.

Authors' contributions

AR designed the study and first drafted the protocol. VR, LKS, and YY contributed to the study design and critically revised the manuscript. All authors have read and approved the manuscript.

Funding

The author(s) received no specific funding for this work.

Availability of data and materials

All data relevant to the study are included in the article or uploaded as additional files.

Declarations

Competing interests

The authors declare that they have no competing interests.

Author details

¹Centre Population et Développement (IRD, Université Paris Cité, Inserm ERL 1244 SAGESUD), 75006 Paris, France. ²Institut de recherche pour le

développement, Marseille, France. ³French Collaborative Institute on Migrations, CNRS, Aubervilliers, France. ⁴UMR 1137, Inserm, University of Paris, IAME, Paris, France.

Received: 31 May 2021 Accepted: 27 April 2022
Published online: 06 May 2022

References

- World Health Organization (2007) Resolutions WHA 60.24 Health promotion in a globalized world.
- World Health Organization (2009) Resolutions WHA 62.14 Reducing health inequities through action on the social determinants of health.
- McQueen DV, Anderson LM. Utiliser des données probantes pour évaluer l'efficacité de la promotion de la santé: quelques enjeux fondamentaux. *Promot Educ.* 2004;11:1–6.
- O'Neill M. Le débat international sur l'efficacité de la promotion de la santé: d'où vient-il et pourquoi est-il si important? *Promot Educ.* 2004;11:6–10.
- Piot P, Bartos M, Larson H, Zewdie D, Mane P. Coming to terms with complexity: a call to action for HIV prevention. *Lancet.* 2008;372:845–59.
- Joint United Nations Programme on HIV/AIDS. Fast-tracking combination prevention. Geneva: UNAIDS; 2015.
- Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred Reporting Items For Systematic Review And Meta-Analysis Protocols (PRISMA-P) 2015 statement. *Syst Rev.* 2015;4:1.
- World Health Organization (1986) Ottawa Charter for Health Promotion.
- Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakaravitch C, Song F, et al. Evaluating non-randomised intervention studies. *Health Technol Assess Winch Engl.* 2003;7(iii-x):1–173.
- Bramer WM, Rethlefsen ML, Kleijnen J, Franco OH. Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. *Syst Rev.* 2017;6:245.
- Lang TA, Altman DG. Basic statistical reporting for articles published in Biomedical Journals: the "Statistical Analyses and Methods in the Published Literature" or the SAMPL guidelines. *Int J Nurs Stud.* 2015;52:5–9.
- Hong QN, Fàbregues S, Bartlett G, et al. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Educ Inf.* 2018;34:285–91.
- Monks T, Currie CSM, Onggo BS, Robinson S, Kunc M, Taylor SJE. Strengthening the reporting of empirical simulation studies: introducing the STRESS guidelines. *J Simul.* 2019;13:55–67.
- Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ.* 2008. <https://doi.org/10.1136/bmj.a2390>.
- Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ.* 2010. <https://doi.org/10.1136/bmj.c332>.
- Campbell MK, Piaggio G, Elbourne DR, Altman DG. Consort 2010 statement: extension to cluster randomised trials. *BMJ.* 2012. <https://doi.org/10.1136/bmj.e5661>.
- Montgomery P, Grant S, Mayo-Wilson E, et al. Reporting randomised trials of social and psychological interventions: the CONSORT-SPI 2018 extension. *Trials.* 2018;19:407.
- Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of non-randomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Public Health.* 2004;94:361–6.
- von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening of Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Int J Surg.* 2014;12:1495–9.
- Pace R, Pluye P, Bartlett G, Macaulay AC, Salsberg J, Jagosh J, et al. Testing the reliability and efficiency of the pilot Mixed Methods Appraisal Tool (MMAT) for systematic mixed studies review. *Int J Nurs Stud.* 2012;49:47–53.
- Souto RQ, Khanassov V, Hong QN, Bush PL, Vedel I, Pluye P. Systematic mixed studies reviews: updating results on the reliability and efficiency of the mixed methods appraisal tool. *Int J Nurs Stud.* 2015;52:500–1.
- Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Syst Rev.* 2021;10:89.
- Cartwright N, Hardie J. Evidence-ranking schemes, advice guides, and choosing effective policies. In: *Evid.-Based Policy Pract. Guide Doing It Better*: Oxford University Press; 2012. p. 135–43.
- Gertler PJ, Martinez S, Premand P, Rawlings LB, Vermeersch CMJ. *Impact Evaluation in Practice*. 2nd ed: World Bank Publications; 2016.
- Van Belle S, Wong G, Westhorp G, Pearson M, Emmel N, Manzano A, et al. Can "realist" randomised controlled trials be genuinely realist? *Trials.* 2016;17:313.
- Deaton A, Cartwright N. Understanding and misunderstanding randomized controlled trials. *Soc Sci Med.* 2018;210:2–21.
- McQueen DV, Anderson LM. What counts as evidence: issues and debates. In: *Eval. Health Promot. Princ. Perspect.*, WHO Regional Publications: WHO Regional Office Europe; 2001. p. 63–81.
- Potvin L, McQueen DV. Practical dilemmas for health promotion evaluation. In: Potvin L, McQueen DV, Hall M, de Salazar L, Anderson LM, Hartz ZMA, editors. *Health Promot Eval. Pract. Am. Values Res.* New York, NY: Springer; 2009. p. 25–45.
- de Salazar L, Hall M. Developing evaluation questions: beyond the technical issues. In: Potvin L, McQueen DV, Hall M, de Salazar L, Anderson LM, Hartz ZMA, editors. *Health Promot. Eval. Pract. Am. Values Res.* New York, NY: Springer; 2009. p. 49–62.
- Kurth AE, Celum C, Baeten JM, Vermund SH, Wasserheit JN. Combination HIV prevention: significance, challenges, and opportunities. *Curr HIV/AIDS Rep.* 2011;8:62–72.
- Johnson B, Christensen LB. Quantitative, qualitative, and mixed research. In: *Educ. Res. Quant. Qual. Mix. Approaches*. 7th ed. Los Angeles: SAGE; 2020. p. 29–54.
- Joint United Nations Programme on HIV/AIDS. Strategic guidance for evaluating HIV prevention programmes. Geneva: UNAIDS; 2010.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



Annexe 2

Article 2 : Revue systématique concernant les méthodes quantitatives pour évaluer l'impact des interventions dans la lutte contre le VIH



Quantitative Methods Used to Evaluate Impact of Combination HIV Prevention Intervention: A Methodological Systematic Review

Andrainolo Ravalihasy^{1,2,3} · Pearl Anne Ante-Testard^{4,5} · Lidia Kardas-Sloma³ · Yazdan Yazdanpanah⁶ · Manuela De Allegri⁷ · Valéry Ridde^{1,2}

Accepted: 18 January 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Combination HIV prevention aims to provide the right mix of biomedical, behavioral and structural interventions, and is considered the best approach to curb the HIV pandemic. The impact evaluation of combined HIV prevention intervention (CHPI) provides critical information for decision making. We conducted a systematic review of the literature to map the designs and methods used in these studies. We searched original articles indexed in Web of Science, Scopus and PubMed. Fifty-eight studies assessing the impact of CHPI on HIV transmission were included. Most of the studies took place in Asia or sub-Saharan Africa and were published from 2000 onward. We identified 36 (62.1%) quasi-experimental studies (posttest, pretest–posttest and nonequivalent group designs) and 22 (37.9%) experimental studies (randomized designs). The findings suggest that diverse methods are already rooted in CHPI impact evaluation practices as recommended but should be better reported. CHPI impact evaluation would benefit from more comprehensive approaches.

Keywords Combination HIV prevention intervention · Impact evaluation · Decision-making · Evidence-based

Introduction

In 2021, UNAIDS [1] emphasized that the progress in the fight against the HIV/AIDS pandemic is slowing down and is even in jeopardy owing to the effects of COVID-19

crisis on health systems. This statement implies a call for action and the continuation of efforts to curb the pandemic while recognizing that HIV prevention is still a major public health issue. Despite the progress in biomedical prevention research and tools, it is already acknowledged that biomedical approaches alone are not sufficient to curb the epidemic [2]. Besides, numerous behavioral and structural interventions have shown to be effective in improving intermediate outcomes that potentially block the pathway to HIV transmission, such as inducing changes in sexual behaviors. Although such interventions struggle to show any impact on HIV incidence [3, 4], dealing with structural and behavioral determinants of HIV spread alongside the use of biomedical prevention tools is necessary when designing HIV prevention interventions [5, 6]. Therefore, combination HIV prevention, “a dynamic, rights-based approach to providing the right mix of biomedical, behavioral and structural interventions aiming to have the greatest, sustained effort on reducing new HIV infections” [7], is considered the best approach to curb the HIV pandemic [8]. A combination HIV prevention intervention (CHPI) then relates to any intervention that aims to reduce HIV transmission by using strategies that deal with behavioral and structural health determinants supplemented by biomedical prevention tools [5]. Whereas

✉ Andrainolo Ravalihasy
andrainolo.ravalihasy@ird.fr

¹ Centre Population et Développement (Ceped), Institut de Recherche pour le Développement (IRD) et Université Paris Cité, Inserm ERL 1244, 45 Rue Des Saints-Pères, 75006 Paris, France

² French Collaborative Institute On Migrations, CNRS, Aubervilliers, France

³ UMR 1137, Inserm, Université Paris Cité, IAME, Paris, France

⁴ MESuRS and Unité PACRI, Institut Pasteur, Conservatoire National des Arts et Métiers, Paris, France

⁵ Francis I. Proctor Foundation, University of California, San Francisco, San Francisco, CA, USA

⁶ AP-HP, Hôpital Bichat, Service, 75018 Paris, France

⁷ Heidelberg Institute of Global Health, University Hospital and Faculty of Medicine, University of Heidelberg, Heidelberg, Germany

biomedical and behavioral components are individually focused approaches, structural components are designed to affect environmental conditions outside the individuals' control (economic conditions, policies, programmatic vulnerabilities, social inequalities, discrimination, societal norms) [9, 10]. CHPIs add up many components and mobilize all involved parties to account for specific risks and vulnerabilities. By doing so, they take into consideration the contextual needs and conditions of people and communities. CHPIs are expected to prevent HIV transmission by considering their components' effectiveness and relevant hypotheses on how they interact with one other, a credible program impact pathway and a program theory that is able to deal with pragmatic issues [11]. Hence, the impact evaluation of CHPIs raises methodological challenges owing to their multicomponent and complex nature.

Impact evaluations contribute to HIV-related decision making by generating evidence from CHPIs about effective strategies to prevent HIV infections. Indeed, impact evaluations are primarily expected to quantify the extent to which the intervention to be evaluated achieved the intended outcomes. In that sense, they are supposed to establish a causal relationship between the set of activities undertaken during the considered intervention and the improvements of the beneficiaries' circumstances. Beyond the question of whether an intervention is effective, impact evaluations are expected to provide comprehensive evidence that informs the decisions on the implementation, the scale-up, as well as on the continuation or the interruption of an intervention [11, 12]. Currently, impact evaluations are often based on quantitative methods applied in the framework of a Campbellian validity model where the impacts of intervention are quantified in controlled settings (efficacy) and then in real-world settings (effectiveness) [13, 14]. These methods have proven their relevance, which has rooted their use in evaluation practices and has legitimated a form of hierarchy among methods in terms of evidence [15]. However, they present their own limits [14, 16, 17], inter alia concerning CHPIs impact evaluation [11]. The current recommendations for CHPIs impact evaluation acknowledge the relevance of diverse quantitative methods and approaches depending on the context of these interventions [11]. For these reasons, this study is conducted to map and critically review the quantitative methods used to assess the efficacy or the effectiveness of CHPIs on HIV transmission. It will help to address the gap between the recommendations about the use of these methods and what is actually occurring.

Methods

The systematic review was conducted according to the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) statement [18] (Online Appendix 1). The protocol was registered and published with PROSPERO (CRD42020210825) and details are presented in Ravalihasy et al. [19].

Search Strategy

We searched original articles in English and French indexed in Web of Science, Scopus and PubMed from inception to August 2022. The joint use of these databases is expected to uncover most of the studies relevant to this review [20]. Database-specific search strategies were developed using key terms associated with (i) HIV transmission (prevalence or incidence), (ii) impact of interventions and (iii) prevention and sexual risk exposure [19]. We added terms related to HIV transmission and target populations in order to improve the search strategy (Online Appendix 2). We conducted the literature search in English and then in French after translating text terms (i.e. not index terms). All records were retrieved on August 25, 2022, and imported to a reference manager library (Zotero).

Eligibility

A first screening step was conducted based on the title and abstract. References were deemed relevant if:

- The study focused on the impact evaluation of CHPI.
- The analyses were conducted using data gathered from the intervention beneficiaries during these interventions.
- The data allowed to assess the impact of the intervention including the behavioral/structural components.
- HIV incidence, prevalence or averted infections was an intervention outcome.

A second screening step based on full text was then conducted and allowed to exclude all studies that:

- did not assess an impact on HIV transmission (hereafter, irrelevant outcome).
- referred to an intervention that did not include any behavioral or structural component (hereafter, irrelevant intervention).
- where the data did not allow to assess the impact of the behavioral or structural components (hereafter, irrelevant data).

- are not an intervention impact evaluation (hereafter, irrelevant evaluation study).
- were based on simulated data.
- did not report any quantitative method (statistical methods, mathematical modeling or both) to assess the intervention impact on HIV transmission.

In each step, the studies were screened independently by AR and PAA-T. The disagreements were resolved by LKS, MDA or VR.

Data Extraction

A systematic review management software, Covidence (www.covidence.org), was used for data management and extraction. General information (authors, title, date of publication, location where the studies were carried out, purpose and results of the studies) were extracted from the included studies. Specific information about quantitative methods were extracted using a grid developed for this purpose [19]. When necessary, other documents referenced in the full texts, such as the study protocol or pilot study, were used as complementary supports during the data extraction process. Also, each study was classified as an “efficacy” or “effectiveness” study according to the study objectives as stated in the full texts, or according to the fact that the CHPI was already implemented or scaled-up, or using items from the PRECIS-2 tool [21].

Data Analysis

Evaluation Design

Distinctions are made between experimental and quasi-experimental approaches for evaluation design. An experimental design relates to studies where researchers assign the participants to different intervention conditions according to a randomization scheme. A quasi-experimental design relates to studies where researchers do not control intervention allocation or do not use a randomization scheme for intervention allocation [22]. Quasi-experimental designs include posttest design, pretest–posttest design and its extensions (such as interrupted time-series or regression discontinuity), nonequivalent group design, and any combination of the former [23]. Items from the Mixed Methods Appraisal Tool [24, 25] were adapted to extract information about study design reporting (Online Appendix 3).

Statistical Methods

Information about statistical methods were assessed using items developed from the guidelines for Statistical Analyses and Methods in the Published Literature [26]. Information

about sample size are assessed through two items to verify if the data allowed the detection of the expected gain from the CHPIs (expected effect size) with sufficient precision in estimates. The three remaining items allow to check if the studies reported: (i) how the methods fit to the data structure (statistical validity condition), (ii) a measure of precision (confidence and credible intervals) alongside the impact measure, and (iii) how the analysis accounted for the evaluation design [19]. We verified if the studies reported each item and all items together.

Mathematical Modeling

Information about mathematical modeling were extracted using items developed from the guidelines for Strengthening The Reporting of Empirical Simulation Studies [27]. Two items allow the investigation of information about model outputs and their precision. Two items allow the investigation of information about the models’ assumptions. Two items allow the investigation of information about the data used for modeling. One item allows the investigation of information about models implementation [19]. We verified if the studies reported each item and all items together.

Results

The literature search identified 2335 articles, of which 154 were considered relevant based on their titles and abstracts. A total of 58 articles [28–85] concerning 46 CHPI satisfied the inclusion criteria and were included in our review (Fig. 1).

All the 46 considered CHPIs (58 articles) presented behavioral components, of which eighteen (39.1%) presented structural components. Table 1 reports the characteristics of the included studies. Among 53 studies where HIV transmission was a primary outcome, 26 (49.1%) reported a significant reduction of HIV transmission. Among five studies where HIV transmission was a secondary outcome, two (40.0%) reported a significant reduction of HIV transmission. Thirty-six studies (62.1%) including three cost-effectiveness studies were conducted in real-life settings, of which 21 (58.3%) reported a reduction of HIV transmission as intended. Twenty-two studies (37.9%) were conducted in controlled settings, of which seven (31.8%) reported a reduction of HIV transmission as intended. Most of the included studies targeted female sex workers (39.7%), were published after 2000 (94.8%), and were conducted in Asia (36.2%) or in sub-Saharan Africa (44.8%). Data concerning the study characteristics are provided in Online Appendix 4.

Table 2 presents the evaluation design characteristics of the included studies. Among the included studies, 22 (37.9%) used experimental design: 12 (54.6%) were

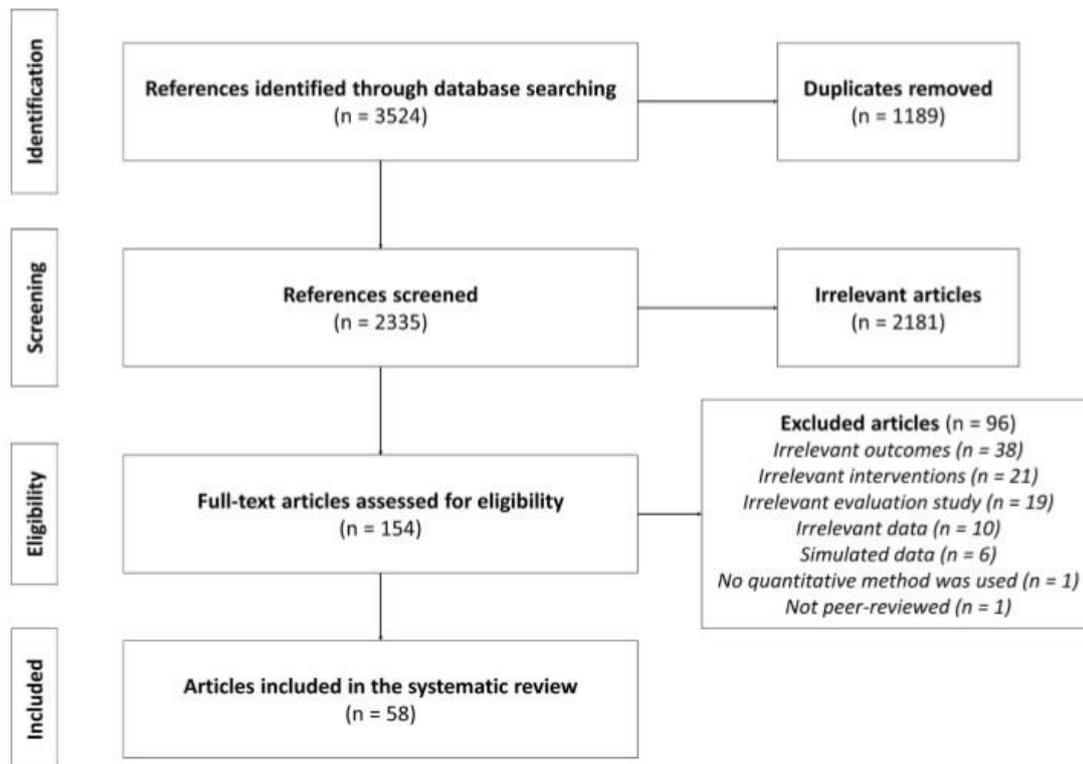


Fig. 1 PRISMA flow diagram for study selection process

cluster-randomized, 9 (40.9%) were individually randomized, and one (4.5%) was a multilevel randomization (i.e., a combination of individual and cluster-randomization). The remaining 36 (62.1%) studies used quasi-experimental designs, including 6 (16.7%) posttest designs, 6 (16.7%) nonequivalent group designs, 8 (22.2%) pretest–posttest designs and 16 (44.4%) combinations of pretest–posttest and nonequivalent group designs. Items concerning sampling (67.2%) and treatment allocation or exposure (82.8%) were more frequently reported than other items. Quasi-experimental studies more frequently reported how confounders and measurement biases were accounted for in the design than in the experimental studies (50.0% vs 36.4%, respectively), but this difference was not significant ($\chi^2 = 1.02$, $p = 0.311$). In particular, among the experimental studies, eight reported blinded procedures: all kinds of outcome assessor [44, 49, 71]; biological outcome assessor [35, 48]; and interviewers [72] or investigators [40, 55] were blinded in these studies. Experimental studies more frequently reported information about intervention administration and adherence than did quasi-experimental studies (63.6% vs 13.9%, respectively, $\chi^2 = 15.3$, $p < 0.001$).

Among the included studies, 51 (87.9%) used statistical methods to assess to what extent the interventions reduced HIV transmission as intended. Among the former,

38 (74.5%) used regression-based methods, twelve (23.5%) used hypothesis testing and one (2.0%) used an analysis of variance. Table 3 presents the statistical methods used by the included studies. The measures of precision and the consideration of evaluation design were the most frequently reported items (80.4%). The latter involved any methods to account for the data generating process [86, 87] (adjustment, stratification, matching, weighting). Confidence and credible intervals were reported in 41 studies (80.4%), while p-values only were reported in five studies (9.8%). Twenty out of twenty-four studies used an expected effect size related to HIV transmission to compute sample size: 15 were experimental and five were quasi-experimental studies. Overall and for each considered item, statistical methods were more frequently reported in experimental studies.

Among the included studies, nine quasi-experimental studies (15.5%) used mathematical models to assess to what extent the interventions reduced HIV transmission as intended. Table 4 presents the reported mathematical models characteristics in the included studies. Information about model implementation were the least reported.

Table 1 Characteristics of the 58 included studies

Study characteristics	n	%
Impact on HIV transmission as primary outcome of the study		
No	5	8.6
Yes	53	91.4
Significant reduction of HIV transmission reported		
No	30	51.7
Yes	28	48.3
Study type (in real life settings)		
Efficacy (in controlled settings)	22	37.9
Effectiveness (in real-world settings)	36	62.1
Study population		
Female sex workers	23	39.7
Drug users	6	10.4
*Female sex workers and drug users	1	1.7
Men who have sex with men	5	8.6
Adolescent and young adults	6	10.4
Businesses employees	1	1.7
Couples	3	5.2
General population	13	22.4
Year of publication		
< 2000	3	5.2
2000–2009	19	32.7
2010 onward	36	62.1
Study location		
Asia	21	36.2
Caribbean	1	1.7
Eastern Europe	1	1.7
North America	4	6.9
South/Latin America	3	5.2
Sub-Saharan Africa	26	44.8
**Multi-country	2	3.5

*The intervention was directed toward female sex workers and drug users

**One study was conducted in South Africa; Tanzania; Zimbabwe and Thailand; one study was conducted in USA and in Thailand

Discussion

This study contributes to the literature on CHPI impact evaluation by giving a broad view of the quantitative methods typically being used and how they are reported. Diverse quantitative designs and methods are currently being implemented, depicting the intrinsic complexity and the contexts of these interventions. To derive an intervention impact estimate on HIV transmission, one can use common procedures, opt for specific procedures [88–91] as can be seen in four included studies [36, 38, 44, 50], or develop procedures when relevant as can be seen in one CHPI [92]. Moreover, CHPIs potentially include structural level activities, raising the demand for more comprehensive impact evaluation studies [11]. In our review, two studies [39, 66] used causal pathway analysis (i.e. accounted for the hypothesized relation between intervention components) to derive estimates of CHPI impact. Some other studies adjusted their estimates for intervention implementation outcomes such as coverage and acceptance, or reported such information [58, 67, 75, 93]. In light of the above, diverse methods are already rooted in CHPI impact evaluation practices as recommended [11]. Still, some effort should be made to better report the methods related to the results of these studies in order to best inform theories and practices.

This review identified more quasi-experimental studies than in other reviews on behavioral and structural interventions to prevent HIV infection [94–99]. It is consistent with the fact that such design may be more appropriate to assess the impact on HIV incidence or prevalence [11, 100, 101]. Furthermore, some ethical, political and resources issues as well as the nature of CHPIs make the implementation of randomized designs less feasible [102]. When feasible, some features adaptation pertaining to randomization or blinding procedures often apply, as reported in our review. Indeed, the complex nature of CHPIs challenges the translation into practice of the theoretical properties of randomized designs,

Table 2 Proportion of reported design characteristics in the 58 included studies

Items [19, 25]	Quasi-experimental (N = 36)		Experimental (N = 22)		Overall (N = 58)		χ^2 †	p-value††
	n	%	N	%	n	%		
Sampling	20	55.6	19	86.4	39	67.2	5.88	0.015
Treatment allocation	29	80.6	19	86.4	48	82.8	0.32	0.570
Outcome integrity	5	13.9	3	13.6	8	13.8	0.001	0.978
Consideration of potential confounders and measurement biases	18	50.0	8	36.4	26	44.8	1.02	0.311
Administration/adherence	5	13.9	14	63.6	19	32.8	15.34	<0.001
Reported of all items listed above	1	2.9	0	0.0	1	1.7	0.62	0.430

†Chi-square statistic

††p-value associated with the chi-square statistic

Table 3 Proportion of reported items among 51 studies concerned by statistical methods

Items [19, 26]	Quasi-experi- mental (N=29)		Experimental (N=22)		Overall (N=51)		χ^2 †	p-value†)
	n	%	n	%	n	%		
Expected difference between groups for sample size calculation	8	27.6	16	72.7	24	47.1	10.23	0.001
Estimate level of precision for sample size calculation	9	31.0	16	72.7	25	49.0	8.70	0.003
Statistical validity condition	3	10.3	4	18.2	7	13.7	0.65	0.421
Reported intervention impacts measures of precision	23	79.3	18	81.8	41	80.4	0.05	0.823
Consideration of evaluation design	19	65.5	22	100.0	41	80.4	9.43	0.002
Reported of all items listed above	0	0.0	3	13.6	3	5.9	4.20	0.04

†Chi-square statistic

††p-value associated with the chi-square statistic

Table 4 Proportion of accurately reported items among 9 studies concerned by mathematical models

Items [19, 27]	n	%
Model outputs	8	88.9
Sensitivity analysis	8	88.9
Model components	7	77.8
Model hypothesis	9	100.0
Data sources	9	100.0
Model inputs	7	77.8
Model implementation	2	22.2
Reported of all items listed above	2	22.2

which shall ensure unbiasedness and precision [103, 104]. These results illustrate the recommendation that no single methodology should be applied as a gold standard to evaluate CHPIs [11], particularly since confounders may impact experimental as well as quasi-experimental designs [105, 106].

Our review contributes to the literature and stands out for giving insights about the methods used to assess the effectiveness of these interventions. Numerous methods are used to assess the impact of CHPIs. This diversity allows to account for evaluation contexts, and practically all types of evaluation designs. Our results shows that although data from one-group posttest designs seem irrelevant for impact evaluation [107], mathematical modeling allows a counterfactual analysis of intervention outcomes. This review highlights that the reported sampling strategy or the data used in most of the quasi-experimental studies and in a few experimental studies was not intended for impact evaluation on HIV transmission. In addition, information about how well the statistical methods in use suit the data structure are not frequently reported. Therefore, the reporting of these methods should be improved in order to make clearer the relevance of the sample in relation to the methods used to derive impact estimates. Indeed, this will help to better

understand the significance of the results given the diversity of the designs and the methods that are actually used.

Implications for Impact Evaluation Studies

This systematic review shows that the availability of diverse approaches, methods and designs allows us to challenge the complexity of CHPIs impact evaluation. The use of pathways analysis may overcome the need for more comprehensive approaches to impact evaluation. In order to go further, the outcomes of impact evaluation should incorporate contextual and implementation outcomes [108]. In that sense, different tools [109, 110] or approaches, such as theory-driven outcomes evaluation [111, 112], may help to enhance impact evaluation designs. By doing so, impact evaluation may improve the generalizability or the transferability of the findings. Nevertheless, while many approaches are already rooted in impact evaluation practices and the future direction for improvement is identified, poor study reporting may hamper the credibility of the findings. The lack of information on design and quantitative methods implementation means we cannot firmly rely on the findings. Many reporting guidelines have been developed to enhance the reporting of health research studies [113–115], and while their use might be cumbersome, their uptake is critical [116].

The lack of reporting observed in this review may be related to major methodological issues [117] given the items that are reported in the data extraction grid. Our results question the sufficiency and the completeness of the procedures of impact evaluation in the included studies, especially concerning the extent to which the data in use and the intervention administration are relevant. Thus, this review points to the need for updating some key principles that guide the planning, the processing, and the reporting of impact evaluation studies in order to take advantage of the strengths and weaknesses of the designs and the methods in use. In light of the above, these principles should deal with three non-exhaustive but essential questions.

First of all, these principles should address the question of the primary recipients of the evaluation findings (e.g. beneficiaries, stakeholders, funders). Although impact evaluation studies share the aim of establishing causal relationship between programs and outcomes, they may have different purposes ranging from testing the relevance of a program within a specific setting to influencing political decisions [12]. Indeed, the evaluation strategy and constraints may differ according to whoever is interested and involved in the evaluation process. It should be clear whether the evaluation studies are intended to only apply and be restituted within the initial program context or to have implications beyond. Indeed, the clarification of this point gives concrete indications on the scope of the evaluation outcomes.

Second, impact evaluation studies should be able to provide information about and account for the data generating process. Here, accounting for data generating process means identifying to what extent the data allows to derive impact estimates and if not, what kind of adjustment are needed. The data generating process constrains certain methodological aspects of the impact evaluation by shaping the data and the sample characteristics, the intervention allocation or exposure, and the confounding and/or the intervention contextual factors. For example, the intervention allocation may constrain the design or the quantitative analysis methods depending on whether the data collection was specifically planned to allow an assessment on the outcome of interest, such as HIV incidence. Moreover, some additional quantitative processes such as power analysis should be performed when the data collection was not planned specifically for impact evaluation purpose.

Third, some implementation outcomes, especially fidelity, should be assessed alongside impacts. Implementation fidelity is a multidimensional concept [108, 118] that encompasses not only the quality of the delivery or the adherence to the intervention, but also the exposure to the intervention, the beneficiaries' responsiveness and the program components differentiation. This outcome deals with theoretical issues such as the program theory and pathways as well as practical issues such as the stakeholders and beneficiaries' participation. Hence, such an outcome constitutes a key intermediate factor for attaining the expected effects from the intervention.

Taking into account these three questions allows to move towards a more comprehensive manner to consider the intervention impact, capturing the efficacy-effectiveness continuum [21]. These questions also allow to introduce the notion of transferability which focuses on a more practical way to consider the generalizability of the findings [110, 119] without questioning the necessity of the probability statements on which the campbellian generalizability relates to. Indeed, the former one deals with the impact variation that depends on the beneficiaries, the context and the implementation,

while the latter one warrants the relevance of the impact estimates. By addressing these three questions, the evaluation process takes advantage of the intervention theories of action and change [120] and will enable the production of more actionable results.

Strengths and Limitations

To our knowledge, this review is the first to focus on the impact evaluation methodology of CHPI. Hence, our review included different intervention designs and settings, thus limiting the possibility of a sharper methodological analysis of impact evaluation. However, while the grid we developed [19] only include information that are common to each design, it allows to adopt a conservative approach to the analysis by focusing on the items that are reported correctly.

Some references may have been missed by the search equation: we identified few studies published before 2000 and found no study conducted in Western Europe. Nevertheless, we expanded the initial literature search strategy in order to include index terms and text terms that are expected to increase the likelihood of detecting eligible studies. Furthermore the focus on CHPI and HIV transmission, the diversity of included study and the fact that almost half of the included studies did not show a significant impact on HIV transmission is comforting with respect to publication bias [121].

This review also included efficacy as well as effectiveness studies that may have different purposes. Nevertheless, decision makers often equally use evidence from these studies [11, 14]. Also, the specific context of CHPIs may affect the generalizability of the findings in the same way [122].

Modeling studies may account for complementary approaches to impact evaluation [12]. However, these studies were included thanks to their common use in measuring HIV incidence or prevalence and to examine potential intervention impact [123]. Therefore, our study highlights the relevance of mathematical models as tools for CHPI impact evaluation.

Conclusion

This study highlights that diverse methods are already rooted in CHPI impact evaluation practices. Still, some effort have yet to be made to accurately report these methods to allow a better understanding of the findings' significance. In addition, CHPI impact evaluation may benefit from more comprehensive approaches such as path analysis or theory-driven evaluation. Such approaches allow the quantification of the impact of these interventions, while also taking into account the pragmatic issues and causal theories underlying these

interventions. Indeed, the success of a CHPI is supposed to rely on the interaction between the intervention components implementation, and so should the impact evaluation. These findings contribute to inform future directions for impact evaluation practices in order to make available more transferable and generalizable insights into CHPI.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10461-023-04000-8>.

Acknowledgements We thank Elisabeth Adjadj (DESP, Inserm, France) for her assistance in developing the search strategy. We also thank Emilie Brunet and Laurence Goury (IST, IRD, France) for their assistance in finding full texts of articles. We thank Kevin Jean (MESuRS, Cnam, France) for his precious advice and help during the study selection process.

Author Contributions AR designed the study and VR, LKS, YY and MDA contributed to the study design. AR, PAA-T, MDA and LKS performed the literature search. AR and PAA-T extracted the data and conducted the analyses. AR first drafted the manuscript. PAA-T, MDA, VR, LKS and YY critically revised the manuscript. All authors have read and approved the manuscript.

Funding This work was done in the course of the employment of all authors.

Data Availability All data relevant to the study are included in the article or uploaded as supplementary information.

Code Availability Not applicable.

Declarations

Conflict of interest We declare no conflicts of interest.

Ethical Approval Not applicable.

Consent to Participate Not applicable.

Consent for Publication Not applicable.

References

1. Joint United Nations Programme on HIV/AIDS. Unequal, unprepared, under threat: why bold action against inequalities is needed to end AIDS, stop COVID-19 and prepare for future pandemics. Geneva: Joint United Nations Programme on HIV/AIDS; 2021.
2. Padian NS, Buvé A, Balkus J, Serwadda D, Cates W. Biomedical interventions to prevent HIV infection: evidence, challenges, and way forward. *Lancet*. 2008;372:585–99.
3. Gliberman J, Mitra S, Gogolishvili D, Rueda S, Schoffel L, Gangbar K, Shi Q, Rourke SB. HIV/STI prevention interventions: a systematic review and meta-analysis. *Open Med*. 2017;12:450–67.
4. Iskarpatyoti BS, Lebov J, Hart L, Thomas J, Mandal M. Evaluations of structural interventions for HIV prevention: a review of approaches and methods. *AIDS Behav*. 2018;22:1253–64.
5. Kurth AE, Celum C, Baeten JM, Vermund SH, Wasserheit JN. Combination HIV prevention: significance, challenges, and opportunities. *Curr HIV/AIDS Rep*. 2011;8:62–72.
6. Piot P, Bartos M, Larson H, Zewdie D, Mane P. Coming to terms with complexity: a call to action for HIV prevention. *Lancet*. 2008;372:845–59.
7. Joint United Nations Programme on HIV/AIDS. Combination HIV prevention: tailoring and coordinating biomedical. Geneva: Behavioural and Structural Strategies to Reduce New HIV Infections; 2007.
8. Joint United Nations Programme on HIV/AIDS. Fast-tracking combination prevention. Geneva: UNAIDS; 2015.
9. Sumartojo E. Structural factors in HIV prevention: concepts, examples, and implications for research. *AIDS*. 2000;14:S3.
10. Auerbach JD, Parkhurst JO, Cáceres CF. Addressing social drivers of HIV/AIDS for the long-term response: conceptual and methodological considerations. *Glob Public Health*. 2011;6:S293–309.
11. Joint United Nations Programme on HIV/AIDS. Strategic guidance for evaluating HIV prevention programmes. Geneva: UNAIDS; 2010.
12. Gertler PJ, Martinez S, Premand P, Rawlings LB, Vermeersch CMJ. *Why evaluate? Impact Eval. Pract. Second Ed.*; 2016
13. Campbell DT, Stanley JC. *Experimental and quasi-experimental designs for research*. Belmont: Wadsworth; 2011.
14. Chen HT. The bottom-up approach to integrative validity: a new perspective for program evaluation. *Eval Program Plann*. 2010;33:205–14.
15. Cartwright N, Hardie J. *Evidence-ranking schemes, advice guides, and choosing effective policies*. Oxford: Oxford University Press; 2012. p. 135–43.
16. Cartwright N, Hardie J. *What are RCTs good for?* Oxford: Oxford University Press; 2012. p. 122–34.
17. Deaton A, Cartwright N. Understanding and misunderstanding randomized controlled trials. *Soc Sci Med*. 2018;210:2–21.
18. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Syst Rev*. 2021;10:89.
19. Ravalihasy A, Kardaś-Słoma L, Yazdanpanah Y, Ridde V. Quantitative methods used to evaluate impact of health promotion interventions to prevent HIV infections: a methodological systematic review protocol. *Syst Rev*. 2022;11:87.
20. Bramer WM, Rethlefsen ML, Kleijnen J, Franco OH. Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. *Syst Rev*. 2017;6:245.
21. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ*. 2015;350: h2147.
22. de Vocht F, Katikireddi SV, McQuire C, Tilling K, Hickman M, Craig P. Conceptualising natural and quasi experiments in public health. *BMC Med Res Methodol*. 2021;21:32.
23. Reichardt CS. *Quasi-experimentation: a guide to design and analysis*. London: The Guilford Press; 2019.
24. Hong QN, Fàbregues S, Bartlett G, et al. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Educ Inf*. 2018;34:285–91.
25. Hong QN, Pluye P, Fàbregues S, et al. *Mixed Methods Appraisal Tool (MMAT) version 2018, User guide*; 2018
26. Lang TA, Altman DG. Basic statistical reporting for articles published in Biomedical Journals: The “Statistical Analyses and Methods in the Published Literature” or the SAMPL Guidelines. *Int J Nurs Stud*. 2015;52:5–9.
27. Monks T, Currie CSM, Onggo BS, Robinson S, Kunc M, Taylor SJE. Strengthening the reporting of empirical simulation studies: introducing the STRESS guidelines. *J Simul*. 2019;13:55–67.

28. Alary M, Banandur P, Rajaram SP, Thamattoor UK, Mainkar MK, Paranjape R, Adhikary R, Duchesne T, Isac S, Moses S. Increased HIV prevention program coverage and decline in HIV prevalence among female sex workers in South India. *Sex Transm Dis.* 2014;41:380–7.
29. Baeten JM, Heffron R, Kidoguchi L, et al. Integrated delivery of antiretroviral treatment and pre-exposure prophylaxis to HIV-1-serodiscordant couples: a prospective implementation study in Kenya and Uganda. *PLoS Med.* 2016;13: e1002099.
30. Banandur P, Mahajan U, Potty RS, Isac S, Duchesne T, Abdous B, Ramesh BM, Moses S, Alary M. Population-level impact of Avahan in Karnataka State, South India using multilevel statistical modelling techniques. *JAIDS-J Acquir Immune Defic Syndr.* 2013;62:239–45.
31. Bhave G, Lindan CP, Hudes ES, Desai S, Wagle U, Tripathi SP, Mandel JS. Impact of an intervention on HIV, sexually transmitted diseases, and condom use among sex workers in Bombay, India. *AIDS.* 1995;9(Suppl 1):S21–30.
32. Birdthistle I, Kwaro D, Shahmanesh M, et al. Evaluating the impact of DREAMS on HIV incidence among adolescent girls and young women: a population-based cohort study in Kenya and South Africa. *PLoS Med.* 2021;18: e1003837.
33. Boily M-C, Pickles M, Lowndes CM, et al. Positive impact of a large-scale HIV prevention programme among female sex workers and clients in South India. *AIDS.* 2013;27:1449–60.
34. Chabata ST, Hensen B, Chiyaka T, et al. The impact of the DREAMS partnership on HIV incidence among young women who sell sex in two Zimbabwean cities: results of a non-randomised study. *BMJ Glob Health.* 2021;6: e003892.
35. Coates TJ, Kulich M, Celentano DD, et al. Effect of community-based voluntary counselling and testing on HIV incidence and social and behavioural outcomes (NIMH Project Accept; HPTN 043): a cluster-randomised trial. *Lancet Glob Health.* 2014;2:E267–77.
36. Corbett EL, Makamure B, Cheung YB, Dauya E, Matambo R, Bandason T, Munyati SS, Mason PR, Butterworth AE, Hayes RJ. HIV incidence during a cluster-randomized trial of two strategies providing voluntary counselling and testing at the workplace, Zimbabwe. *AIDS.* 2007;21:483–9.
37. Cowan FM, Pascoe SJ, Langhaug LF, et al. The Regai Dzive Shiri project: results of a randomized trial of an HIV prevention intervention for youth. *AIDS.* 2010;24:2541–52.
38. Doyle AM, Ross DA, Maganja K, et al. Long-term biological and behavioural impact of an adolescent sexual health intervention in Tanzania: follow-up survey of the community-based MEMA kwa Vijana Trial. *PLOS Med.* 2010. <https://doi.org/10.1371/journal.pmed.1000287>.
39. Eaton LA, Kalichman SC, Kenny DA, Harel O. A reanalysis of a behavioral intervention to prevent incident HIV infections: including indirect effects in modeling outcomes of Project EXPLORE. *AIDS Care.* 2013;25:805–11.
40. El-Bassel N, Gilbert L, Terlikbayeva A, et al. Effects of a couple-based intervention to reduce risks for HIV, HCV, and STIs among drug-involved heterosexual couples in Kazakhstan: a randomized controlled trial. *JAIDS-J Acquir Immune Defic Syndr.* 2014;67:196–203.
41. El-Bassel N, McCrimmon T, Mergenova G, et al. A cluster-randomized controlled trial of a combination HIV risk reduction and microfinance intervention for female sex workers who use drugs in Kazakhstan. *J Int AIDS Soc.* 2021;24: e25682.
42. Geidelberg L, Mitchell KM, Alary M, et al. Mathematical model impact analysis of a real-life pre-exposure prophylaxis and treatment-as-prevention study among female sex workers in Cotonou, Benin. *JAIDS J Acquir Immune Defic Syndr.* 2021;86: e28.
43. Ghys PD, Diallo MO, Ettiègne-Traoré V, et al. Effect of interventions to control sexually transmitted disease on the incidence of HIV infection in female sex workers. *AIDS.* 2001;15:1421–31.
44. Go VF, Frangakis C, Minh NL, et al. Efficacy of a multi-level intervention to reduce injecting and sexual risk behaviors among HIV-infected people who inject drugs in Vietnam: a four-arm randomized controlled trial. *PLoS ONE.* 2015. <https://doi.org/10.1371/journal.pone.0125909>.
45. Goswami P, Medhi GK, Armstrong G, et al. An assessment of an HIV prevention intervention among People Who Inject Drugs in the states of Manipur and Nagaland, India. *Int J Drug Policy.* 2014;25:853–64.
46. Grabowski MK, Serwadda DM, Gray RH, et al. HIV prevention efforts and incidence of HIV in Uganda. *N Engl J Med.* 2017;377:2154–66.
47. Gregson S, Adamson S, Papaya S, Mundingu J, Nyamukapa CA, Mason PR, Garnett GP, Chandiwana SK, Foster G, Anderson RM. Impact and process evaluation of integrated community and clinic-based HIV-1 control: a cluster-randomised trial in eastern Zimbabwe. *PLoS Med.* 2007;4:545–55.
48. Hayes RJ, Donnell D, Floyd S, et al. Effect of universal testing and treatment on HIV incidence—HPTN 071 (PopART). *N Engl J Med.* 2019;381:207–18.
49. Hoffman IF, Latkin CA, Kukhareva PV, et al. A peer-educator network HIV prevention intervention among injection drug users: results of a randomized controlled trial in St. Petersburg, Russia. *AIDS Behav.* 2013;17:2510–20.
50. Jewkes R, Nduna M, Levin J, Jama N, Dunkle K, Puren A, Duvvury N. Impact of stepping stones on incidence of HIV and HSV-2 and sexual behaviour in rural South Africa: cluster randomised controlled trial. *BMJ.* 2008;337: a506.
51. Kagaayi J, Chang LW, Ssempijja V, et al. Impact of combination HIV interventions on HIV incidence in hyperendemic fishing communities in Uganda: a prospective cohort study. *Lancet HIV.* 2019;6:e680–7.
52. Kerrigan D, Mbwambo J, Likindikoki S, et al. Project Shikamana: community empowerment-based combination HIV prevention significantly impacts HIV incidence and care continuum outcomes among female sex workers in Iringa, Tanzania. *JAIDS-J Acquir Immune Defic Syndr.* 2019;82:141–8.
53. Koblin BA. Effects of a behavioural intervention to reduce acquisition of HIV infection among men who have sex with men: the EXPLORE randomised controlled study. *Lancet.* 2004;364:41–50.
54. Konate I, Traore I, Ouedraogo A, et al. Linking HIV prevention and care for community interventions among high-risk women in Burkina Faso-The ARNS 1222 “Yerelon” cohort. *JAIDS-J Acquir Immune Defic Syndr.* 2011;57:S50–4.
55. Latkin CA, Donnell D, Metzger D, et al. The efficacy of a network intervention to reduce HIV risk behaviors among drug users and risk partners in Chiang Mai, Thailand and Philadelphia, USA. *Soc Sci Med.* 2009;68:740–8.
56. Lee MB, Leibowitz A, Rotheram-Borus MJ. Cost-effectiveness of a behavioral intervention for seropositive youth. *AIDS Educ Prev.* 2005;17:105–18.
57. Luchters S, Chersich MF, Rinyiru A, Barasa M-S, King’ola N, Mandaliya K, Bosire W, Wambugu S, Mwarogo P, Temmerman M. Impact of five years of peer-mediated interventions on sexual behavior and sexually transmitted infections among female sex workers in Mombasa, Kenya. *BMC Public Health.* 2008;1:11. <https://doi.org/10.1186/1471-2458-8-143>.
58. Matovu J, Gray R, Makumbi F, Wawer M, Serwadda D, Kigozi G, Sewankambo N, Nalugoda F. Voluntary HIV counseling and testing acceptance, sexual risk behavior and HIV incidence in Rakai, Uganda. *AIDS.* 2005;19:503–11.

59. Matovu JKB, Gray RH, Kiwanuka N, Kigozi G, Wabwire-Mangen F, Nalugoda F, Serwadda D, Sewankambo NK, Wawer MJ. Repeat voluntary HIV counseling and testing (VCT), sexual risk behavior and HIV incidence in Rakai, Uganda. *AIDS Behav.* 2007;11:71–8.
60. Moses S, Plummer FA, Ngugi EN, Nagelkerke NJ, Anzala AO, Ndinya-Achola JO. Controlling HIV in Africa: effectiveness and cost of an intervention in a high-frequency STD transmitter core group. *AIDS.* 1991;5:407–11.
61. Moses S, Ramesh BA, Nagelkerke NJD, et al. Impact of an intensive HIV prevention programme for female sex workers on HIV prevalence among antenatal clinic attenders in Karnataka state, south India: an ecological analysis. *AIDS.* 2008;22:S101–8.
62. Ng M, Gakidou E, Levin-Rector A, Khera A, Murray CJL, Dandona L. Assessment of population-level effect of Avahan, an HIV-prevention initiative in India. *Lancet.* 2011;378:1643–52.
63. Nkambule R, Philip NM, Reid G, et al. HIV incidence, viremia, and the national response in Eswatini: two sequential population-based surveys. *PLoS ONE.* 2021;16: e0260892.
64. Patterson TL, Mausbach B, Lozada R, et al. Efficacy of a brief behavioral intervention to promote condom use among female sex workers in Tijuana and Ciudad Juarez, Mexico. *Am J Public Health.* 2008;98:2051–7.
65. Peak A, Rana S, Maharjan SH, Jolley D, Crofts N. Declining risk for HIV among injecting drug users in Kathmandu, Nepal: the impact of a harm-reduction programme. *AIDS.* 1995;9:1067–70.
66. Pickles M, Boily M-C, Vickerman P, et al. Assessment of the population-level effectiveness of the Avahan HIV-prevention programme in South India: a preplanned, causal-pathway-based modelling analysis. *Lancet Glob Health.* 2013;1:E289–99.
67. Ramesh BM, Beattie TSH, Shajy I, Washington R, Jagannathan L, Reza-Paul S, Blanchard JF, Moses S. Changes in risk behaviours and prevalence of sexually transmitted infections following HIV preventive interventions among female sex workers in five districts in Karnataka state, south India. *Sex Transm Infect.* 2010;86:117–24.
68. Reza-Paul S, Beattie T, Syed HUR, et al. Declines in risk behaviour and sexually transmitted infection prevalence following a community-led HIV preventive intervention among female sex workers in Mysore, India. *AIDS.* 2008;22:S91–100.
69. Roland ME, Neilands TB, Krone MR, Coates TJ, Franses K, Chesney MA, Kahn JS, Martin JN. A randomized noninferiority trial of standard versus enhanced risk reduction and adherence counseling for individuals receiving post-exposure prophylaxis following sexual exposures to HIV. *Clin Infect Dis.* 2011;53:76–83.
70. Sabido M, Giardina F, Hernandez G, Hugo Fernandez V, Ernesto Monzon J, Ortiz R, Montoliu A, Casabona J. The UALE Project: decline in the incidence of HIV and sexually transmitted infections and increase in the use of condoms among sex workers in Guatemala. *JAIDS-J Acquir Immune Defic Syndr.* 2009;51:S35–41.
71. Safren SA, Thomas B, Biello KB, et al. Strengthening resilience to reduce HIV risk in Indian MSM: a multicity, randomised, clinical efficacy trial. *Lancet Glob Health.* 2021;9:e446–55.
72. Sherman SG, Sutcliffe C, Sirojnj B, Latkin CA, Aramratanna A, Celentano DD. Evaluation of a peer network intervention trial among young methamphetamine users in Chiang Mai, Thailand. *Soc Sci Med.* 2009;68:69–79.
73. Subramanian T, Ramakrishnan L, Aridoss S, et al. Increasing condom use and declining STI prevalence in high-risk MSM and TGs: evaluation of a large-scale prevention program in Tamil Nadu, India. *BMC Public Health.* 2013. <https://doi.org/10.1186/1471-2458-13-857>.
74. Sweat M, Kerrigan D, Moreno L, Rosario S, Gomez B, Jerez H, Weiss E, Barrington C. Cost-effectiveness of environmental-structural communication interventions for HIV prevention in the female sex industry in the Dominican Republic. *J Health Commun.* 2006;11:123–42.
75. Thilakavathi S, Boopathi K, Girish Kumar CP, et al. Assessment of the scale, coverage and outcomes of the Avahan HIV prevention program for female sex workers in Tamil Nadu, India: is there evidence of an effect? *BMC Public Health.* 2011;11(Suppl 6):S3.
76. Thirumurthy H, Bair EF, Ochwal P, Marcus N, Putt M, Maman S, Napierala S, Agot K. The effect of providing women sustained access to HIV self-tests on male partner testing, couples testing, and HIV incidence in Kenya: a cluster-randomised trial. *Lancet HIV.* 2021;8:e736–46.
77. Thuong NV, Van Nghia K, Hau TP, et al. Impact of a community sexually transmitted infection/HIV intervention project on female sex workers in five border provinces of Vietnam. *Sex Transm Infect.* 2007;83:376–82.
78. Tinajeros F, Miller WM, Castro L, Artilles N, Flores F, Evans JL, Mendoza S, Urquia M, Rodriguez X, Paz-Bailey G. Declining sexually transmitted infections among female sex workers: the results of an HIV and sexually transmitted infection prevention strategy in Honduras, 2006–08. *Int J STD AIDS.* 2012;23:88–93.
79. Traore IT, Meda N, Hema NM, et al. HIV prevention and care services for female sex workers: efficacy of a targeted community-based intervention in Burkina Faso. *J Int AIDS Soc.* 2015. <https://doi.org/10.7448/IAS.18.1.20088>.
80. Vickerman P, Terris-Prestholt F, Delany S, Kumaranayake L, Rees H, Watts C. Are targeted HIV prevention activities cost-effective in high prevalence settings? Results from a sexually transmitted infection treatment project for sex workers in Johannesburg, South Africa. *Sex Transm Dis.* 2006;33:S122–32.
81. Wagman JA, Gray RH, Campbell JC, et al. Effectiveness of an integrated intimate partner violence and HIV prevention intervention in Rakai, Uganda: analysis of an intervention in an existing cluster randomised cohort. *Lancet Glob Health.* 2015;3:e23–33.
82. Wechsberg WM, Zule WA, El-Bassel N, Doherty IA, Minnis AM, Novak SD, Myers B, Carney T. The male factor: Outcomes from a cluster randomized field experiment with a couples-based HIV prevention intervention in a South African township. *Drug Alcohol Depend.* 2016;161:307–15.
83. Williams JR, Alary M, Lowndes CM, et al. Positive impact of increases in condom use among female sex workers and clients in a medium HIV prevalence epidemic: modelling results from Project SIDA 1/2/3 in Cotonou, Benin. *PLoS ONE.* 2014. <https://doi.org/10.1371/journal.pone.0102643>.
84. Ye S, Xiao Y, Jin C, Cassell H, Blevins M, Sun J, Vermund SH, Qian H-Z. Effectiveness of integrated HIV prevention interventions among Chinese men who have sex with men: evaluation of a 16-city public health program. *PLoS ONE.* 2012;7: e50873.
85. Yu J, Zhang Y, Jiang J, et al. Implementation of a “county-Township-Village” Allied HIV Prevention and Control Intervention in Rural China. *AIDS Patient Care STDs.* 2017;31:384–93.
86. Pearl J. *Causality: models, reasoning, and inference.* New York: Cambridge University Press; 2000.
87. Pearl J. *Causal inference in statistics: an overview.* *Stat Surv.* 2009;3:96–146.
88. Bennett S, Parpia T, Hayes R, Cousens S. Methods for the analysis of incidence rates in cluster randomized trials. *Int J Epidemiol.* 2002;31:839–46.
89. Hayes RJ, Moulton LH. *Cluster randomised trials, second edition.* 2017. <https://doi.org/10.4324/9781315370286>
90. Murray DM. *Design and analysis of group-randomized trials.* Oxford: Oxford University Press; 1998.
91. Wu Z, Frangakis CE, Louis TA, Scharfstein DO. Estimation of treatment effects in matched-pair cluster randomized trials by

- calibrating covariate imbalance between clusters. *Biometrics*. 2014;70:1014–22.
92. Chandrasekaran P, Dallabetta G, Loo V, et al. Evaluation design for large-scale HIV prevention programmes: the case of Avahan, the India AIDS initiative. *AIDS*. 2008;22:S1.
 93. Ye S, Xiao Y, Jin C, Cassell H, Blevins M, Sun J, Vermund SH, Qian H-Z. Effectiveness of integrated HIV prevention interventions among Chinese men who have sex with men: evaluation of a 16-city public health program. *PLoS ONE*. 2012. <https://doi.org/10.1371/journal.pone.0050873>.
 94. Crepaz N, Tungol-Ashmon MV, Higa DH, et al. A systematic review of interventions for reducing HIV risk behaviors among people living with HIV in the United States, 1988–2012. *AIDS*. 2014;28:633–56.
 95. Lyles CM, Kay LS, Crepaz N, Herbst JH, Passin WF, Kim AS, Rama SM, Thadiparthi S, DeLuca JB, Mullins MM. Best-evidence interventions: findings from a systematic review of HIV behavioral interventions for US populations at high risk, 2000–2004. *Am J Public Health*. 2007;97:133–43.
 96. Ota E, Wariki WM, Mori R, Hori N, Shibuya K. Behavioral interventions to reduce the transmission of HIV infection among sex workers and their clients in high-income countries. *Cochrane Database Syst Rev*. 2011. <https://doi.org/10.1002/14651858.CD006045.pub3>.
 97. Sipe TA, Barham TL, Johnson WD, Joseph HA, Tungol-Ashmon ML, O’Leary A. Structural interventions in HIV prevention: a taxonomy and descriptive systematic review. *AIDS Behav*. 2017;21:3366–430.
 98. Townsend L, Mathews C, Zembe Y. A systematic review of behavioral interventions to prevent HIV infection and transmission among heterosexual, adult men in low-and middle-income countries. *Prev Sci*. 2013;14:88–105.
 99. Wariki WM, Ota E, Mori R, Koyanagi A, Hori N, Shibuya K. Behavioral interventions to reduce the transmission of HIV infection among sex workers and their clients in low- and middle-income countries. *Cochrane Database Syst Rev*. 2012. <https://doi.org/10.1002/14651858.CD005272.pub3>.
 100. Padian NS, McCoy SI, Balkus JE, Wasserheit JN. Weighing the gold in the gold standard: challenges in HIV prevention research. *AIDS*. 2010;24:621–35.
 101. Brookmeyer R. Measuring the HIV/AIDS epidemic: approaches and challenges. *Epidemiol Rev*. 2010;32:26–37.
 102. Petticrew M, Cummins S, Ferrell C, Findlay A, Higgins C, Hoy C, Kearns A, Sparks L. Natural experiments: an underused tool for public health? *Public Health*. 2005;119:751–7.
 103. Stephenson J, Imrie J. Why do we need randomised controlled trials to assess behavioural interventions? *BMJ*. 1998;316:611–3.
 104. Sullivan GM. Getting off the “gold standard”: randomized controlled trials and education research. *J Grad Med Educ*. 2011;3:285–9.
 105. Reichardt CS. Cause and effect. New York: The Guilford Press; 2019. p. 11–25.
 106. Reichardt CS. Threats to validity. New York: The Guilford Press; 2019. p. 26–44.
 107. Reichardt CS. One-group posttest-only designs. New York: The Guilford Press; 2019. p. 94–8.
 108. Proctor E, Siltmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, Griffey R, Hensley M. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health Ment Health Serv Res*. 2011;38:65–76.
 109. Cambon L, Minary L, Ridde V, Alla F. A tool to facilitate transferability of health promotion interventions: ASTAIRE. *Sante Publique (Bucur)*. 2014;26:783–6.
 110. Wang S, Moss JR, Hiller JE. Applicability and transferability of interventions in evidence-based public health. *Health Promot Int*. 2006;21:76–83.
 111. Chen HT. The theory-driven approach to outcome evaluation. 2nd ed. Los Angeles: SAGE Publications; 2015. p. 304–39.
 112. Cambon L, Alla F. Understanding the complexity of population health interventions: assessing intervention system theory (ISyT). *Health Res Policy Syst*. 2021;19:95.
 113. Altman DG, Simera I. A history of the evolution of guidelines for reporting medical research: the long road to the EQUATOR Network. *J R Soc Med*. 2016;109:67–77.
 114. Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research reporting guidelines. *PLoS Med*. 2010;7:e1000217.
 115. Simera I, Moher D, Hirst A, Hoey J, Schulz KF, Altman DG. Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network. *BMC Med*. 2010;8:24.
 116. Altman DG, Simera I. Responsible reporting of health research studies: transparent, complete, accurate and timely. *J Antimicrob Chemother*. 2010;65:1–3.
 117. Page MJ, McKenzie JE, Higgins JPT. Tools for assessing risk of reporting biases in studies and syntheses of studies: a systematic review. *BMJ Open*. 2018;8: e019703.
 118. Dane AV, Schneider BH. Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clin Psychol Rev*. 1998;18:23–45.
 119. Victora CG, Habicht J-P, Bryce J. Evidence-based public health: moving beyond randomized trials. *Am J Public Health*. 2004;94:400–5.
 120. Chen HT. Logic models and the action model/ change model schema (program theory). 2nd ed. Los Angeles: SAGE Publications; 2015. p. 58–93.
 121. Jessani NS, Williamson RT, Choonara S, Gautier L, Hoe C, Jafar SK, Khalid AF, Rodríguez Salas I, Turcotte-Tremblay A-M, Rodríguez DC. Evidence attack in public health: Diverse actors’ experiences with translating controversial or misrepresented evidence in health policy and systems research. *Glob Public Health*. 2022;0:1–17.
 122. Craig P, Di Ruggiero E, Frohlich KL, et al (2018) Taking account of context in population health intervention research: guidance for producers, users and funders of research. <https://doi.org/10.3310/CIHR-NIHR-01>
 123. McGill E, Er V, Penney T, et al. Evaluation of public health interventions from a complex systems perspective: a research methods review. *Soc Sci Med*. 2021;272: 113697.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Appendix

Appendix 1: PRISMA checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Title page
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	L 1-14
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	L 25-66
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	L 66-69
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	L 83-103
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	L 71-82
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	L 71-82 L 104-113
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	L 74-81 L 83-103
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	L 102-139
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	L 104-139
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	L 104-139
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Not applicable
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	Not applicable
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	L 104-139
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Not applicable
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	L 104-139 Ref #19

Section and Topic	Item #	Checklist item	Location where item is reported
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	L 104-139 Ref #19
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	L 104-139 Ref #19
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	Not applicable
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	Not applicable
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Not applicable
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Figure 1
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	L 134-136
Study characteristics	17	Cite each included study and present its characteristics.	L 141-143 Appendix 4
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Not applicable
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Not applicable
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	L 140-184
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	L 140-184 Table 1 to 4
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	L 140-184
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	Not applicable
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	Not applicable
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Not applicable
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	L 185-222 L 279-293
	23b	Discuss any limitations of the evidence included in the review.	L 279-293
	23c	Discuss any limitations of the review processes used.	L 279-293
	23d	Discuss implications of the results for practice, policy, and future research.	L 223-278
OTHER INFORMATION			
Registration	24a	Provide registration information for the review, including register name	L 71-73

Section and Topic	Item #	Checklist item	Location where item is reported
and protocol		and registration number, or state that the review was not registered.	
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	L 71-73
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	Not applicable
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Not applicable
Competing interests	26	Declare any competing interests of review authors.	Title page
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Not applicable

Appendix 2: Search strategy

Databases	Search terms
Web of Science	<p>i) Terms associated with prevalence and incidence of HIV infection</p> <p>TS=(HIV OR "HIV infection" OR "human immunodeficiency virus" OR AIDS)</p> <p>AND TS=(prevalence OR incidence OR "HIV acquisition" OR "risk of infection" OR "infection rate" OR "rate of infection" OR "averted infections" OR "infections averted")</p>
	<p>ii) Terms associated with the notion of effectiveness of interventions</p> <p>AND TS=(effect OR effects OR efficacy OR effectiveness OR impact*)</p> <p>AND TS=(intervention* OR program* OR project OR trial*)</p>
	<p>iii) Terms associated with prevention interventions and sexual risk exposure</p> <p>AND TS=(prevention OR prophylaxis OR "primary prevention")</p>

	<p>AND TS=("unsafe sex" OR "safe sex" OR "sex work" OR "sex worker" OR "sex workers" OR "unprotected sex" OR "protected sex" OR "high-risk sex" OR "high risk sex" OR "sexual risk")</p> <p>iv) Language AND LA=(English OR French)</p> <p>v) Document type NOT DT=("RETRACTED PUBLICATION" OR BOOK OR "book chapter" OR "PROCEEDINGS PAPER" OR "MEETING ABSTRACT")</p>
PubMed	<p>i) Terms associated with prevalence and incidence of HIV infection (HIV[Title/Abstract] OR "HIV infection"[Title/Abstract] OR "human immunodeficiency virus"[Title/Abstract] OR "AIDS"[Title/Abstract]) AND ("prevalence"[Title/Abstract] OR "incidence"[Title/Abstract] OR ("averted"[Title/Abstract] AND "infections"[Title/Abstract]) OR ("rate"[Title/Abstract] AND "infections"[Title/Abstract]) OR ("risk"[Title/Abstract] AND "infections"[Title/Abstract]))</p> <p>ii) Terms associated with the notion of effectiveness of interventions AND (effect [Title/Abstract] OR effects [Title/Abstract] OR efficacy [Title/Abstract] OR effectiveness [Title/Abstract] OR impact*[Title/Abstract]) AND (intervention*[Title/Abstract] OR program*[Title/Abstract] OR project[Title/Abstract] OR trial*[Title/Abstract])</p>

	<p>iii) Terms associated with prevention interventions and sexual risk exposure</p> <p>AND (prevention[Title/Abstract] OR prophylaxis[Title/Abstract] OR "primary prevention"[MeSH])</p> <p>AND ("unsafe sex"[MeSH Terms] OR "safe sex"[MeSH Terms] OR "sex work"[MeSH Terms] OR "unprotected sex"[Title/Abstract] OR "protected sex"[Title/Abstract] OR "sexual risk"[Title/Abstract] OR "high-risk sex"[Title/Abstract] OR "high-risk sex"[Title/Abstract] OR ("women"[MeSH Terms] AND "hiv infections/prevention and control"[MeSH Terms]) OR ("young adult"[MeSH Terms] AND "hiv infections/prevention and control"[MeSH Terms]))</p>
	<p>iv) Language</p> <p>AND (english[Filter] OR french[Filter])</p>
	<p>v) Document type</p> <p>NOT (bookdocs[Filter] OR congress[Filter] OR retractedpublication[Filter])</p>
<p>Scopus</p>	<p>i) Terms associated with prevalence and incidence of HIV infection</p> <p>TITLE-ABS-KEY(hiv OR "HIV infection" OR "human immunodeficiency virus" OR AIDS) AND TITLE-ABS-KEY(prevalence OR incidence OR "averted infections" OR "infections averted" OR "infection rate" OR "rate of infection" OR "risk of infection" OR "HIV acquisition")</p> <p>ii) Terms associated with the notion of effectiveness of interventions</p>

	<p>AND TITLE-ABS-KEY(effect OR effects OR efficacy OR effectiveness OR impact*) AND TITLE-ABS-KEY(intervention* OR program* OR project OR trial*)</p>
	<p>iii) Terms associated with prevention interventions and sexual risk exposure</p> <p>AND TITLE-ABS-KEY (prevention OR prophylaxis OR "primary prevention")</p> <p>AND (TITLE-ABS-KEY ("unsafe sex" OR "safe sex" OR "sex work" OR "sex worker" OR "sex workers" OR "unprotected sex" OR "protected sex" OR "high-risk sex" OR "high risk sex" OR "sexual risk") OR KEY ("young adult" AND "prevention" AND "HIV") OR KEY ("women" AND "prevention" AND "HIV"))</p>
	<p>iv) Language</p> <p>AND (LIMIT-TO(LANGUAGE, "English") OR LIMIT-TO (LANGUAGE, "French"))</p>
	<p>v) Document type</p> <p>AND (EXCLUDE(DOCTYPE, "tb") OR EXCLUDE(DOCTYPE, "ch") OR EXCLUDE(DOCTYPE, "cp"))</p>

Appendix 3: Details of the items used to assess the reporting of the design

Items	Quasi-experimental studies	Experimental studies
Sampling	Are the participants representative of the target population?	Is randomization appropriately performed?
Treatment allocation	Are measurements appropriate regarding both the outcome and intervention (or exposure)?	Are the groups comparable at baseline?
Outcome integrity	Are there complete outcome data?	Are there complete outcome data?
Consideration of potential confounders and measurement biases	Are the confounders accounted for in the design ?	Are outcome assessors blinded to the intervention provided ?
Administration / adherence	During the study period, is the intervention administered (or exposure occurred) as intended?	Did the participants adhere to the assigned intervention?

Annexe 3

Article 3 : Développement et validation d'une échelle de mesure de l'empowerment en santé sexuelle dans le contexte de l'étude MAKASI

ELEMENT SOUS DROIT, DIFFUSION NON AUTORISEE

Annexe 4

Article 4 : Modélisation de l'impact d'une intervention de prévention du VIH basée sur la réduction des vulnérabilités et des inégalités sociales sur le risque de transmission chez les immigrants originaires d'Afrique subsaharienne vivant en Île-de-France

Auteurs: Andrainolo Ravalihasy^{1,2,3} *et al**

¹Ceped (IRD-Université Paris Cité-Inserm ERL 1244), Paris (France)

²DeSCID, IAME UMR1137 (Université Paris Cité-Inserm), Paris (France)

³Institut Convergence Migration, Aubervilliers, France

* Ce travail étant en cours, la liste des co-auteurs reste à définir

Contexte

En France, les immigrés originaires d’Afrique subsaharienne constituent le deuxième groupe le plus touché par l’épidémie VIH. Une étude parue en 2018 estime le taux d’incidence à 19.2‰ et 25.1‰ personnes-années respectivement chez les hommes et les femmes hétérosexuelles issues de cette population, témoignant d’une incidence allant jusqu’à 100 fois celle de la population générale (Marty et al., 2018). Malgré l’existence de différents dispositifs en faveur de l’accès aux soins et à la prévention, plusieurs études montrent que cette population demeure difficile à atteindre (Cazein et al., 2020). On remarque un retard à l’accès à une prise en charge antirétrovirale chez les personnes vivant avec le VIH nées à l’étranger (Cazein et al., 2019). Il en va de même pour l’utilisation de la prophylaxie pré et post exposition du fait de la méconnaissance de ces outils (Carillon & Gosselin, 2020; Coulibaly, Gosselin, et al., 2023; Hadj et al., 2017). L’insuffisance de l’utilisation des outils de prévention classiques tels que le dépistage et le préservatif est également à souligner. Actuellement, les stratégies de promotion de l’utilisation de ces outils ne ciblent pas suffisamment cette population. La diversification des outils de prévention est pourtant pertinente pour freiner la propagation de l’infection auprès d’immigrés originaires d’Afrique subsaharienne et encourage la mise en place de stratégies de promotion adaptées.

Par ailleurs, l’intrication de plusieurs facteurs explique la vulnérabilité par rapport à l’épidémie dans cette population. Il faut notamment remarquer l’influence des facteurs socioéconomiques et structurels sur l’épidémie qui démontrent la complexité de la prévention et de la prise en charge du VIH. En effet, ces immigrés font face à des situations de précarité associées directement et indirectement à l’acquisition du VIH et l’exposition au VIH (Desgrees-du-Lou et al., 2015, 2016). Ce constat souligne l’importance des actions de prévention combinant l’utilisation d’outils biomédicaux à la prise en considération des facteurs comportementaux et les problématiques structurelles. Ce type d’action de prévention est en ligne avec les recommandations de l’Organisation mondiale de la Santé en termes de programmes de prévention du VIH et de la définition même de prévention combinée du VIH (Joint United Nations Programme on HIV/AIDS, 2007, 2015). Dans le cadre de cette étude, nous étudions l’efficacité de l’intervention MAKASI sur le risque d’acquisition du VIH auprès des immigrés originaires d’Afrique subsaharienne vivant en France. L’intervention MAKASI propose d’améliorer l’autonomie dans l’utilisation des ressources, outils et dispositifs existant dans le système de soins et la capacité de protection dans les interactions sexuelles et relationnelles (Gosselin et al., 2019). Pour ce faire, un entretien motivationnel individuel est proposé pour

améliorer la capacité d'action (empowerment) des individus à prendre en main leur santé sexuelle et leur situation sociale. Dans le cadre de cet entretien, un bilan personnalisé en santé sexuelle assorti d'une orientation voire un accompagnement dans le système de santé et social est effectué. L'amélioration de l'accès aux droits et aux soins est pensée pour réduire les situations de précarité. L'entretien motivationnel et le bilan personnalisé en santé sexuelle sont pensés pour réduire la fréquence des rapports sexuels non protégés et l'occurrence des relations transactionnelles. Dans cette étude de modélisation, nous explorons l'impact de l'intervention sur le risque d'acquisition du VIH à travers l'utilisation du préservatif, l'utilisation de la prévention combinée et l'accès aux ressources sociales.

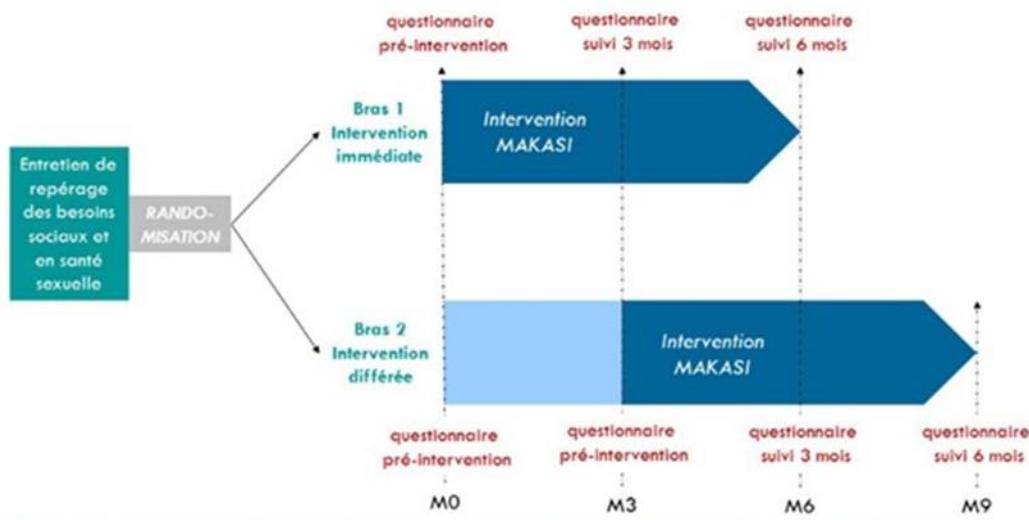
Données

Données du projet MAKASI

La mise en œuvre de l'intervention s'effectue dans le cadre d'une action de sensibilisation à la santé sexuelle et d'offre de test rapide d'orientation diagnostique du VIH menée par l'équipe mobile d'une association, Afrique Avenir, sur des lieux stratégiques de passage (marchés, places, gares de RER...). Le recrutement des participants a eu lieu entre le 01/04/2018 et le 31/12/2020. L'intervention est proposée aux personnes nées en Afrique subsaharienne ne vivant pas avec le VIH qui déclarent i) des situations sexuelles qui les exposent particulièrement à l'infection VIH, ou ii) des situations de précarité sociale (instabilité résidentielle, absence de titre de séjour ou titre de séjour de moins d'un an, isolement social), ou iii) des difficultés d'accès aux services de santé (absence de couverture maladie, méconnaissance du système de santé). La population ciblée lors de l'intervention se distingue par la fréquence des situations de précarité en comparaison avec la population générale des immigrants originaires d'Afrique subsaharienne vivant en France. Les informations concernant ce projet sont disponibles en ligne (<https://www.projet-makasi.fr/>).

Les participants à l'étude sont répartis aléatoirement dans deux bras : un bras intervention immédiate incluant les personnes bénéficiant de l'intervention le jour de leur recrutement et un bras intervention différée incluant les personnes recevant l'intervention trois mois après avoir été recrutées. Les participants à l'étude sont suivis six mois pour observer les indicateurs au moment où ils bénéficient de l'intervention (M0) ainsi que trois mois (M3) et six mois (M6) après en avoir bénéficié.

Figure 1 : Schéma de répartition des groupes pour l'évaluation de l'intervention MAKASI



Les informations collectées sont d'ordre différent entre les éligibles et les participants à l'étude. Un questionnaire de repérage recueille des informations concernant tous les éligibles, dont les caractéristiques sociodémographiques et les critères d'inclusion (âge, lieu de naissance, exposition aux risques sexuels, précarité et isolement social, accès aux services de santé et de prévention). Un questionnaire d'inclusion administré au moment de bénéficier de l'intervention (M0) recueille les données sur les pratiques contraceptives et la sexualité, les connaissances des ressources sociales et de santé, les compétences en santé sexuelle, l'empowerment et le recours aux soins durant les trois derniers mois. Le questionnaire d'inclusion, ainsi que des questionnaires de suivi (M3 et M6) contiennent des informations sur la situation sociale et de santé, et permettent d'observer les évolutions en lien avec l'intervention.

Données du recensement 2017

Le recensement de la population générale française effectué par l'INSEE permet de définir de manière précise les caractéristiques démographiques et sociales des individus et des ménages. Il est réalisé au niveau des communes et au sein des ménages issus de ces communes. En fonction de la taille des communes où s'effectue l'enquête, l'échantillonnage est exhaustif (<10000 habitants) ou se fait par sondage (≥ 10000 habitants). Dans le cadre de cette modélisation, les données utilisées sont restreintes à celles concernant les personnes nées étrangères dans un pays d'Afrique subsaharienne recensées en Île-de-France. Les informations utilisées sont celles permettant de décrire les caractéristiques sociodémographiques et migratoires (sexe, âge, niveau d'étude et durée de séjour en France) des immigrants originaires d'Afrique subsaharienne vivant en Île-de-France.

Données du projet PARCOURS

L'enquête ANRS-Parcours (Agence Nationale de Recherche sur le Sida et les hépatites, 2013) est une enquête observationnelle transversale avec une approche biographique conduite en 2012-2013 auprès de 3 groupes de personnes originaires d'Afrique subsaharienne en Île-de-France : des personnes infectées par le VIH, des personnes porteuses d'une hépatite B chronique, et des individus n'ayant aucune de ces deux infections. L'enquête a été menée dans des structures de soin recensées de façon exhaustive. Pour chaque groupe, un échantillonnage aléatoire stratifié à deux degrés proportionnels au nombre d'immigrés sub-sahariens suivis a été réalisé. L'objectif de l'étude ANRS-Parcours était de comprendre chez les migrants subsahariens comment les facteurs sociaux et individuels se conjuguent au cours du parcours de migration et d'installation en France pour augmenter ou diminuer les risques d'infection, favoriser ou réduire l'accès à la prévention et aux soins, optimiser ou limiter l'efficacité de la prise en charge pour les deux maladies VIH et hépatite B. Pour tous les participants, des informations détaillées sur les caractéristiques sociodémographiques, les conditions de la migration et de la vie en France, l'histoire de vie relationnelle, sexuelle et reproductive, l'histoire du dépistage et du parcours de soin ont été collectées à l'aide d'un questionnaire biographique administré en face-à-face par un enquêteur spécialisé. Les informations cliniques étaient documentées à partir des dossiers médicaux par le personnel soignant. Les données étaient collectées de façon anonyme. Nous redressons ces données avec celles du recensement de la population de l'année 2017 concernant les personnes nées étrangères dans un pays d'Afrique subsaharienne et vivant en Île-de-France en faisant un calage sur marge (Kolenikov, 2019) tenant compte du sexe, de l'âge, du niveau d'étude et de la durée de séjour en France. Le résultat permet d'obtenir des distributions identiques à celles observées dans les données du recensement relatives aux variables citées précédentes.

Données du baromètre santé 2016

Le Baromètre santé est une enquête périodique représentative de la population française, menée depuis 1992, qui vise à mieux connaître les connaissances, les attitudes, les croyances et les comportements des Français en matière de santé. L'enquête est réalisée à l'aide d'un système d'interview par téléphone assistée par ordinateur. Un double échantillonnage aléatoire avec chevauchement (par téléphone fixe et par téléphone mobile) a ainsi été adopté pour améliorer la couverture théorique de l'enquête. Les informations recueillies concernent les

caractéristiques sociodémographiques, la santé en général (tabagisme, drogues, ...), les comportements de dépistage et de vaccination, et la santé sexuelle (Blineau et al., 2016).

Données de la littérature

Une revue de la littérature a été effectuée afin d'identifier les études qui présentent des estimations correspondant aux paramètres à inclure dans le modèle. Peu d'études quantitatives se penchent sur les comportements sexuels des immigrants originaires d'Afrique subsaharienne. Nous n'avons pas pu trouver des études présentant des estimations concernant le nombre de partenaires sexuels, le nombre de rapports sexuels ou la prévalence du VIH chez les différents types de partenaires. Concernant ce dernier paramètre, seule une étude parue en 2017 présente des estimations chez une population immigrée originaire d'Afrique subsaharienne en Europe (Loos et al., 2017). En ce qui concerne les paramètres biologiques à savoir l'infectivité du VIH et l'efficacité du préservatif, des études sont disponibles et présentent des estimations variées. Dans notre étude, nous utilisons les résultats de méta-analyses concernant ces différentes études (Boily et al., 2009a; Giannou et al., 2016; Patel et al., 2014).

Méthodes

Modèle théorique

Le risque d'acquisition du VIH est modélisé à l'aide d'un processus de Bernoulli tel que décrit par Pinkerton et Abramson (Pinkerton & Abramson, 1993) où l'acquisition du VIH dépend du nombre de partenaires sexuels, du nombre de relations sexuelles avec ces derniers, de la prévalence du VIH chez les partenaires, de l'infectivité du virus (probabilité de transmission par relation sexuelle). Cette équation s'écrit comme suit pour un individu i ayant une relation monogame :

$$P_i = 1 - [(1 - \pi) + \pi(1 - \alpha)^n]^m$$

(1)

Où :

- m est le nombre de partenaires sexuels et égal à 1 pour relation monogame
- n est le nombre de relations sexuelles
- α est l'infectivité du virus
- π est la prévalence chez les partenaires

En factorisant par π , cette équation se réduit à un modèle similaire à celui décrit par Allard (Allard, 1990) et s'écrit comme suit :

$$P_i = 1 - [1 - \pi(1 - (1 - \alpha)^n)]^m$$

(2)

Afin d'étudier l'efficacité de l'intervention MAKASI, nous apportons des précisions sur certains paramètres de ce modèle. Ainsi, nous avons choisi :

- De prendre en compte la mixité relationnelle en distinguant deux types de partenaires sexuels au lieu du nombre de partenaires : ceux provenant des immigrants originaires d'Afrique subsaharienne, et ceux qui n'en proviennent pas. Cette prise en compte est nécessaire étant donné les niveaux de mixité dans la population d'étude ainsi que les niveaux de prévalence différents en fonction de ces types de partenaires.
- De prendre en compte l'efficacité du préservatif lorsqu'il est utilisé
- De distinguer deux types de relations hétérosexuelles : les relations sexuelles vaginales pénétrantes (transmission de la femme à l'homme) et les relations sexuelles vaginales réceptives (transmission de l'homme à la femme).

La prise en compte de ces différents éléments aboutit à l'équation suivante :

pour l'homme :

$$P_i = 1 - [1 - \pi_a(1 - (1 - \alpha_h)^{n_1})][1 - \pi_a(1 - (1 - (1 - \varepsilon)\alpha_h)^{n_2})][1 - \pi_p(1 - (1 - \alpha_h)^{n_3})][1 - \pi_p(1 - (1 - (1 - \varepsilon)\alpha_h)^{n_4})]$$

et pour la femme :

$$P_i = 1 - [1 - \pi_a(1 - (1 - \alpha_f)^{n_1})][1 - \pi_a(1 - (1 - (1 - \varepsilon)\alpha_f)^{n_2})][1 - \pi_p(1 - (1 - \alpha_f)^{n_3})][1 - \pi_p(1 - (1 - (1 - \varepsilon)\alpha_f)^{n_4})]$$

(3)

Où :

- π_a et π_p sont les prévalences du VIH respectivement chez les partenaires immigrants originaires d'Afrique subsaharienne et chez les partenaires non immigrants originaires d'Afrique subsaharienne
- n_1, n_2, n_3, n_4 sont les nombres de relations sexuelles respectivement non protégées avec un partenaire immigré originaire d'Afrique subsaharienne, protégées avec un partenaire immigré originaire d'Afrique subsaharienne, non protégées avec un partenaire non-originaire d'Afrique subsaharienne, et protégées avec un partenaire non-originaire d'Afrique subsaharienne.

- α_h et α_f sont les infectivités respectivement lors de relations sexuelles vaginales pénétrantes et lors de relations sexuelles vaginales réceptives.
- ε représente l'efficacité du préservatif

Le modèle est stratifié en fonction du sexe afin de refléter le niveau d'incidence et les comportements sexuels différents selon ce paramètre. Nous considérons également les groupes d'âge des personnes de 18 à 24 ans, de 25 à 34 ans et de 35 ans et plus pour stratifier selon les comportements sexuels à ces âges.

Les paramètres du modèle

Afin d'estimer au mieux les risques d'infection au VIH et les incertitudes liées à ces derniers, nous implémentons des simulations de Monte Carlo. Cette méthode permet d'estimer le risque d'acquisition du VIH en faisant varier les différents paramètres identifiés dans le modèle théorique précédent afin de quantifier aux mieux les incertitudes.

Avant d'utiliser le modèle sur les données de l'intervention, il est nécessaire de s'assurer que ce dernier est apte à quantifier ces risques de manière robuste. Pour ce faire, nous avons calibré le modèle de manière à ce qu'il puisse reproduire les risques d'acquisition du VIH chez les immigrés originaires d'Afrique subsaharienne vivant en Île-de-France.

Les comportements sexuels des immigrés originaires d'Afrique subsaharienne

L'exercice de modélisation auquel nous nous livrons nécessite des données concernant les comportements sexuels de la population, à savoir : le niveau de mixité/endogamie au sein de la population (proportions de personnes immigrées originaires d'Afrique subsaharienne et non originaire d'Afrique subsaharienne parmi les partenaires), le nombre de partenaires sexuels par mois au cours du suivi dans le modèle, la fréquence des rapports sexuels par mois.

Mixité/endogamie relationnelle

Le niveau de mixité concernant les partenaires sexuels est un paramètre important étant donné la dynamique de l'épidémie au sein de la population immigrée originaire d'Afrique subsaharienne différente comparée à la population générale. L'estimation de l'incidence du VIH dans cette population (Marty et al., 2018) témoigne de cette dynamique, en l'absence de données concernant la prévalence, celles observées dans d'autres pays en Europe (Dias, Gama, Abrantes, et al., 2020; Dias, Gama, Loos, et al., 2020; Loos et al., 2017) laisse présager le même constat. La proportion de partenaires sexuels provenant de la population immigrée originaire d'Afrique subsaharienne devient un paramètre d'autant plus important qu'il a été

constaté une forte endogamie dans cette population (Le Guen, 2018; Marsicano, 2012). Afin de disposer d'une estimation du niveau d'endogamie au sein des partenaires nous utilisons les données de l'enquête PARCOURS.

Nombre de partenaires et fréquence des rapports sexuels

Peu d'études quantitatives ont été menées pour comprendre les comportements sexuels des immigrants originaires d'Afrique subsaharienne en France. Si l'hypothèse de l'adoption des comportements observés dans le pays de naissance ou ceux observés dans le pays d'accueil peuvent être posés, les comportements observés dans cette population semblent plus complexes et dépendent notamment du contexte de socialisation sexuelle (Marsicano, 2012). S'il est difficile d'émettre des hypothèses sur les comportements à partir des données disponibles, en ce qui concerne la mise en couple, les données de l'enquête ANRS-Parcours ont permis d'observer que l'adoption des façons de se mettre en couple en France pourraient dépendre des caractéristiques sociales, démographiques et migratoires des individus (Le Guen, 2018).

Face à la complexité des comportements et le manque d'indications à partir des données disponibles, nous émettons dans notre modèle l'hypothèse que les comportements sexuels en termes de choix et de nombre de partenaires ainsi que la fréquence des rapports sexuels peuvent correspondre autant à ce qui est observé dans les pays de naissance que dans le pays d'immigration. Le fait d'adopter le comportement dans les pays de naissance ou dans le pays d'immigration est aléatoire dans notre modèle. En ce qui concerne l'adoption des comportements observés dans le pays d'accueil, nous nous basons sur des estimations effectuées à partir des données de l'enquête Baromètre Santé 2016. Plus précisément, nous estimons la probabilité d'avoir eu 0, 1, 2, 3 ou 4 partenaires sexuels dans le mois ainsi que le nombre moyen de rapports sexuels au cours du mois. Nous estimons également la probabilité d'avoir rencontré aucun, un ou plusieurs partenaires à partir des données de l'enquête ANRS Parcours afin de disposer d'estimations effectuées dans une population immigrée originaire d'Afrique subsaharienne. Enfin, en ce qui concerne la fréquence des rapports sexuels au cours du mois observés dans les pays de naissance, nous utilisons les données de la littérature (Amoo et al., 2020; Brown, 2000; Delva et al., 2013; George et al., 2019; Johnson et al., 2009; Smith et al., 2014).

Transmission du VIH

Au-delà des comportements sexuels, l'acquisition du virus est tributaire de la probabilité de transmission à chaque acte sexuel avec une personne infectée. Les estimations de ces probabilités dans la littérature sont hétérogènes, hétérogénéité souvent associée à des facteurs de risque connus : le sexe, l'exposition aux relations transactionnelles, le contexte et le pays de réalisation des études (Boily et al., 2009b; Chemaitelly et al., 2014; Powers et al., 2008). Afin de rendre compte de ces situations hétérogènes qui sont également observées dans notre population d'étude, nous utilisons une estimation issue d'une méta-analyse impliquant indifféremment des populations exposées et non exposées à ces facteurs de risque (Boily et al., 2009b). Nous utilisons une estimation commune aux deux sexes car il n'a pas été montré de différences significatives selon cette variable.

La transmission du virus est limitée lors de l'utilisation des outils de la prévention combinée. Dans la présente étude, nous considérons trois outils : le préservatif, la prophylaxie pré exposition (PrEP) et la prophylaxie post exposition (PEP). L'estimation de l'efficacité du préservatif est issue d'une méta-analyse d'études effectuées auprès de populations de couples hétérosexuels sérodiscordants (Giannou et al., 2016). De même, l'estimation de l'efficacité de la PrEP provient d'une méta-analyse d'études effectuées auprès de populations de couples hétérosexuels sérodiscordants (O Murchu et al., 2022). En ce qui concerne la PEP, peu d'études ont été en mesure d'estimer des mesures précises de l'efficacité (DeHaan et al., 2022), nous utilisons donc les résultats d'une méta-analyse portant sur des populations primates non humaines (Irvine et al., 2015), seule étude en mesure de procurer des estimations précises. Nous utilisons également les estimations présentées dans Lasry et al. (Lasry et al., 2014) concernant la probabilité cumulée d'infection rapportée au nombre moyen de rapports sexuels en une année pour explorer l'efficacité de l'utilisation conjointe des trois outils de la prévention combinée considérés.

Protection lors des interactions sexuelles

On observe au sein de la population des immigrés originaires d'Afrique subsaharienne vivant en France des vulnérabilités structurelles et sociales ainsi que des dynamiques et trajectoires sexuelles les exposant au risque d'infection au VIH (Desgrees-du-Loû et al., 2016; Marsicano, 2012; Pannetier et al., 2018). Les comportements de protection liés à la prévention du VIH sont donc intrinsèquement associés à ces paramètres. Afin d'en tenir compte dans notre modèle, nous estimons les probabilités d'adopter 3 types de comportements de protection (protection systématique, protection occasionnelle, non protection) en fonction de situations sociales et structurelles. Pour estimer dans quelles mesures ces situations influencent les

comportements de protection, nous utilisons les données de l'enquête ANRS-Parcours. Nous mettons en place des modèles de probabilité linéaire prenant successivement pour variable à expliquer la probabilité d'adopter un des comportements de protection cités précédemment, ajusté sur l'occurrence des relations stables, l'occurrence des relations occasionnelles, l'occurrence des relations transactionnelles, la situation administrative, la situation de logement et la durée de séjour en France. Nous utilisons ensuite les résultats de ces modèles pour estimer la distribution des comportements de protection dans les populations qui nous intéressent en fonction des situations sociales et structurelles. Nous aboutissons aux équations suivantes :

Chez les hommes :

$$\begin{cases} \Pr(P_{sys}) = \alpha_{sys} + \beta_{sys1} \times ts + \beta_{sys2} \times resid + \beta_{sys3} \times rl + \beta_{sys4} \times occas + \beta_{sys5} \times sejour \\ \Pr(P_{nsys}) = \alpha_{nsys} + \beta_{nsys1} \times ts + \beta_{nsys2} \times resid + \beta_{nsys3} \times rl + \beta_{nsys4} \times occas + \beta_{nsys5} \times sejour \\ \Pr(P_n) = \alpha_n + \beta_{n1} \times ts + \beta_{n2} \times resid + \beta_{n3} \times rl + \beta_{n4} \times occas + \beta_{n5} \times sejour \end{cases}$$

Chez les femmes :

$$\begin{cases} \Pr(P_{sys}) = \alpha_{sys} + \beta_{sys1} \times ts + \beta_{sys2} \times resid + \beta_{sys3} \times rl + \beta_{sys4} \times occas + \beta_{sys5} \times trans + \beta_{sys6} \times sejour \\ \Pr(P_{nsys}) = \alpha_{nsys} + \beta_{nsys1} \times ts + \beta_{nsys2} \times resid + \beta_{nsys3} \times rl + \beta_{nsys4} \times occas + \beta_{nsys5} \times trans + \beta_{nsys6} \times sejour \\ \Pr(P_n) = \alpha_n + \beta_{n1} \times ts + \beta_{n2} \times resid + \beta_{n3} \times rl + \beta_{n4} \times occas + \beta_{n5} \times trans + \beta_{n6} \times sejour \end{cases}$$

Où :

- $\Pr(P_{sys})$, $\Pr(P_{nsys})$, $\Pr(P_n)$ sont respectivement les probabilités de se protéger systématiquement, de se protéger non systématiquement, et de ne pas se protéger
- ts et $resid$ correspondent respectivement au fait de disposer d'un document administratif de séjour et au fait de disposer d'un logement
- rl , $occas$, $trans$ correspondent respectivement au fait d'avoir un partenaire stable, d'avoir un partenaire sexuel occasionnel, et d'être dans une relation transactionnelle
- $sejour$ correspond à la durée depuis l'arrivée en France

Les résultats des régressions et les coefficients associés sont présentés dans les tableaux suivants.

Tableau 1 : Coefficients associés à la probabilité de se protéger systématiquement, non systématiquement ou de ne pas se protéger chez les hommes

Protection systématique		Protection non systématique		Non protection	
Paramètre	Valeur	Paramètre	Valeur	Paramètre	Valeur
α_{sys}	0.238	α_{nsys}	0.556	α_n	0.206
β_{sys1}	0.250	β_{nsys1}	-0.369	β_{n1}	0.119
β_{sys2}	0.030	β_{nsys2}	0.021	β_{n2}	-0.051
β_{sys3}	- 0.206	β_{nsys3}	0.066	β_{n3}	0.140
β_{sys4}	0.087	β_{nsys4}	0.119	β_{n4}	-0.206
β_{sys5}	0.001	β_{nsys5}	-0.009	β_{n5}	0.008

Tableau 2 : Coefficients associés à la probabilité de se protéger systématiquement, non systématiquement ou de ne pas se protéger chez les femmes

Protection systématique		Protection non systématique		Non protection	
Paramètre	Valeur	Paramètre	Valeur	Paramètre	Valeur
α_{sys}	0.194	α_{nsys}	0.558	α_n	0.247
β_{sys1}	-0.087	β_{nsys1}	-0.048	β_{n1}	0.135
β_{sys2}	0.102	β_{nsys2}	0.037	β_{n2}	-0.139
β_{sys3}	-0.191	β_{nsys3}	0.040	β_{n3}	0.150
β_{sys4}	0.203	β_{nsys4}	0.022	β_{n4}	-0.225
β_{sys5}	-0.308	β_{nsys5}	0.195	β_{n5}	0.114
β_{sys6}	0.001	β_{nsys6}	-0.009	β_{n6}	0.008

Prévalence du VIH chez les partenaires originaires d'Afrique subsaharienne

La prévalence du VIH parmi les partenaires est le paramètre dont on dispose de moins d'informations dans la littérature. En effet, en France, peu de travaux font état de la prévalence du VIH dans les sous-groupes de population. En 2022, Marty et al. (Marty et al., 2022) estiment le nombre de personnes adultes vivant avec le VIH à 161125, parmi lesquelles 33% sont des femmes. Cette estimation ne donne cependant pas d'indications sur les sous-groupes de population les plus touchées par l'épidémie. Dans le cadre du modèle à mettre en place, elle ne permet donc pas de distinguer les personnes hétérosexuelles et celles originaires d'Afrique subsaharienne ; ces deux informations étant primordiales car les estimations à effectuer concernent des personnes hétérosexuelles issues d'une population où on observe une forte endogamie (Le Guen et al., 2017). D'une part, les résultats de l'enquête VESPA-2 permettent d'émettre l'hypothèse que 23% des personnes adultes vivant avec le VIH (Dray-Spira et al., 2013) sont originaires d'Afrique subsaharienne, et d'autre part, les données de l'INSEE permettent d'émettre l'hypothèse que l'effectif de la population adulte totale d'immigrés originaires d'Afrique subsaharienne est de 1200000 (INSEE, 2023b) dont près de la moitié sont des femmes (Beauchemin, 2020), parmi 53684803 personnes adultes (INSEE, 2023a, 2023c) en France. Malgré la possibilité d'en déduire que la population immigrée

originaires d’Afrique subsaharienne est 16 fois plus touchée par l’épidémie, ces estimations sont trop imprécises pour être utilisées dans le cadre de cette modélisation d’autant plus qu’elles ne reflètent pas forcément la dynamique de l’épidémie chez les partenaires sexuels de cette population. Nous avons donc recours au calcul bayésien approché pour estimer ce paramètre, en posant l’hypothèse que le rapport de prévalence selon le sexe, déduit des données citées précédentes, est maintenue au sein des partenaires issues des populations originaires ou non originaires d’Afrique subsaharienne.

Ainsi, on note $\frac{\pi_{am}}{\pi_{pm}} = 7.5$ et $\frac{\pi_{am}}{\pi_{pf}} = 40$, où π_{am} , π_{pm} , π_{af} , π_{pf} sont respectivement la prévalence du VIH chez les partenaires masculins originaires d’Afrique subsaharienne, la prévalence du VIH chez les partenaires masculins non originaires d’Afrique subsaharienne, la prévalence du VIH chez les partenaires féminins originaires d’Afrique subsaharienne, la prévalence du VIH chez les partenaires féminins non originaires d’Afrique subsaharienne, et les rapports obtenus sont les rapports de prévalence correspondants. Nous estimons donc successivement par calcul bayésien approché les paramètres π_{am} et π_{af} , nous en déduisons ensuite π_{pm} et π_{pf} . En utilisant le langage Python 3.9 et la bibliothèque PyMC 5.0.1, et une distribution a priori uniforme $\pi_a \sim U[0 ; 0.2]$, nous obtenons $\pi_{am} = 0.0584$ [95% CI: 0.0569 – 0.0598] et $\pi_{af} = 0.0549$ [95% CI: 0.0529 – 0.0569], donc $\pi_{pm} = 0.0078$ [95% CI: 0.0076 – 0.0080] et $\pi_{pf} = 0.0014$ [95% CI: 0.0013 – 0.0014]. Les valeurs de la variance inter et intra-chaînes ($\hat{R} < 1.05$) et le rapport entre la taille d’échantillon efficace et le nombre d’échantillon simulé ($\frac{ESS}{N} \cong 1$) rassurent sur la qualité des résultats obtenus. Les résultats obtenus sont satisfaisants pour reproduire les risques d’infections estimés dans cette population (19.6‰ et 28.6‰ personnes-années respectivement chez les hommes et chez les femmes vivant en Île-de-France (Marty et al., 2018)) dans cette population, à savoir 19.8‰ [95%CI : 17.9‰ - 21.6‰] chez les hommes et 28.7‰ [95%CI : 27.2‰ - 30.1‰] chez les femmes.

Encadré : présentation des méthodes de calcul bayésien approché

Méthode d'estimation en absence d'un estimateur de vraisemblance ou en cas de manque de données

Dans le cas des modèles où l'expression de la vraisemblance est trop complexe pour être calculée numériquement ou lorsqu'on dispose de peu d'informations pour effectuer ce calcul, la méthode du calcul bayésien approché (ABC) devient une alternative pertinente. Il s'agit d'un algorithme d'échantillonnage basé sur la méthode d'acceptation-rejet et implémenté dans le cadre des méthodes bayésiennes. Cet algorithme permet d'effectuer un échantillonnage successif dans la distribution a priori de θ du paramètre d'intérêt pour reconstituer un ensemble de données assez proches de celui engendré par un modèle basé sur une vraisemblance lorsque cette dernière ne peut être évaluée. Les étapes de mise en œuvre de cet algorithme peuvent être résumées comme suit :

Soit : $\pi(\theta|y) \propto \int(y|\theta)\pi(\theta)$ où :

- θ est un vecteur de paramètres à estimer
- y représente les données à disposition
- $\pi(\theta)$ est la distribution a priori de θ
- $\int(y|\theta)$ est la vraisemblance de θ selon les données y
- $\pi(\theta|y)$ est la distribution a posteriori de θ sachant y

Un algorithme ABC appliqué à ce modèle consiste à :

- Échantillonner successivement les vecteurs $\{\theta^{(1)}, \dots, \theta^{(N)}\}$ à partir de $\pi(\theta)$ et ensuite constituer des données \hat{y} à partir de $\int(y|\theta)$
- Garder le vecteur $\theta^{(j)}$ lorsque \hat{y} est assez proche de y en fonction de conditions prédéfinies. Ces conditions se définissent comme un ensemble de distances statistiques permettant de comparer les données obtenues à l'aide du modèle et les données observées. Dans notre cas, ces conditions peuvent être représentées par $d(y, \hat{y}) < \epsilon$ où d représente la distance entre les données observées et simulées et ϵ le seuil en dessous duquel les conditions sont réunies.
- Répéter les étapes 1) et 2) autant de fois que nécessaires pour obtenir un nombre suffisant de vecteurs $\theta^{(j)}$ indépendants et identiquement distribués provenant de la distribution a posteriori $\pi(\theta|y)$

De cette manière, il devient possible de déduire des estimations fiables dans des conditions où les informations concernant la vraisemblance du modèle sont insuffisantes. Il est aisé à partir de cette description de déduire les potentielles utilisations de cet algorithme qui ne se limitent pas aux estimations mais peuvent être étendues à plusieurs applications telles que la calibration de modèles ou la comparaison de modèles, la liste n'étant pas exhaustive. Malgré son aspect pratique, cet algorithme souffre d'imperfections limitant son utilisation dans les cas où on ne peut définir correctement les conditions qui permettent de comparer les données générées par le modèle celles attendues d'autant plus que la distribution a posteriori peut potentiellement être différente de la distribution a priori. Pour contourner ce problème, certains auteurs ont proposé l'utilisation de méthodes de Monte Carlo par chaîne de Markov (MCMC) pour améliorer l'étape correspondant à l'échantillonnage de vecteurs θ à partir de la distribution a priori. C'est une alternative intéressante à condition que la nature corrélée des échantillons obtenus, ainsi que les conditions d'acceptation des échantillons ne soient pas trop strictes au point de causer des problèmes de convergence des chaînes. Une autre alternative, celle la plus satisfaisante actuellement, consiste à appliquer la méthode séquentielle de Monte Carlo (SMC) lors de l'échantillonnage des vecteurs θ (Beaumont et al., 2002, 2009; Marjoram et al., 2003; Sisson et al., 2007; Toni et al., 2008).

Calcul bayésien approché par la méthode séquentielle de Monte Carlo (ABC-SMC)

La méthode ABC-SMC consiste à utiliser une méthode séquentielle de Monte Carlo au lieu d'utiliser une méthode MCMC. L'échantillonnage des vecteurs est basé sur une pondération séquentielle d'une première proposition du vecteur $\theta^{(j)}$ obtenue à partir de la distribution a priori de θ . La pondération est effectuée de manière à obtenir des vecteurs $\theta^{(j)}$ dont la distribution converge vers la distribution a posteriori. Un algorithme de type ABC-SMC peut s'illustrer comme suit (Beaumont et al., 2009; Marin et al., 2012; Sisson et al., 2007; Toni et al., 2008) :

Soient **1)** la séquence itérative $i = 1, \dots, T$ et N le nombre de vecteurs qu'on souhaite obtenir tel que $j = 1, \dots, N$; **2)** $\{\theta_T^{(1)}, \dots, \theta_T^{(N)}\}$ l'échantillon de vecteurs $\theta_T^{(j)}$ qu'on souhaite obtenir; et **3)** $\epsilon_1 > \dots > \epsilon_T > 0$ les seuils successifs utilisés pour constituer les vecteurs $\{\theta_T^{(1)}, \dots, \theta_T^{(N)}\}$ dans le cadre de la méthode ABC-SMC :

- A l'initialisation $i = 1$, échantillonner un vecteur $\theta_1^{(j)}$ à partir de $\pi(\theta)$ pour constituer des données \hat{y} assez proches de y selon le seuil ϵ_1 . On attribue au vecteur $\theta_1^{(j)}$ un poids $W_1^{(j)}$
- A partir de $i = 2$, effectuer un échantillonnage pondéré du vecteur $\theta_i^{(j)}$ à partir de $\theta_{i-1}^{(j)}$ à l'aide de poids $W_{i-1}^{(j)}$ déterminé à l'étape précédente et en utilisant une fonction (kernel) basée sur un processus Markovien tout en respectant les conditions $d(y, \hat{y}) < \epsilon_i$
- Calculer le poids $W_i^{(j)}$ nécessaire à l'échantillonnage à l'itération suivante calculé en fonction de la distribution a priori $\pi(\theta)$ et celle observée $\pi(\theta_i^{(j)})$ s'il s'agit de la première itération ($i = 1$) et calculé en fonction de la distribution observée à l'itération précédente $\pi(\theta_{i-1}^{(j)})$ et celle observée à l'itération considérée $\pi(\theta_i^{(j)})$ à partir de la deuxième itération ($i = 2$).

Cet algorithme est réitéré jusqu'à la constitution de l'échantillon de N vecteurs $\{\theta_T^{(1)}, \dots, \theta_T^{(N)}\}$ suffisants pour effectuer les estimations qui intéressent.

Cette méthode a l'avantage d'aboutir à un échantillon de vecteurs indépendants et identiquement distribués et d'optimiser l'échantillonnage même si la distribution a priori utilisée semble diffuse par rapport à la distribution a posteriori recherchée.

Description des scénarios

Afin d'analyser la pertinence de l'intervention MAKASI, nous explorons deux types de scénarios : un scénario d'implémentation optimale de l'utilisation des outils de la prévention combinée et un scénario d'implémentation raisonnable de l'utilisation des outils de la prévention combinée.

Dans le premier type de scénario, nous explorons successivement l'impact de l'utilisation de la PrEP, de la PEP et de l'utilisation conjointe des 3 outils de la prévention combinée avec une forte adhérence chez toute personne ayant déclaré se protéger lors des interactions sexuelles.

Dans le second type de scénario, nous considérons que les proportions dans la population de personnes utilisant uniquement le préservatif, uniquement les outils médicamenteux (PrEP/PEP) de la prévention combinée, où l'ensemble des outils (préservatif, PrEP et PEP)

suit les distributions présentées dans McMahon et al. (McMahon et al., 2022). En outre, nous considérons que ces distributions se répartissent uniquement dans le sous-groupe de la population qui a été sensibilisée à l'existence et l'accès de ces outils lors de l'intervention MAKASI telle qu'estimée par Coulibaly et al. (Coulibaly, Bousmah, et al., 2023).

Analyses de sensibilité

Afin de vérifier l'impact de l'incertitude liée à la distribution des paramètres du modèle, nous effectuons deux analyses de sensibilité univariées en faisant varier les paramètres des modèles i) dans l'étendue de l'intervalle de confiance de ces derniers et ii) à des valeurs extrêmes. Nous effectuons également une analyse de sensibilité multivariée (Blower & Dowlatabadi, 1994) en faisant varier simultanément les valeurs des paramètres du modèle dans l'étendue de leur intervalle de confiance à l'aide d'un échantillonnage par hypercube latin. Nous effectuons 20000 simulations à partir de combinaisons distinctes des valeurs des paramètres sur leur intervalle de confiance. Nous calculons ensuite les coefficients de corrélation partielles pour déterminer les paramètres ayant le plus d'influence sur les résultats du modèle. Dans cette analyse de sensibilité, les proportions d'utilisation des outils biomédicaux varient en fonction de l'impact de l'intervention sur la connaissance des outils de la prévention combinée.

Résultats

Les résultats présentés concernent 2108 individus éligibles à l'intervention. Parmi eux, 70% étaient des hommes et 30% des femmes. 13% étaient âgés de 18 à 24 ans, 35% de 25 à 34 ans, et 52% 35 ans au moins. 57% n'avaient pas de titre de séjour et 7% n'avaient pas de logement personnel. Les différents paramètres permettant de mettre en œuvre les simulations sont présentés dans le Tableau 3 et les résultats du modèle dans la Figure 2.

Tableau 3 : Paramètres du modèle

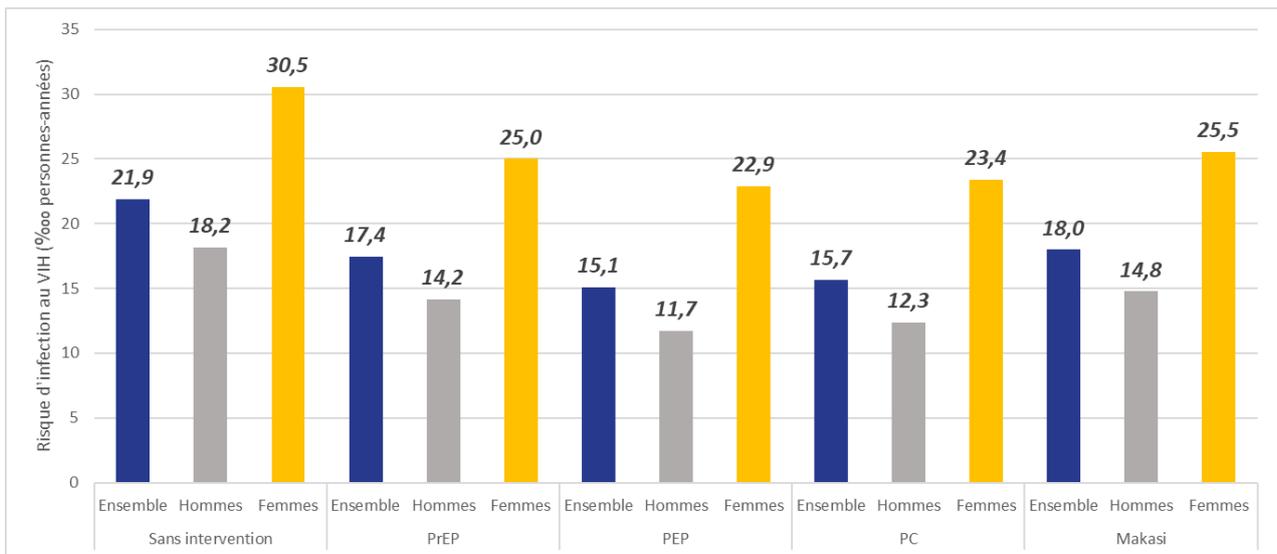
Paramètres	Valeur moyenne	Bornes des analyses de sensibilité	Distribution	Sources
Paramètres biologiques				
Infectivité	0.00182	[0.00110 – 0.00299]	Paramètre fixe	(Boily et al., 2009b)
Prévalence chez les partenaires				
Partenaires féminins	0.0549	[0.012 – 0.100]	Paramètre fixe	Estimé par ABC-SMC
Partenaires masculins	0.0584	[0.012 – 0.100]	Paramètre fixe	Estimé par ABC-SMC
Comportements sexuels				

Nombre de partenaires sexuels par mois _↓				
Hommes 18 – 24 ans	0 à 3 distribuées selon les probabilités [0.101, 0.832, 0.059, 0.008]* ou [0.407, 0.592, 0.001, 0]**	-	Distribution catégorielle	Données du Baromètre Santé 2016 et données de la littérature (Amoo et al., 2020; Brown, 2000; Delva et al., 2013; George et al., 2019; Johnson et al., 2009; Smith et al., 2014).
25 – 34 ans	0 à 3 distribuées selon les probabilités [0.063, 0.873, 0.03, 0.034]* ou [0.321, 0.646, 0.033, 0]**			
35 ans et plus	0 à 3 distribuées selon les probabilités [0.119, 0.854, 0.021, 0.006]* ou [0.267, 0.725, 0.008, 0]**			
Femmes 18 – 24 ans	0 à 3 distribuées selon les probabilités [0.081, 0.883, 0.027, 0.009]* ou [0.307, 0.693, 0.0]**	-	Distribution catégorielle	Données de l'enquête ANRS Parcours et données de la littérature (Amoo et al., 2020; Brown, 2000; Delva et al., 2013; George et al., 2019; Johnson et al., 2009; Smith et al., 2014).
25 – 34 ans	0 à 3 distribuées selon les probabilités [0.051, 0.933, 0.012, 0.004]* ou [0.133, 0.867, 0, 0]**			
35 ans et plus	0 à 3 distribuées selon les probabilités [0.227, 0.764, 0.008, 0.001]* ou [0.281, 0.719, 0, 0]**			
Nombre de rapports sexuels par mois chez les personnes ayant un partenaire*				
Hommes 18 – 24 ans	4.1* ou 3**	[3 – 5]	Distribution binomiale négative	Données du Baromètre Santé 2016 et données de la littérature (Amoo et al., 2020; Brown, 2000; Delva et al., 2013; George et al., 2019; Johnson et al., 2009;
25 – 34 ans	4.4* ou 4**	[3 – 5]		
35 ans et plus	3.7* ou 3**	[3 – 5]		

				Smith et al., 2014).
Femmes 18 – 24 ans 25 – 34 ans 35 ans et plus	3.3* ou 3** 4.7* ou 4** 3.0* ou 3**	[3 – 5] [3 – 5] [3 – 5]	Distribution binomiale négative	Données du Baromètre Santé 2016 et données de la littérature (Amoo et al., 2020; Brown, 2000; Delva et al., 2013; George et al., 2019; Johnson et al., 2009; Smith et al., 2014).
Niveau d'endogamie				
Hommes	68%	[59% – 76%]	Paramètre fixe	Données de l'enquête ANRS Parcours
Femmes	86%	[80% – 90%]	Paramètre fixe	Données de l'enquête ANRS Parcours
Efficacité des méthodes de prévention				
Préservatif	71%	[57% – 80%]	Paramètre fixe	(Giannou et al., 2016)
PrEP	75%	[54% – 86%]	Paramètre fixe	(O Murchu et al., 2022)
PEP	89%	[76% – 95%]	Paramètre fixe	(Irvine et al., 2015)
Utilisation conjointe des trois outils	86%	[80% – 92%]	Paramètre fixe	(Lasry et al., 2014)
Impact de l'intervention MAKASI				
Connaissance de la prévention combinée	60% dans l'ensemble de l'échantillon à l'issue de l'intervention	[40% – 60%]	Paramètre fixe	Données de l'intervention MAKASI
Proportion de personnes parmi celles sensibilisées effectivement utilisant et déclarant se protéger :		-	Paramètre fixe	(McMahon et al., 2022)
Le préservatif exclusivement	20%			
Les outils médicamenteux exclusivement	60%			
L'ensemble des outils	20%			

‡ Les vecteurs sont choisis de manière aléatoire et attribués aux personnes qui ont des comportements se rapprochant à ce qui est observé dans la population générale française ou non, *vecteurs choisis lorsque les personnes adoptent des comportements se rapprochant à ce qui est observé dans la population générale française, ** vecteurs choisis lorsque les personnes adoptent des comportements se rapprochant à ce qui est observé dans les régions de naissance

Figure 2 : Résultats des différents scénarios du modèle sur l'ensemble de la population



PrEP : prophylaxie pré exposition, PEP : prophylaxie post exposition, PC : prévention combinée

Dans les scénarios optimaux l'appropriation des outils médicamenteux de la prévention combinée permettent de diminuer le risque d'acquisition du VIH. Le scénario de diversification des outils de prévention permet de baisser efficacement le risque d'infection.

Le scénario « raisonnable » cohérent avec les résultats de l'intervention MAKASI montrent l'intérêt de ce type d'intervention. Dans notre exercice de modélisation, les résultats de l'intervention se rapprochent de ceux d'une implémentation optimale de la PrEP en termes de réduction de risque.

Les résultats démontrent également des dynamiques différentes selon le sexe : la diminution du risque est plus évidente chez les hommes que chez les femmes dans les différents scénarios explorés.

Les analyses de sensibilité univariées (lorsqu'on fait varier les paramètres un par un) indiquent que l'inféctivité du virus, la prévalence chez les partenaires et le nombre de rapports sexuels mensuels sont les paramètres qui influencent le plus les résultats de ce modèle (Figure 3, Figure 4). Ces résultats sont confirmés par l'analyse de sensibilité multivariée (Tableau 4) en ce qui concerne l'inféctivité du virus et la prévalence chez les partenaires. En revanche, la corrélation des résultats avec le nombre de rapports sexuels, même s'il reste significatif, est plus faible et dépend de l'âge des individus et des types de partenaires sexuels. L'analyse de sensibilité multivariée montre également que l'endogamie au sein de cette population peut jouer un rôle important étant donnée les valeurs prises par le coefficient de corrélation partielle (0.43 dans le cas des deux sexes).

Perspectives

Les résultats de notre modèle montrent que ce type d'intervention présente des potentiels intéressants en perspective avec les scénarios classiques de déploiement des outils biomédicaux. Une offre de prévention en faveur de l'accès à une offre diversifiée d'outils biomédicaux à travers un dispositif communautaire peut aider à atteindre des groupes de populations mobiles, vulnérables et difficiles à atteindre. Les analyses de sensibilité montrent qu'il est important de porter attention à différentes dynamiques complexes que n'illustrent pas forcément les scénarios abordés par le modèle. Par exemple, l'impact du niveau d'endogamie sur les estimations, pas forcément visible dans une analyse univariée, rappelle à quel point la prévalence du VIH chez les partenaires issus des régions de naissance peut affecter la dynamique de transmission de l'épidémie ou rappeler qu'il s'agit d'une population fortement exposée à ce virus.

Il n'est donc pas judicieux de se limiter à la simplicité des scénarios présentés qu'il faut absolument mettre en regard avec les réalités complexes du terrain. Pour répondre à ce défi, deux propositions ont été identifiées pour poursuivre cette analyse : affiner les scénarios à investiguer et affiner les analyses de sensibilité à effectuer. Nous nous proposons de réaliser la première proposition, affiner les hypothèses à tester, en exposant ces scénarios au regard de l'ensemble des acteurs, dont ceux qui travaillent sur le terrain, pour en améliorer la faisabilité, la viabilité et la désirabilité telles que définies par la notion de « *Quantitative Story Telling* » (Saltelli & Giampietro, 2017). De la même manière, nous allons également réaliser la deuxième proposition, affiner les analyses de sensibilité, en mettant à contribution ces acteurs (Saltelli et al., 2013) mais également en mettant nos résultats en regard avec les défis de mise en œuvre de ces programmes identifiés dans la littérature, telle que l'utilisation et l'acceptabilité des outils médicamenteux de la prévention combinée auprès des populations à qui sont destinés les programmes. Ce faisant, nous visons à fournir des estimations cohérentes avec les défis de la mise en œuvre de ces programmes.

Tableau 4: Analyse de sensibilité multivariée

Paramètres			<i>r</i>	<i>p</i>
Endogamie chez les hommes			0.43	<0.001
Endogamie chez les femmes			0.43	<0.001
Prévalence du VIH chez			0.88	<0.001
Prévalence du VIH chez			0.93	<0.001
Infectivité chez les partenaires des hommes			0.71	<0.001
Infectivité chez les partenaires des femmes			0.83	<0.001
Efficacité du préservatif			0.33	<0.001
Efficacité des outils PrEP/PEP			0.28	<0.001
Efficacité de la prévention combinée			0.08	0.011
Connaissance de la prévention combinée			0.03	0.376
Proportion d'utilisation de la PrEP			0.14	<0.001
Proportion d'utilisation du préservatif			0.15	<0.001
Proportion d'utilisation de la prévention combinée			0.24	<0.001
Nombre de rapports sexuels	Dans l'ensemble		0.35	<0.001
	Chez les hommes		0.16	<0.001
	Chez les femmes		0.08	0.017
Nombre de rapports sexuels avec des partenaires non originaires d'Afrique subsaharienne	Hommes âgés de	18-24 ans	0.10	0.001
		25-34 ans	0.07	0.027
		35 ans et plus	0.11	<0.001
	Femmes âgées de	18-24 ans	0.07	0.042
		25-34 ans	0.04	0.192
		35 ans et plus	0.06	0.063
Nombre de rapports sexuels avec des partenaires originaires d'Afrique subsaharienne	Hommes âgés de	18-24 ans	0.07	0.028
		25-34 ans	0.06	0.082
		35 ans et plus	0.12	<0.001
	Femmes âgées de	18-24 ans	0.01	0.717
		25-34 ans	0.06	0.086
		35 ans et plus	0.04	0.228

r : coefficient de corrélation partielle, *p* : p-value du coefficient de corrélation partielle

Figure 3: Analyse de sensibilité univariée aux bornes des intervalles de confiance

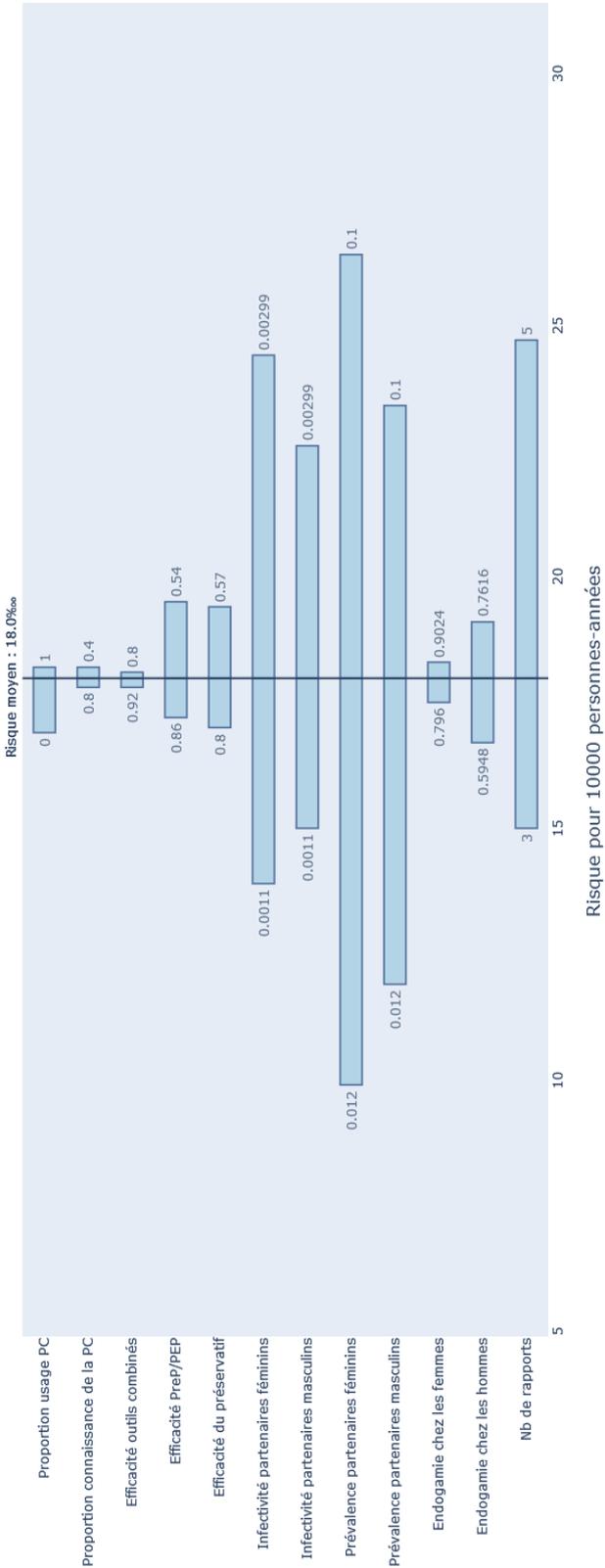
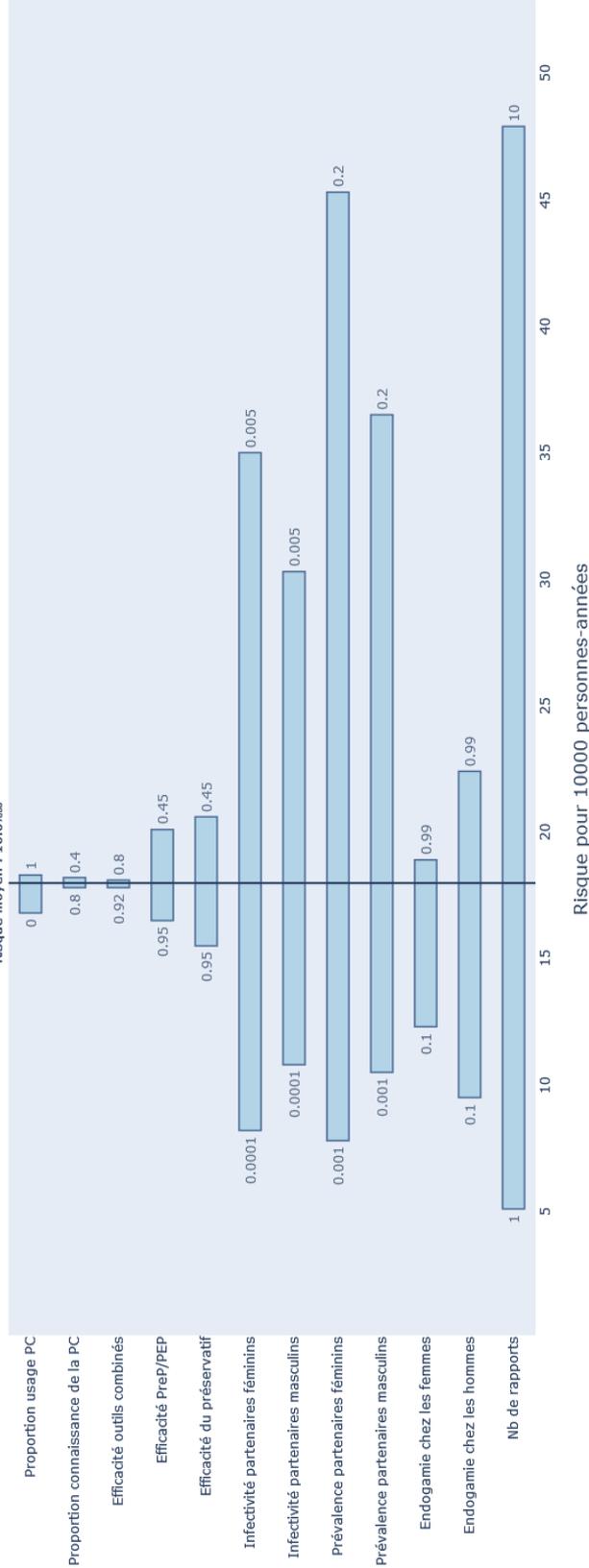


Figure 4: Analyse de sensibilité univariée aux valeurs extrêmes



Bibliographie

- Agence Nationale de Recherche sur le Sida et les hépatites. (2013). *Parcours de vie, VIH/Sida et hépatite B chez les migrants originaires d'Afrique subsaharienne vivant en Île-de-France—2012-2013* [jeu de données]. Progedo-Adisp. <https://doi.org/10.13144/LIL-1195>
- Allard, R. (1990). A Family of Mathematical Models to Describe the Risk of Infection by a Sexually Transmitted Agent. *Epidemiology*, 1(1), 30-33.
- Amoo, E. O., Solanke, B. L., Amoo, A. I., Onipede, W., & Osadolor, U. E. (2020). Descriptive analysis of men's sexual behaviour in sub-Saharan Africa : Simplicity-parsimony approach. *Cogent Arts & Humanities*. <https://www.tandfonline.com/doi/abs/10.1080/23311983.2020.1796227>
- Beauchemin, C. (2020). *Profil démographique des personnes d'origine subsaharienne en France*. <https://hal.science/hal-03097531>
- Beaumont, M. A., Cornuet, J.-M., Marin, J.-M., & Robert, C. P. (2009). Adaptive approximate Bayesian computation. *Biometrika*, 96(4), 983-990. <https://doi.org/10.1093/biomet/asp052>
- Beaumont, M. A., Zhang, W., & Balding, D. J. (2002). Approximate Bayesian Computation in Population Genetics. *Genetics*, 162(4), 2025-2035. <https://doi.org/10.1093/genetics/162.4.2025>
- Blineau, V., Gautier, A., David, C., Malti, F. E., Rahib, D., Richard, J. B., & Lydié, N. (2016). *Faut-il abandonner les appels sur téléphone fixe dans les enquêtes aléatoires en population générale par téléphone ?*
- Blower, S. M., & Dowlatabadi, H. (1994). Sensitivity and Uncertainty Analysis of Complex Models of Disease Transmission : An HIV Model, as an Example. *International Statistical Review / Revue Internationale de Statistique*, 62(2), 229-243. <https://doi.org/10.2307/1403510>
- Boily, M.-C., Baggaley, R. F., Wang, L., Masse, B., White, R. G., Hayes, R. J., & Alary, M. (2009a). Heterosexual risk of HIV-1 infection per sexual act : Systematic review and meta-analysis of observational studies. *The Lancet Infectious Diseases*, 9(2), 118-129. [https://doi.org/10.1016/S1473-3099\(09\)70021-0](https://doi.org/10.1016/S1473-3099(09)70021-0)
- Boily, M.-C., Baggaley, R. F., Wang, L., Masse, B., White, R. G., Hayes, R. J., & Alary, M. (2009b). Heterosexual risk of HIV-1 infection per sexual act : Systematic review and meta-analysis of observational studies. *The Lancet Infectious Diseases*, 9(2), 118-129. [https://doi.org/10.1016/S1473-3099\(09\)70021-0](https://doi.org/10.1016/S1473-3099(09)70021-0)
- Brown, M. S. (2000). Coitus, The Proximate Determinant Of Conception : Inter-Country Variance In Sub-Saharan Africa. *Journal of Biosocial Science*, 32(2), 145-159. <https://doi.org/10.1017/S0021932000001450>
- Carillon, S., & Gosselin, A. (2020). Tensions between Restrictive Migratory Policies and an Inclusive Prevention Programme : An Ethnography of a Biomedical HIV Prevention Programme among Sub-Saharan Africa Immigrants in the Paris Area. In N. Sahraoui (Éd.), *Borders across Healthcare* (p. 62-83). Berghahn Books. <https://doi.org/10.2307/j.ctv1tbhpx5.8>
- Cazein, F., Lot, F., Lydié, N., Velter, A., Marty, L., Supervie, V., Costagliola, D., Dray-Spira, R., Féron, V., Grémy, I., Goyet, F., Lert, F., & Doré, V. (2020). *Épidémiologie de l'infection VIH en France – 2013-2018. Tendances et contribution de la prévention combinée (dépistage, traitement antirétroviral des PVVIH, prévention par le préservatif et la PrEP)*. Agence nationale de recherches sur le sida et les

- hépatites virales, Action coordonnée « Dynamique et contrôle du VIH et des hépatites ». http://www.anrs.fr/sites/default/files/2020-07/epidemiologie-infection-vih-france-2013-2018_0.pdf
- Cazein, F., Sommen, C., Pillonel, J., Bruyand, M., Ramus, C., Pichon, P., Da Costa, C., Benyelles, L., Thierry, D., Desouches, C., Barin, F., & Lot, F. (2019). Activité de dépistage du VIH et circonstances de découverte de l'infection à VIH, France 2018. *Bulletin Épidémiologique Hebdomadaire*, 31-32, 615-624.
- Chemaitelly, H., Awad, S. F., & Abu-Raddad, L. J. (2014). The risk of HIV transmission within HIV-1 sero-discordant couples appears to vary across sub-Saharan Africa. *Epidemics*, 6, 1-9. <https://doi.org/10.1016/j.epidem.2013.11.001>
- Coulibaly, K., Bousmah, M.-Q., Ravalihasy, A., Taéron, C., Mbiribindi, R., Senne, J.-N., Gubert, F., Gosselin, A., & Desgrées du Loû, A. (2023). Bridging the knowledge gap of biomedical HIV prevention tools among sub-saharan african immigrants in France. Results from an empowerment-based intervention. *SSM - Population Health*, 23, 101468. <https://doi.org/10.1016/j.ssmph.2023.101468>
- Coulibaly, K., Gosselin, A., Carillon, S., Taéron, C., Mbiribindi, R., Loû, A. D. D., & Group, on behalf of the M. study. (2023). Low knowledge of antiretroviral treatments for the prevention of HIV among precarious immigrants from sub-Saharan Africa living in the greater Paris area : Results from the Makasi project. *PLOS ONE*, 18(6), e0287288. <https://doi.org/10.1371/journal.pone.0287288>
- DeHaan, E., McGowan, J. P., Fine, S. M., Vail, R., Merrick, S. T., Radix, A., Hoffmann, C. J., & Gonzalez, C. J. (2022). *PEP to Prevent HIV Infection*. Johns Hopkins University. <http://www.ncbi.nlm.nih.gov/books/NBK562734/>
- Delva, W., Meng, F., Beauclair, R., Deprez, N., Temmerman, M., Welte, A., & Hens, N. (2013). Coital frequency and condom use in monogamous and concurrent sexual relationships in Cape Town, South Africa. *Journal of the International AIDS Society*, 16(1), 18034. <https://doi.org/10.7448/IAS.16.1.18034>
- Desgrées-du-Lou, A., Pannetier, J., Ravalihasy, A., Gosselin, A., Supervie, V., Panjo, H., Bajos, N., Lert, F., Lydie, N., Dray-Spira, R., & Grp, P. S. (2015). Sub-Saharan African migrants living with HIV acquired after migration, France, ANRS PARCOURS study, 2012 to 2013. *EUROSURVEILLANCE*, 20(46), 31-38. <https://doi.org/10.2807/1560-7917.ES.2015.20.46.30065>
- Desgrées-du-Loû, A., Pannetier, J., Ravalihasy, A., Le Guen, M., Gosselin, A., Panjo, H., Bajos, N., Lydie, N., Lert, F., Dray-Spira, R., & Group, the P. S. (2016). Is hardship during migration a determinant of HIV infection? Results from the ANRS PARCOURS study of sub-Saharan African migrants in France. *AIDS*, 30(4), 645-656. <https://doi.org/10.1097/QAD.0000000000000957>
- Desgrées-du-Lou, A., Pannetier, J., Ravalihasy, A., Le Guen, M., Gosselin, A., Panjo, H., Bajos, N., Lydie, N., Lert, F., Dray-Spira, R., & Grp, P. S. (2016). Is hardship during migration a determinant of HIV infection ? Results from the ANRS PARCOURS study of sub-Saharan African migrants in France. *AIDS*, 30(4), 645-656. <https://doi.org/10.1097/QAD.0000000000000957>
- Dias, S., Gama, A., Abrantes, P., Gomes, I., Fonseca, M., Reigado, V., Simoes, D., Carreiras, E., Mora, C., Ferreira, A., Akpogheneta, O., & Martins, M. (2020). Patterns of Sexual Risk Behavior, HIV Infection, and Use of Health Services Among Sub-Saharan African Migrants in Portugal. *JOURNAL OF SEX RESEARCH*, 57(7), 906-913. <https://doi.org/10.1080/00224499.2019.1601154>

Dias, S., Gama, A., Loos, J., Roxo, L., Simoes, D., & Nostlinger, C. (2020). The role of mobility in sexual risk behaviour and HIV acquisition among sub-Saharan African migrants residing in two European cities. *PLOS ONE*, *15*(2). <https://doi.org/10.1371/journal.pone.0228584>

Dray-Spira, R., Wilson D'Almeida, K., Aubrière, C., Marcellin, F., Spire, B., & Lert, F. (2013). État de santé de la population vivant avec le VIH en France métropolitaine en 2011 et caractéristiques des personnes récemment diagnostiquées. Premiers résultats de l'enquête ANRS-Vespa2. *Bulletin Epidémiologique Hebdomadaire*, *26-27*, 285-292.

George, G., Maughan-Brown, B., Beckett, S., Evans, M., Cawood, C., Khanyile, D., Govender, K., & Kharsany, A. B. (2019). Coital frequency and condom use in age-disparate partnerships involving women aged 15 to 24 : Evidence from a cross-sectional study in KwaZulu-Natal, South Africa. *BMJ Open*, *9*(3), e024362. <https://doi.org/10.1136/bmjopen-2018-024362>

Giannou, F. K., Tsiara, C. G., Nikolopoulos, G. K., Talias, M., Benetou, V., Kantzanou, M., Bonovas, S., & Hatzakis, A. (2016). Condom effectiveness in reducing heterosexual HIV transmission : A systematic review and meta-analysis of studies on HIV serodiscordant couples. *Expert Review of Pharmacoeconomics & Outcomes Research*, *16*(4), 489-499. <https://doi.org/10.1586/14737167.2016.1102635>

Gosselin, A., Carillon, S., Coulibaly, K., Ridde, V., Taéron, C., Kohou, V., Zouménou, I., Mbiribindi, R., Derche, N., Desgrées du Loû, A., Gubert, F., Melchior, M., Delbe, A., Ebongue, J., El Khoury, F., Gaywahali, C., Lert, F., Lutonadio, B., Plenel, E., ... for the MAKASI Study Group. (2019). Participatory development and pilot testing of the Makasi intervention : A community-based outreach intervention to improve sub-Saharan and Caribbean immigrants' empowerment in sexual health. *BMC Public Health*, *19*(1), 1646. <https://doi.org/10.1186/s12889-019-7943-2>

Hadj, L., Du Loû, A. D., Dupont, J., & Nguyen, V.-K. (2017). Acceptabilité et freins chez les populations africaines et caribéennes vivant en Île-de-France d'une nouvelle offre de prévention du VIH: le Truvada® en prophylaxie pré-exposition (PrEP). Une enquête exploratoire. *Bulletin d'épidémiologie hebdomadaire*, *6*, 110-114.

INSEE. (2023a). *Bilan démographique 2022—Insee Première—1935*. https://www.insee.fr/fr/statistiques/6687000#tableau-figure6_radio1

INSEE. (2023b). *Immigrés et descendants d'immigrés—Immigrés et descendants d'immigrés en France*. <https://www.insee.fr/fr/statistiques/6793391>

INSEE. (2023c). *Population par sexe et groupe d'âges*. <https://www.insee.fr/fr/statistiques/2381474>

Irvine, C., Egan, K. J., Shubber, Z., Van Rompay, K. K. A., Beanland, R. L., & Ford, N. (2015). Efficacy of HIV Postexposure Prophylaxis : Systematic Review and Meta-analysis of Nonhuman Primate Studies. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, *60 Suppl 3*, S165-169. <https://doi.org/10.1093/cid/civ069>

Johnson, L., Dorrington, R., Bradshaw, D., Pillay-Van Wyk, V., & Rehle, T. (2009). Sexual behaviour patterns in South Africa and their association with the spread of HIV : Insights from a mathematical model. *Demographic Research*, *21*, 289-340. <https://doi.org/10.4054/DemRes.2009.21.11>

Joint United Nations Programme on HIV/AIDS. (2007). *Combination HIV Prevention : Tailoring and Coordinating Biomedical, Behavioural and Structural Strategies to Reduce New HIV Infections*.

https://www.unaids.org/sites/default/files/media_asset/JC2007_Combination_Prevention_paper_en_0.pdf

Joint United Nations Programme on HIV/AIDS. (2015). *Fast-tracking combination prevention*. UNAIDS, Geneva. https://www.unaids.org/sites/default/files/media_asset/20151019_JC2766_Fast-tracking_combination_prevention.pdf

Kolenikov, S. (2019). Updates to the ipfraking ecosystem. *The Stata Journal*, 19(1), 143-184. <https://doi.org/10.1177/1536867X19830912>

Lasry, A., Sansom, S. L., Wolitski, R. J., Green, T. A., Borkowf, C. B., Patel, P., & Mermin, J. (2014). HIV sexual transmission risk among serodiscordant couples : Assessing the effects of combining prevention strategies. *AIDS*, 28(10), 1521. <https://doi.org/10.1097/QAD.0000000000000307>

Le Guen, M. (2018). *Genre, conditions de vie et conjugalité autour de deux ruptures biographiques : Migration et diagnostic d'infection au VIH chez les immigré-e-s d'Afrique subsaharienne vivant en Île-de-France* [These de doctorat, Université Paris-Saclay (ComUE)]. <https://www.theses.fr/2018SACLS491>

Le Guen, M., Loû, A. D. du, Bajos, N., & Marsicano, É. (2017). Migration et évolutions des situations conjugales : Entre diversification des partenaires et persistance des asymétries de genre. In *Parcours de vie et santé des Africains immigrés en France* (p. 92-112). La Découverte. <https://www.cairn.info/parcours-de-vie-et-sante-des-africains-immigres--9782707196453-page-92.htm>

Loos, J., Nöstlinger, C., Vuylsteke, B., Deblonde, J., Ndungu, M., Kint, I., Manirankunda, L., Reyniers, T., Adobea, D., Laga, M., & Colebunders, R. (2017). First HIV prevalence estimates of a representative sample of adult sub-Saharan African migrants in a European city. Results of a community-based, cross-sectional study in Antwerp, Belgium. *PLOS ONE*, 12(4), e0174677. <https://doi.org/10.1371/journal.pone.0174677>

Marin, J.-M., Pudlo, P., Robert, C. P., & Ryder, R. J. (2012). Approximate Bayesian computational methods. *Statistics and Computing*, 22(6), 1167-1180. <https://doi.org/10.1007/s11222-011-9288-2>

Marjoram, P., Molitor, J., Plagnol, V., & Tavaré, S. (2003). Markov chain Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences*, 100(26), 15324-15328. <https://doi.org/10.1073/pnas.0306899100>

Marsicano, E. (2012). *Genre, expérience migratoire et condition minoritaire. Les trajectoires sexuelles et préventives des migrant-e-s d'Afrique subsaharienne vivant en France*.

Marty, L., Cazein, F., Panjo, H., Pillonel, J., Costagliola, D., & Supervie, V. (2018). Revealing geographical and population heterogeneity in HIV incidence, undiagnosed HIV prevalence and time to diagnosis to improve prevention and care : Estimates for France. *Journal of the International AIDS Society*, 21(3), e25100. <https://doi.org/10.1002/jia2.25100>

Marty, L., Diawara, Y., Rachas, A., Grabar, S., Costagliola, D., & Supervie, V. (2022). Projection of age of individuals living with HIV and time since ART initiation in 2030 : Estimates for France. *Journal of the International AIDS Society*, 25, e25986. <https://doi.org/10.1002/jia2.25986>

McMahon, J. M., Simmons, J., Braksmajer, A., & LeBlanc, N. (2022). HIV-serodifferent couples' perspectives and practices regarding HIV prevention strategies : A mixed methods study. *PLOS Global Public Health*, 2(8), e0000620. <https://doi.org/10.1371/journal.pgph.0000620>

O Murchu, E., Marshall, L., Teljeur, C., Harrington, P., Hayes, C., Moran, P., & Ryan, M. (2022). Oral pre-exposure prophylaxis (PrEP) to prevent HIV : A systematic review and meta-analysis of clinical effectiveness, safety, adherence and risk compensation in all populations. *BMJ Open*, *12*(5), e048478. <https://doi.org/10.1136/bmjopen-2020-048478>

Pannetier, J., Ravalihasy, A., Lydié, N., Lert, F., & Loû, A. D. du. (2018). Prevalence and circumstances of forced sex and post-migration HIV acquisition in sub-Saharan African migrant women in France : An analysis of the ANRS-PARCOURS retrospective population-based study. *The Lancet Public Health*, *3*(1), e16-e23. [https://doi.org/10.1016/S2468-2667\(17\)30211-6](https://doi.org/10.1016/S2468-2667(17)30211-6)

Patel, P., Borkowf, C. B., Brooks, J. T., Lasry, A., Lansky, A., & Mermin, J. (2014). Estimating per-act HIV transmission risk : A systematic review. *AIDS*, *28*(10), 1509-1519. <https://doi.org/10.1097/QAD.0000000000000298>

Pinkerton, S. D., & Abramson, P. R. (1993). Evaluating the Risks : A Bernoulli Process Model of HIV Infection and Risk Reduction. *Evaluation Review*, *17*(5), 504-528. <https://doi.org/10.1177/0193841X9301700503>

Powers, K. A., Poole, C., Pettifor, A. E., & Cohen, M. S. (2008). Rethinking the heterosexual infectivity of HIV-1 : A systematic review and meta-analysis. *The Lancet Infectious Diseases*, *8*(9), 553-563. [https://doi.org/10.1016/S1473-3099\(08\)70156-7](https://doi.org/10.1016/S1473-3099(08)70156-7)

Saltelli, A., & Giampietro, M. (2017). What is wrong with evidence based policy, and how can it be improved? *Futures*, *91*, 62-71. <https://doi.org/10.1016/j.futures.2016.11.012>

Saltelli, A., Guimarães Pereira, Â., Van der Sluijs, J. P., & Funtowicz, S. (2013). What do I make of your latinorum? Sensitivity auditing of mathematical modelling. *International Journal of Foresight and Innovation Policy*, *9*(2-3-4), 213-234. <https://doi.org/10.1504/IJFIP.2013.058610>

Sisson, S. A., Fan, Y., & Tanaka, M. M. (2007). Sequential Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences*, *104*(6), 1760-1765. <https://doi.org/10.1073/pnas.0607208104>

Smith, J., Nyamukapa, C., Gregson, S., Lewis, J., Magutshwa, S., Schumacher, C., Mushati, P., Hallett, T., & Garnett, G. (2014). The Distribution of Sex Acts and Condom Use within Partnerships in a Rural Sub-Saharan African Population. *PLOS ONE*, *9*(2), e88378. <https://doi.org/10.1371/journal.pone.0088378>

Toni, T., Welch, D., Strelkowa, N., Ipsen, A., & Stumpf, M. P. H. (2008). Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of The Royal Society Interface*, *6*(31), 187-202. <https://doi.org/10.1098/rsif.2008.0172>

Annexe 5

Article 5 : L'évaluateur et l'évaluation d'impact d'une intervention comme levier potentiel de l'*empowerment* au niveau communautaire : un exercice réflexif à partir de l'intervention MAKASI

Author: Andrainolo Ravalihasy^{1,2,3}

¹Ceped (IRD-Université Paris Cité-Inserm ERL 1244), Paris (France)

²DeSCID, IAME UMR1137 (Université Paris Cité-Inserm), Paris (France)

³Institut Convergence Migration, Aubervilliers, France

Correspondance :

Andrainolo Ravalihasy

Centre Population et Développement (Ceped), Institut de recherche pour le développement (IRD) et Université Paris Cité, 45 Rue des Saints-Pères, 75006 Paris, France

andrainolo.ravalihasy@ird.fr

Remerciements :

Je tiens à remercier Valéry Ridde pour ses conseils judicieux qui ont abouti à cette étude. Je remercie également Karna Coulibaly pour le temps consacré à la relecture de ce manuscrit.

Je remercie le groupe MAKASI et les participants à l'étude. Le projet MAKASI a été financé par l'Agence régionale de Santé d'Ile de France et l'Agence nationale de recherche contre le sida et les hépatites virales (ANRS).

Financements :

Aucun financement n'a été requis dans la réalisation de cet article

Conflits d'intérêts :

Pas de conflits d'intérêts

INTRODUCTION

Les recherches en santé publique visant à répondre aux besoins de santé des populations ont fait croître l'intérêt pour les interventions au design complexe (1). En effet, les approches biomédicales, à elles seules, présentent des limites qui encouragent l'utilisation d'approches holistiques tenant compte des principaux déterminants de santé (2,3). Reconnaître l'importance de ces déterminants et prendre en compte les mécanismes agissant sur leur distribution permet de mettre en place des stratégies non seulement efficaces mais aussi équitables (4,5), en améliorant le pouvoir d'agir des populations sur leur santé globale. La complexité inhérente à ce type de stratégie, ayant attrait de prime abord aux différentes composantes des interventions, est également inhérente à l'organisation des activités qui en découlent. Plusieurs guides pertinents pour l'évaluation des interventions en santé publique, dont le MRC guidance (6), proposent un ensemble de principes inspirés de l'évaluation des interventions biomédicales. Dans le cas du MRC guidance, ces principes ont évolué pour répondre de manière plus concrète aux enjeux propres aux interventions non biomédicales (7,8). Bien que pertinente, cette évolution ne répond que partiellement aux défis liés à la nature de ces interventions, notamment lorsqu'elles visent à promouvoir la santé (9,10). Par exemple, les estimations obtenues à l'aide des méthodes quantitatives sous-entendent souvent la maîtrise du contexte et des conditions d'administration de l'intervention, limitant la portée des résultats en situation réelle. Ce type d'estimation contribue à nier dans quelles mesures l'évaluation peut être influencée par son contexte (11). Dans le champ de la promotion de la santé, une analyse critique des contours de l'évaluation est en effet utile pour mieux comprendre sa mise en œuvre. Dans quelles mesures les méthodes utilisées respectent-elles les notions d'empowerment ? Les intérêts de quels intervenants sont respectés ? Comment la posture de l'évaluateur s'imbrique dans ce processus ? Répondre à autant de questions est nécessaire pour rendre compte de la pertinence de l'évaluation dans son contexte. Une des

voies proposées pour aborder ces questions complexes est l'adoption d'une démarche réflexive (12). Ce type de démarche a le potentiel de mettre à profit les expériences des intervenants pour développer une meilleure compréhension de leur rôle et améliorer leur pratique. L'évaluation des programmes, comme toute activité en promotion de la santé, bénéficierait d'une telle démarche (12). Dans la présente étude, nous proposons une analyse réflexive de l'usage des méthodes quantitatives utilisées dans l'évaluation d'impact de l'intervention MAKASI.

METHODES

L'analyse réflexive proposée dans cette étude repose sur l'expérience du premier auteur seul lors de sa participation à l'évaluation d'impact de l'intervention MAKASI. Nous adoptons une approche narrative hybride (13,14) où les processus amenant au choix des méthodes quantitatives sont identifiés de manière inductive et analysés à l'aide d'une grille déductive basée sur des cadres théoriques de la réflexivité en promotion de la santé. Nous les analysons en perspective avec les caractéristiques de l'auteur, le contexte de l'intervention et les principes de la promotion de la santé. Dans la suite de l'étude, nous utiliserons les termes : acteur pour désigner l'ensemble des personnes impliquées dans la mise en œuvre des interventions, intervenant pour désigner les acteurs qui ne sont pas issus du monde académique, bénéficiaire pour désigner les personnes à qui l'intervention est destinée, et enfin participant pour désigner l'ensemble des acteurs et bénéficiaires.

Contexte de l'étude MAKASI

Les immigrants originaires d'Afrique subsaharienne en situations de précarité sont une population clé dans la lutte contre l'épidémie de VIH en France (15,16). L'intervention MAKASI (17) propose d'améliorer leur autonomie dans l'utilisation des dispositifs socio-sanitaires existants et leur capacité à se protéger dans les interactions relationnelles. Elle se

base sur un cadre théorique de l'empowerment individuel (18) et une approche participative impliquant des chercheurs, une association communautaire (Afrique Avenir) et une organisation non gouvernementale (ARCAT). L'intervention s'insère dans une action de sensibilisation à la santé sexuelle et d'offre de test rapide d'orientation diagnostique du VIH menée par l'équipe mobile d'Afrique Avenir, sur des lieux stratégiques de passage. Les intervenants repèrent à l'aide d'un questionnaire les personnes à qui l'intervention est destinée et leur propose un entretien motivationnel (19) personnalisé avec une médiatrice de santé afin de les aider à hiérarchiser leurs besoins, et à les orienter dans le système de santé et d'aide sociale.

Une expérimentation aléatoire avec permutation séquentielle (20) a été adoptée pour analyser l'impact de l'intervention. La mise en œuvre des méthodes quantitatives est sous la supervision de chercheurs issus de différentes disciplines : un statisticien, deux démographes, trois économistes et un chercheur en santé publique. Parmi ces derniers, un statisticien, un démographe et un économiste sont chargés de la mise en œuvre technique de méthodes quantitatives. L'interprétation des résultats implique la participation de ces chercheurs et de six intervenants des associations précédentes.

Caractéristiques de l'évaluateur

L'auteur de l'article est un homme né dans un pays d'Afrique subsaharienne et installé en France depuis huit ans au début du projet MAKASI. Il a une formation initiale en pharmacie de son pays de naissance et en santé publique en France. Il occupe un poste d'ingénieur statisticien dans une structure de recherche et apporte un appui concernant l'utilisation des méthodes quantitatives dans le cadre de projets coordonnés par des chercheurs. Au sein du projet MAKASI, il supervise l'utilisation des méthodes quantitatives et prépare une thèse de doctorat en santé publique.

Démarche réflexive

Nous adoptons des positions réflexives en référence à la typologie proposée par Alexander et al. (12) pour analyser la mise en œuvre des méthodes quantitatives. La réflexivité « in action » correspond aux ajustements immédiats pendant la mise en œuvre des méthodes quantitatives. La réflexivité « on action » désigne l'analyse des méthodes alors qu'elles ont déjà été mises en œuvre. Enfin, la réflexivité « underlying action » consiste à mettre en perspective cette mise en œuvre avec les dynamiques sous-jacentes relevant des normes de l'évaluation, des interactions entre les différents acteurs, et du système dans lequel ils évoluent. Nous distinguons également un niveau de réflexivité « formative » (discuter ce qui a attiré à la méthodologie de l'évaluation) et de réflexivité « critique » (discuter ce qui relève des interactions entre acteurs dans le contexte de l'évaluation) (21).

Mise en perspective avec les principes de la promotion de la santé

L'adoption d'une démarche réflexive permet aux acteurs en promotion de la santé de comprendre les dynamiques sous-jacentes à leurs actions pour améliorer les pratiques dans le domaine (12). Afin de veiller à ne pas contribuer à exacerber les inégalités sociales de santé, il est nécessaire de mettre en perspective les principes d'équité, de justice sociale, de participation et de pouvoir d'agir des participants (22). Dans chaque résultat présenté, nous serons attentifs à trois questions :

- Dans quelles mesures les méthodes mobilisées répondent aux besoins des communautés participantes plutôt qu'aux normes imposées par l'évaluation ?
- Dans quelles mesures les communautés ont participé ou auraient pu participer à la mobilisation des méthodes ?
- Dans quelles mesures les communautés participantes peuvent s'appropriier les connaissances résultant du processus d'évaluation ?

RESULTATS

Les leçons apprises de cette analyse réflexive sont présentées dans le Tableau 1 et détaillées dans les paragraphes suivants. Nous prenons les cas d'une échelle de mesure de l'empowerment développé spécifiquement dans le cadre de l'intervention (23) et de la stratégie adoptée dans l'analyse statistique des impacts de l'intervention. Dans le cadre de ces cas, sept réunions concernant la mise en œuvre technique des méthodes quantitatives, et dix réunions concernant l'interprétation des impacts de l'intervention ont été tenues. Les résultats présentés proviennent de l'expérience du premier auteur dans l'exercice spécifique de ses fonctions d'évaluateur dans l'intervention, et au cours de ses interactions avec les chercheurs et les intervenants lors de ces réunions. Les positions réflexives présentées à la première personne évoquent les expériences et actions individuelles de l'auteur.

Réflexivité « in action »

Dans le cas de l'intervention MAKASI, une échelle (Annexe 1) a été développée pour disposer d'un indicateur permettant de mesurer l'empowerment chez les bénéficiaires. L'indicateur obtenu place le niveau d'empowerment d'un individu dans un intervalle linéaire où la plus basse et la plus haute valeurs indiquent respectivement le pire et le meilleur niveaux mesurables (24). L'acculturation à ce concept m'a cependant interpellé dans l'imbrication, la temporalité et la non-linéarité des processus en jeu (18). J'ai donc persuadé les acteurs de l'évaluation de la nécessité d'une analyse approfondie de l'évolution de chaque dimension mesurée par l'échelle pour mieux en comprendre l'usage. En conséquence, seule la dimension correspondant à la connaissance des traitements VIH et de l'accès à ces traitements a été utilisée dans un premier temps dans l'analyse d'impact de l'intervention (25,26). La question de l'utilisation de cette échelle pour mesurer l'empowerment sous ses différents aspects fait l'objet d'une réflexion commune entre les différents acteurs pour une valorisation spécifique.

Comme dans toute expérimentation aléatoire, les méthodes statistiques utilisées dans les expérimentations avec permutation séquentielle mesurent l'impact comme une différence entre des personnes recevant et ne recevant pas l'intervention, estimée sur des indicateurs prédéfinis. Hemming et al. (27) ont fait état des défis liés à l'analyse d'impact dans ce type d'expérimentation sans orienter le choix vers une méthode précise. Des problématiques liées à la sélection de l'échantillon et son attrition au cours du temps se sont ajoutés à cette imprécision. L'influence des disciplines des chercheurs impliqués dans l'évaluation a permis de recenser les méthodes utilisées pour traiter ces questions. La méthode d'Heckman (économétrie), celle des différences de différences (économétrie) (28) et les modèles marginaux structuraux (épidémiologie) (29) ont été proposés pour estimer l'impact de l'intervention en tenant compte de ces derniers. La vulgarisation de chacune des méthodes précitées dans les disciplines concernées les présente potentiellement comme une norme dans le traitement des données, mettant au second plan les enjeux statistiques. Cela peut constituer un défi lors des processus pluri/transdisciplinaires en révélant les limites de ces méthodes au-delà des disciplines d'origine. Devant la nécessité d'aller au-delà des polarisations entre disciplines, ma réaction a été d'explicitier leur spécificité statistique aux acteurs (Annexe 2) afin de rendre les choix cohérents avec les données et chaque question de recherche, à l'instar de l'impact de l'intervention sur la connaissance des outils biomédicaux (25).

Réflexivité « on action »

Le concours des intervenants dans l'analyse d'impact a permis de mettre en perspective les expériences de terrain dans la mise en œuvre de l'intervention. Cette mutualisation des savoirs techniques et expérientiels a permis d'affiner le choix des méthodes et l'interprétation des résultats ainsi que la présentation des résultats (<https://www.projet-makasi.fr/Publications-valorisation>) sous des formes adaptées aux chercheurs (comme les publications scientifiques) et aux intervenants (comme les notes de recherches).

Les indicateurs inclus dans l'échelle développée ont été recensés grâce à une revue de la littérature et ensuite adaptés au contexte de l'intervention grâce à une approche participative. Certaines questions de l'échelle, notamment celles concernant l'estime de soi, pointaient un paradoxe : bien qu'elles aient montré les propriétés psychométriques attendues (23), elles étaient difficiles à administrer et souvent sujets à reformulation. L'outil a néanmoins suscité l'intérêt des intervenants, illustré notamment par une valorisation commune des acteurs (30), pour identifier les usagers en besoin afin de mieux organiser et penser les actions de terrain, soulignant l'importance de l'adoption de l'outil par les intervenants.

La mise en commun, dans le cadre du processus participatif, des résultats statistiques a été à la fois une contrainte et une opportunité pour l'évaluation d'impact de l'intervention. A travers la création d'un espace de concertation entre acteurs, les problématiques statistiques et les normes de l'évaluation se sont incrustées dans le quotidien des intervenants. L'exploration des méthodes à utiliser a amené à présenter l'impact de l'intervention sous différentes formes dans un langage technique durant les réunions d'interprétation de données. La diversité des acteurs impliqués a permis de pointer la difficulté à concilier les normes scientifiques et le pragmatisme du terrain. Il a été nécessaire de créer en amont un espace technique restreint aux acteurs mettant en œuvre les méthodes pour traiter intégralement les discours techniques et les traduire ensuite dans un langage facilitant les échanges dans l'espace de concertation. En réaction, les savoirs expérientiels ont donné sens à certaines observations de prime abord contre-intuitives, expliquant parfois l'absence d'impact de l'intervention par le poids des inégalités structurelles dont font objet certains bénéficiaires.

Réflexivité « underlying action »

Mes actions tout au long de l'évaluation ont été influencées par des enjeux intérieurs liés à ma position dans l'intervention et des enjeux extérieurs liés aux disciplines, aux normes de

l'évaluation, et aux acteurs. Mon intégration comme statisticien du projet MAKASI a été facilitée par des expériences professionnelles communes avec certains acteurs. J'ai pris mes fonctions avec une légitimité professionnelle acquise et à maintenir. Au cours de la recherche, cette légitimité s'est manifestée et maintenue une participation plus importante à la supervision des analyses de données dont la mise en œuvre technique a été déléguée à deux chercheurs démographe et économiste.

La préparation de ma thèse m'amène à remettre en question les normes de l'évaluation face au contexte de l'intervention. L'acculturation à la notion d'empowerment au cœur de l'intervention amène à questionner dans quelles mesures les méthodes mobilisées sont aptes à refléter les réels besoins des bénéficiaires. Cette démarche scientifique a amené à adapter les stratégies d'analyse de données à l'image de la prise en compte des temporalités des processus d'empowerment.

Enfin, mon parcours m'a permis de comprendre certaines situations rapportées par les intervenants concernant les bénéficiaires. Ces situations liées au statut migratoire et au système socio-sanitaire en France (31) engendrent des situations communément vécues avec les bénéficiaires, telles que la difficulté à obtenir un titre de séjour pérenne. Malgré cette proximité vis-à-vis des bénéficiaires, je n'ai pu me substituer aux intervenants dans le processus d'évaluation et l'interprétation des résultats. Ma contribution résidait dans l'effort de vulgarisation des méthodes d'évaluation pour améliorer la littératie des intervenants de manière à faciliter leur participation et mettre à profit leur expérience.

Tout au long de l'évaluation, ma position vis-à-vis de l'usage des méthodes a évolué. La prise de fonction au sein du projet était accompagnée de normes liées au métier d'ingénieur statisticien et à l'évaluation dans le cadre des expérimentations aléatoires. Cependant, la préparation d'une thèse et ma proximité vis-à-vis des bénéficiaires m'ont incité à me

positionner et repositionner mes actions afin de garantir non seulement la légitimité de mes fonctions mais aussi la cohérence méthodologique vis-à-vis du contexte de l'intervention. Les marges de manœuvre acquises conséquemment m'ont permis de discuter et parfois intégrer des enjeux de l'empowerment, tels que la distinction des dimensions mesurés par l'échelle, dans les stratégies d'évaluation.

DISCUSSION

A l'instar de Ryan et Golden (32), nous avons montré que les démarches réflexives sont pertinentes dans les recherches quantitatives pour resituer les acteurs dans les dynamiques sous-jacentes sans nuire à la rigueur des résultats. Cet article s'intéresse à la position d'un acteur particulièrement concerné par les enjeux de données quantitatives, le statisticien. Son intégration dans un processus participatif constitue une opportunité pour améliorer les pratiques méthodologiques dans les programmes de santé publique. Les résultats évoqués plus haut permettent de distinguer des enjeux « formatifs » à travers les analyses « in » et « on » action, et des enjeux « critiques » de la réflexivité à travers les analyses « underlying » action.

Des enjeux formatifs

Le processus de mise en œuvre des méthodes quantitatives a permis de mettre en évidence quelques limites liées aux normes de l'évaluation et aux disciplines mobilisées pour restituer certains enjeux liés à la participation et l'empowerment des communautés concernées.

En premier lieu, notre étude illustre que l'évaluation de l'empowerment en tant que résultat ne peut se penser indépendamment du processus, à l'image de sa double nature indissociable déjà soulignée par différents auteurs à l'instar de Wallerstein ou Zimmerman (33,34). Anticiper et prendre en compte l'organisation des processus en jeu passe impérativement par

l'acculturation de l'évaluateur au concept via un processus participatif du dispositif de recherche.

En second lieu, les difficultés liées à l'administration de certaines questions de l'échelle montrent que les qualités psychométriques (35) des indicateurs dans les programmes gagneraient à être complétés par des critères pragmatiques (36). L'amélioration de la qualité pragmatique des indicateurs faciliterait alors la réalisation des activités par les intervenants.

Enfin, mené à bien, le processus d'évaluation a le potentiel de faciliter l'appropriation des outils et résultats issus de la recherche par les acteurs communautaires. Nos résultats montrent deux exemples qui illustrent ce potentiel du processus d'évaluation : l'intérêt que suscite l'échelle de mesure de l'empowerment en tant qu'outil opérationnel et la mutualisation des savoirs expérientiels comme réaction à la simplification du langage scientifique et technique. Dans notre cas, l'évaluation de l'empowerment et ses principes décrits auparavant par Fetterman (37,38) ont émergé comme un résultat du processus d'évaluation d'impact qui dépasse alors sa première vocation (quantifier l'impact de l'intervention sur la situation des bénéficiaires) pour révéler sa capacité à transformer la portée des résultats et les pratiques au niveau des acteurs.

Nos résultats réflexifs montrent qu'il est possible de distinguer le choix et la rigueur dans l'application des méthodes : choisir la méthode la plus cohérente avec la question de recherche ne remet en aucun cas sa validité statistique. En définitive, les différents réajustements méthodologiques décrits permettent d'éviter les erreurs inhérentes au choix des méthodes et recontextualiser leur usage.

Des enjeux critiques

Les normes liées à l'évaluation de programmes dépeignent l'évaluateur comme un acteur neutre (39). L'adoption d'une démarche réflexive permet de le resituer au cœur des dynamiques engendrées par son identité et le contexte des programmes.

Tel que montré ailleurs (40), notre analyse révèle des avantages et des limites liés à l'identification de chercheurs issus de la communauté bénéficiaire de l'intervention. L'appartenance communautaire dans le contexte de l'intervention MAKASI fait référence à une origine géographique. Cette identification est suffisante pour révéler les vulnérabilités structurelles au niveau des bénéficiaires (16) mais présente des limites quant à l'identification des acteurs communautaires dans la mesure où on attend une implication allant au-delà de l'expérience de ces vulnérabilités. Ainsi, l'application de ce type d'identification au niveau des acteurs de l'intervention présente les mêmes avantages et limites observés dans l'identification des bénéficiaires (41,42). Les avantages se situent dans les principes d'équité, de diversité et d'inclusion à condition que les identités des acteurs soient pensées, voire déterminées, selon une logique intersectionnelle (43). Procéder de cette manière permet de tenir compte des enjeux de pouvoir dans le système qui régit les programmes de santé. Si l'utilisation du cadre théorique de l'intersectionnalité dans la recherche en santé publique présente des défis organisationnels (44), notre étude est un exemple en faveur de sa pertinence pour une inclusion des acteurs communautaires qui ne soit pas réduite à l'état de formalité politique. Les limites, quant à elles, se situent dans les risques de stigmatisation liés à l'identification des participants. Il revient aux acteurs développant les interventions de se rappeler et rappeler que les termes « label » (41) identifiant les communautés participantes sont choisis pour révéler le poids des inégalités entravant la prise en main de leur santé. Ainsi, faire reposer la participation des acteurs communautaires sur ces termes contribue potentiellement à nier l'ampleur des inégalités structurelles à leur niveau. L'identification des différents participants aux programmes se doit de ne pas contribuer à les désigner comme

vecteurs de problématiques sanitaires, ou comme participants symboliques dans le cas des acteurs communautaires. L'implication de ces derniers dans l'évaluation est pertinente (45) mais il convient de penser en amont la proximité vis-à-vis des bénéficiaires.

Notre analyse réflexive montre que le positionnement de l'évaluateur est amené à évoluer dans le contexte de l'intervention. Ses fonctions lui confèrent une marge de manœuvre déterminante dans l'orientation de l'évaluation. L'engagement dans un processus participatif permet de mettre cette marge de manœuvre au profit d'une évaluation ancrée dans les besoins des communautés. Bien que les espaces favorisant la participation des communautés ont été privilégiés en amont de la mise en œuvre de l'intervention (46), notre analyse montre que la participation de l'évaluateur dans ces espaces a été déterminante. On constate alors que ce type d'espace dialogique favorise l'empowerment des communautés à travers les intervenants qui, ce faisant, sont mieux investis dans l'évaluation. L'intégration de l'évaluateur dans un processus participatif est donc une opportunité pour faire de l'évaluation un des leviers de l'empowerment au niveau communautaire.

PERSPECTIVES

Nombreuses sont les réserves à l'égard de l'application d'une démarche réflexive dans le cadre d'une enquête quantitative (39). Elle s'avère ici pertinente pour éviter les erreurs méthodologiques et améliorer la mise en œuvre des méthodes. L'analyse menée dans le cadre de cet article est exposé à différents biais puisqu'elle est centrée sur un seul acteur (47). Néanmoins, nous avons limité ces biais grâce à l'utilisation d'une typologie de la réflexivité et de principes inhérents à la promotion de la santé. Cet exercice a pu être réalisé grâce à la préparation d'une thèse, cependant l'engagement dans les interventions en promotion de la santé ne permettent pas toujours de dégager du temps et des ressources pour le réaliser. A travers notre étude, nous illustrons le fait que la réflexivité est une compétence nécessaire aux

évaluateurs des interventions de santé publique (48). Les exemples en faveur de sa mise en œuvre encouragent donc son intégration en amont des processus d'évaluation.

RÉFÉRENCES

1. Cambon L, Alla F. Current challenges in population health intervention research. *J Epidemiol Community Health*. 1 nov 2019;73(11):990-2.
2. Dahlgren G, Whitehead M. Policies and strategies to promote social equity in health. Background document to WHO - Strategy paper for Europe. Arbetsrapport [Internet]. déc 1991 [cité 2 févr 2023]; Disponible sur: https://ideas.repec.org//p/hhs/ifswps/2007_014.html
3. Dahlgren G, Whitehead M. The Dahlgren-Whitehead model of health determinants: 30 years on and still chasing rainbows. *Public Health*. oct 2021;199:20-4.
4. Diderichsen F, Evans T, Whitehead M. The Social Basis of Disparities in Health. In: Evans T, Whitehead M, Diderichsen F, Bhuiya A, Wirth M, éditeurs. *Challenging Inequities in Health: From Ethics to Action* [Internet]. Oxford University Press; 2001 [cité 1 févr 2023]. p. 0. Disponible sur: <https://doi.org/10.1093/acprof:oso/9780195137408.003.0002>
5. Diderichsen F, Hallqvist J, Whitehead M. Differential vulnerability and susceptibility: how to make use of recent development in our understanding of mediation and interaction to tackle health inequalities. *Int J Epidemiol*. 1 févr 2019;48(1):268-74.
6. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. *BMJ*. 16 sept 2000;321(7262):694-6.

7. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ*. 29 sept 2008;a1655.
8. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *BMJ*. 30 sept 2021;374:n2061.
9. de Salazar L, Hall M. Developing Evaluation Questions: Beyond the Technical Issues. In: Potvin L, McQueen DV, Hall M, de Salazar L, Anderson LM, Hartz ZMA, éditeurs. *Health Promotion Evaluation Practices in the Americas: Values and Research* [Internet]. New York, NY: Springer; 2009 [cité 27 avr 2021]. p. 49-62. Disponible sur: https://doi.org/10.1007/978-0-387-79733-5_4
10. Potvin L, McQueen DV. Practical Dilemmas for Health Promotion Evaluation. In: Potvin L, McQueen DV, Hall M, de Salazar L, Anderson LM, Hartz ZMA, éditeurs. *Health Promotion Evaluation Practices in the Americas: Values and Research* [Internet]. New York, NY: Springer; 2009 [cité 27 avr 2021]. p. 25-45. Disponible sur: https://doi.org/10.1007/978-0-387-79733-5_3
11. Rootman I, Goodstadt M. *Evaluation in Health Promotion: Principles and Perspectives*. WHO Regional Office Europe; 2001. 564 p.
12. Alexander SA, Jones CM, Tremblay MC, Beaudet N, Rod MH, Wright MT. Reflexivity in Health Promotion: A Typology for Training. *Health Promot Pract*. 1 juill 2020;21(4):499-509.
13. Fereday J, Muir-Cochrane E. Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. *Int J Qual Methods*. 1 mars 2006;5(1):80-92.

14. Roberts K, Dowell A, Nie JB. Attempting rigour and replicability in thematic analysis of qualitative research data; a case study of codebook development. *BMC Med Res Methodol*. 28 mars 2019;19(1):66.
15. Desgrées-du-Loû A, Pannetier J, Ravalihasy A, Gosselin A, Supervie V, Panjo H, et al. Sub-Saharan African migrants living with HIV acquired after migration, France, ANRS PARCOURS study, 2012 to 2013. *Eurosurveillance*. 19 nov 2015;20(46):30065.
16. Desgrees-du-Lou A, Pannetier J, Ravalihasy A, Le Guen M, Gosselin A, Panjo H, et al. Is hardship during migration a determinant of HIV infection? Results from the ANRS PARCOURS study of sub-Saharan African migrants in France. *AIDS Lond Engl*. 20 févr 2016;30(4):645-56.
17. Gosselin A, Carillon S, Coulibaly K, Ridde V, Taéron C, Kohou V, et al. Participatory development and pilot testing of the Makasi intervention: a community-based outreach intervention to improve sub-Saharan and Caribbean immigrants' empowerment in sexual health. *BMC Public Health*. 5 déc 2019;19(1):1646.
18. Ninacs W. Empowerment et service social: approches et enjeux. *Serv Soc*. 1995;44(1):69-93.
19. Miller WR, Rollnick S. *L'entretien motivationnel - 2e éd. - Aider la personne à engager le changement: Aider la personne à engager le changement*. 2e édition. Paris: InterEditions; 2013. 448 p.
20. French National Agency for Research on AIDS and Viral Hepatitis. MAKASI. An Intervention for African and Caribbean Migrants' Empowerment in Sexual Health in Order to Reduce Their Social and Health Vulnerabilities in Paris Greater Area [Internet].

clinicaltrials.gov; 2020 juill [cité 6 mars 2023]. Report No.: NCT04468724. Disponible sur: <https://clinicaltrials.gov/ct2/show/NCT04468724>

21. Tremblay MC, Richard L, Brousselle A, Beaudet N. Learning reflexively from a health promotion professional development program in Canada. *Health Promot Int.* 1 sept 2014;29(3):538-48.

22. World Health Organization. Ottawa Charter for Health Promotion. In Ottawa; 1986.

23. Ravalihasy A, Rude N, Yazdanpanah Y, Kardas-Sloma L, Desgrées du Loû A, Gosselin A, et al. Development and Validation of an HIV/AIDS Empowerment Scale for Impact Intervention Evaluation. An Example from the MAKASI Intervention. *Am J Health Educ.* 3 sept 2021;52(5):296-306.

24. Likert R. A technique for the measurement of attitudes. *Arch Psychol.* 1932;22 140:55-55.

25. Coulibaly K, Bousmah M al Q, Ravalihasy A, Taéron C, Mbiribindi R, Senne JN, et al. Bridging the knowledge gap of biomedical HIV prevention tools among sub-saharan african immigrants in France. Results from an empowerment-based intervention. *SSM - Popul Health.* 1 sept 2023;23:101468.

26. Bousmah M al Q. Impact d'une intervention d'empowerment en santé sur l'accès à la couverture maladie chez des personnes immigrées en situation de précarité (projet Makasi). In: SP9 Santé sexuelle [Internet]. France; 2022. Disponible sur: <https://afrahiv2022.org/replays>

27. Hemming K, Haines TP, Chilton PJ, Girling AJ, Lilford RJ. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. *BMJ.* 6 févr 2015;350:h391.

28. Angrist JD, Krueger AB. Chapter 23 - Empirical Strategies in Labor Economics. In: Ashenfelter OC, Card D, éditeurs. Handbook of Labor Economics [Internet]. Elsevier; 1999 [cité 12 mars 2023]. p. 1277-366. Disponible sur: <https://www.sciencedirect.com/science/article/pii/S1573446399030047>
29. Robins JM, Hernán MÁ, Brumback B. Marginal Structural Models and Causal Inference in Epidemiology. *Epidemiology*. sept 2000;11(5):550.
30. Ravalihasy A. Une échelle de mesure de l'empowerment en santé sexuelle : un outil adapté et adaptable aux interventions et intervenants auprès des immigrants originaires d'Afrique subsaharienne exposés au VIH. In: Session Posters discutés 1 [Internet]. France; 2022. Disponible sur: <https://livebyglevents.key4register.com/key4register/schedule.aspx?view=scheduler&details=expand&agenda=1&e=84>
31. Gosselin A, du Loû AD, Lelièvre E, Lert F, Dray-Spira R, Lydié N, et al. Understanding Settlement Pathways of African Immigrants in France Through a Capability Approach: Do Pre-migratory Characteristics Matter? *Eur J Popul Rev Eur Démographie*. 2018;34(5):849-71.
32. Ryan L, Golden A. 'Tick the Box Please': A Reflexive Approach to Doing Quantitative Social Research. *Sociology*. 1 déc 2006;40(6):1191-200.
33. Wallerstein N. Powerlessness, Empowerment, and Health: Implications for Health Promotion Programs. *Am J Health Promot*. 1 janv 1992;6(3):197-205.
34. Zimmerman MA. Taking aim on empowerment research: On the distinction between individual and psychological conceptions. *Am J Community Psychol*. 1 févr 1990;18(1):169-77.

35. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Front Public Health* [Internet]. 2018 [cité 7 mars 2023];6. Disponible sur: <https://www.frontiersin.org/articles/10.3389/fpubh.2018.00149>
36. Glasgow RE, Riley WT. Pragmatic Measures: What They Are and Why We Need Them. *Am J Prev Med*. 1 août 2013;45(2):237-43.
37. Fetterman DM. Empowerment Evaluation Principles in Practice: Assessing Levels of Commitment. In: *Empowerment evaluation principles in practice*. New York, NY, US: Guilford Press; 2005. p. 42-72.
38. Wandersman A, Snell-Johns J, Lentz BE, Fetterman DM, Keener DC, Livet M, et al. The Principles of Empowerment Evaluation. In: *Empowerment evaluation principles in practice*. New York, NY, US: Guilford Press; 2005. p. 27-41.
39. Jamieson MK, Govaart GH, Pownall M. Reflexivity in quantitative research: A rationale and beginner's guide. *Soc Personal Psychol Compass* [Internet]. 2 févr 2023 [cité 13 mars 2023]; Disponible sur: <https://onlinelibrary.wiley.com/doi/10.1111/spc3.12735>
40. Salway S, Chowbey P, Such E, Ferguson B. Researching health inequalities with Community Researchers: practical, methodological and ethical challenges of an 'inclusive' research approach. *Res Involv Engagem*. 13 août 2015;1(1):9.
41. Cloos P. La racialisation de la santé publique aux États-Unis : entre pouvoir sur la vie et droit de laisser mourir. *Glob Health Promot*. 1 mars 2012;19(1):68-75.
42. Fassin D. Généalogies. Entre pouvoirs et savoirs. In: *Faire de la santé publique* [Internet]. Rennes: Presses de l'EHESP; 2008 [cité 9 juin 2023]. p. 15-34. (Hors collection). Disponible sur: <https://www.cairn.info/faire-de-la-sante-publique--9782859529673-p-15.htm>

43. Kelly C, Dansereau L, Sebring J, Aubrecht K, FitzGerald M, Lee Y, et al. Intersectionality, health equity, and EDI: What's the difference for health researchers? *Int J Equity Health*. 19 déc 2022;21(1):182.
44. Holman D, Salway S, Bell A, Beach B, Adebajo A, Ali N, et al. Can intersectionality help with understanding and tackling health inequalities? Perspectives of professional stakeholders. *Health Res Policy Syst*. 25 juin 2021;19(1):97.
45. Robert E, Merry L, Benoît M, Guimaraes DB, Ruiz-Casares M. Rien ne doit se faire pour eux sans eux : renforcer la participation des demandeurs d'asile, réfugiés et migrants sans statut et des organismes communautaires dans la recherche en santé. *Can J Public Health*. 1 juin 2018;109(3):312-5.
46. Ridde V, Carillon S, Desgrées du Loû A, Sombié I. Analyzing implementation of public health interventions : a need for rigor, and the challenges of stakeholder involvement. *Rev D'Épidémiologie Santé Publique*. juill 2022;S0398762022003650.
47. Argyris C, Schön DA. *Theory in practice: increasing professional effectiveness*. San Francisco: Jossey-Bass publ; 1992. (The Jossey-Bass higher and adult education series).
48. van Draanen J. *Introducing Reflexivity to Evaluation Practice: An In-Depth Case Study*. *Am J Eval*. 1 sept 2017;38(3):360-75.

Tableau 1 : Leçons apprises

Position	Exemples	Implications	Recommandations
réflexive			
« in action » : ajustements immédiats pendant la mise en œuvre des méthodes quantitatives	<ul style="list-style-type: none"> • Prise en compte de la temporalité des dimensions de l'<i>empowerment</i> dans les analyses • Clarification des enjeux statistiques dans le choix des méthodes pour dépasser les positions disciplinaires 	<ul style="list-style-type: none"> • La connaissance du contexte et des objectifs de l'intervention est un atout dans les choix méthodologiques • Les dialogues interdisciplinaires permettent d'identifier les méthodes tandis que les enjeux statistiques sont primordiaux • L'usage des méthodes 	<ul style="list-style-type: none"> • Choisir des méthodes cohérentes avec la recherche et appliquer la rigueur statistique dans leur mise en œuvre • Créer un espace de concertation dialogique bénéficiant à tous les acteurs pour améliorer les pratiques évaluatives et l'utilisation des

<p>« on action » : analyse des méthodes après leur mise en œuvre</p>	<ul style="list-style-type: none"> • Création d'un espace technique restreint aux évaluateurs pour identifier les enjeux liés au choix des méthodes • Création d'un espace de concertation communs à tous les acteurs pour mutualiser les connaissances 	<ul style="list-style-type: none"> • La création d'espaces non mixtes permet de traiter de manière exhaustive les questions liées aux fonctions de chaque acteur • La création d'un espace mixte de mutualisation des informations est un impératif pour concilier les intérêts des participants 	<p>produits de l'évaluation</p> <ul style="list-style-type: none"> • Utiliser des approches intersectionnelles pour identifier les participants selon leur position dans le système des programmes de santé
<p>« underlying action » : mise en perspective de l'usage des</p>	<ul style="list-style-type: none"> • Délégation des tâches techniques pour 	<ul style="list-style-type: none"> • Une contribution effective de l'évaluateur 	

<p>méthodes avec les interactions entre acteurs et l'environnement de l'intervention</p>	<p>maintenir une marge de manœuvre dans l'orientation de l'évaluation</p> <ul style="list-style-type: none"> • Identification des spécificités liées au statut des acteurs communautaires grâce aux échanges dans l'espace de concertation 	<p>aux objectifs d'une intervention est inhérente sa marge de manœuvre dans le processus d'évaluation</p> <ul style="list-style-type: none"> • Les enjeux d'identification des participants communautaires différents selon les enjeux structurels liés à leur position dans la mise en œuvre des programmes 	
------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--

ANNEXES

Annexe 1 : Questions utilisées dans l'échelle de mesure d'*empowerment*

Dimensions de l' <i>empowerment</i> selon Ninacs	Dimensions and contenu questions abordées dans l'intervention MAKASI
Participation	Communication sexuelle
	Dans les 3 derniers mois, j'ai été capable de parler librement de sexualité à un ami, un partenaire, un membre de ma famille
	Dans les 3 derniers mois, j'ai été capable de refuser des relations sexuelles qui pouvaient comporter des risques (grossesse non désirée, IST)
	Je crois que je peux parler à mon partenaire du VIH
	Je crois que je peux être honnête avec mon partenaire à propos de mes relations sexuelles passées
	Je crois que je peux refuser d'avoir des rapports sexuels avec mon partenaire s'il ou elle refuse de se faire dépister pour le VIH
Estime de soi	Contrôle perçu
	Je contrôle les décisions qui affectent ma vie
	Je suis satisfait(e) du degré de contrôle que j'ai sur les décisions qui affectent ma vie

Conscience critique	Connaissances traitements VIH et l'accès à ces traitements
	Je crois que grâce aux traitements, une personne qui a le VIH peut mener sa vie normalement
	Je crois que quelqu'un qui a le VIH et qui prend bien son traitement ne le transmet pas pendant un rapport sexuel
	En tant que personne vivant en France, je sais que je peux avoir droit à une protection sociale et à un certain nombre d'aides sociales
Compétences techniques	Compréhension des professionnels de santé et messages de prévention
	Il est facile de comprendre ce que votre médecin vous dit
	Il est facile de suivre les instructions de votre médecin ou votre pharmacien
	Il est facile de savoir où trouver de l'aide en cas de problèmes psychologiques comme le stress ou la dépression
	Il est facile de comprendre les messages liés au tabac, à l'abus d'alcool et au manque d'activité physique
	Il est facile de comprendre l'intérêt des dépistages du VIH et des hépatites

	Il est facile de décider comment vous protéger du VIH et des IST à partir des informations données par les médias (radios, télé, affiches, internet)
--	------------------------------------------------------------------------------------------------------------------------------------------------------

Annexe 2 : Description et contraintes statistiques liées aux méthodes statistiques évoquées

Méthodes	Description	Contraintes statistiques
Méthode d'Heckman	Cette méthode consiste à déterminer, d'une part, les caractéristiques des personnes sélectionnées et non sélectionnées dans l'échantillon analyse d'impact en fonction de caractéristiques observées à l'aide d'un modèle statistique (type <i>probit</i>). D'autre part, l'impact de l'intervention est estimé en tenant compte de la probabilité d'être sélectionné dans l'échantillon d'analyse pondéré par la répartition de cette probabilité chez les personnes sélectionnées et non sélectionnées	<ul style="list-style-type: none"> • Les termes d'erreurs du modèle estimant la probabilité d'être sélectionné dans l'échantillon d'analyse et celui estimant l'impact de l'intervention doivent suivre une loi normale bivariée¹ • Il est nécessaire de disposer d'informations suffisantes concernant les personnes non sélectionnées¹ • Il est nécessaire de disposer de variables exclusives au modèle de sélection de l'échantillon et de variables exclusives au modèle d'estimation d'impact¹

<p>Modèles marginaux structuraux</p>	<p>Ces modèles fournissent des estimations de l'impact de l'intervention qui sont pondérés par la probabilité de recevoir ou non l'intervention. Cette probabilité est quant à elle estimée en fonction de caractéristiques observables mais également de caractéristiques liées à l'exposition à l'intervention dans le temps et des variables de confusion dépendantes du temps dans l'intervention. Des modèles classiques comme les modèles <i>logit</i> peuvent être mobilisés pour sa mise en œuvre</p>	<ul style="list-style-type: none"> • Il est nécessaire de disposer d'informations suffisantes et représentatives de la population d'intérêt pour calculer la probabilité d'être inclus dans l'échantillon, de recevoir l'intervention en tenant compte de l'évolution de l'exposition et des biais de confusion dans le temps. La collecte de telles informations devrait être anticipée en amont de la recherche¹
<p>Méthode des différences de différences</p>	<p>Cette méthode consiste à comparer la différence entre les personnes recevant et ne recevant pas l'intervention avant et après la mise en œuvre de celle-ci. L'ampleur de la différence avant l'intervention est comparée à l'ampleur de la</p>	<ul style="list-style-type: none"> • La validité statistique est tributaire du modèle utilisé (par exemple dans le cas d'une régression linéaire, les résidus suivent une distribution gaussienne) • Les estimations obtenues sont pertinentes à condition que

	différence après l'intervention pour en évaluer l'impact. Des modèles classiques comme les modèles linéaires ou les modèles de type <i>logit</i> peuvent être mobilisés pour sa mise en œuvre	l'échantillon d'analyse permette de faire une inférence directe sur la population d'intérêt sinon les biais de statistiques doivent être pris en compte éventuellement à l'aide de méthodes supplémentaires
--	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

¹Infante-Rivard C, Cusson A. Reflection on modern methods: selection bias—a review of recent developments.

Int J Epidemiol. 1 oct 2018;47(5):1714-22.

Liste des éléments sous droits

Titre du document	Page(s) dans la thèse
Development and Validation of an HIV/AIDS Empowerment Scale for Impact Intervention Evaluation. An Example from the MAKASI Intervention https://doi.org/10.1080/19325037.2021.1955230	144-154