

Mémoire de maîtrise

Les essais contrôlés randomisés et les politiques publiques:

L'apport de la méthodologie qualitative sur la portée de la validité externe

Étude comparative sur l'expansion de la couverture sanitaire dans le secteur informel en Indonésie et au

Nicaragua

Soumis par

Fanta Diaby

Sous la direction de la professeure

Geranda Notten

Dans le cadre du programme de maîtrise en

Affaires Publiques et Internationales

Université d'Ottawa

Le 16 juillet 2020



Sommaire

Cette étude analyse l'apport des méthodes qualitatives dans le design des évaluations contrôlées randomisées (ECR) comme levier dans le renforcement de la validité externe. Celle-ci est un des défis majeurs affectant l'aspect *gold standard* des ECR dont le design ne permet pas la production d'explications causales, à savoir le *pourquoi et comment*, propre à la « boîte noire », une intervention arrive à produire un résultat précis. En effet, si les ECR démontrent la relation causale entre une intervention, ou une politique, et un changement observé grâce à leur méthode de randomisation, elles peinent à offrir une explication claire des facteurs ayant contribué à un tel changement. D'où le problème de reproductibilité des essais, d'une part, mais surtout de contribution au développement de cadres analytiques sur les enjeux étudiés, d'autre part. La source du problème étant ainsi le design des ECR, cette étude procède à une étude comparative du design expérimental de deux cas d'études portant sur la problématique de l'expansion de la couverture universelle chez les travailleurs du secteur informel en Indonésie et au Nicaragua.

En décrivant comment les ECR sont devenus un *gold standard*, et en présentant leurs faiblesses et limites, l'étude sur la couverture sanitaire universelle (CSU) en Indonésie et en Nicaragua, montre que les méthodes qualitatives, intégrées dans le design des évaluations expérimentales, renforcent notre compréhension des résultats des expériences en analysant l'expérience des répondants et les relations entre les partenaires d'implémentation des essais expérimentaux.

Ainsi, les méthodes qualitatives renforcent la validité des données quantitatives en corroborant et en renforçant celles-ci. Cependant, pour mieux informer les politiciens, les méthodes qualitatives doivent explorer les mécanismes d'activation des impacts, inclus dans la « boîte noire », pour faciliter la réplique des politiques dans des contextes différents. Pour ce faire, les méthodes qualitatives doivent être considérées comme part intégrante du design des évaluations au même titre que celles quantitatives.

Table des matières

Sommaire	1
Introduction	4
2. Champ historique: les courants de pensées autour du <i>gold standard</i>?	7
2.1 La perspective de <i>l'evidence-based medicine</i>	7
2.2 La perspective de <i>l'evidence-based policy</i>	11
2.3 La perspective du développement international.....	16
3. Le cadre théorique des essais contrôlés randomisés	20
3.1 Les forces et faiblesses de la randomisation	20
3.1.1 Les forces.....	20
3.1.2 Les Faiblesses	21
3.2 Avantage et limite de la randomisation	23
3.2.1 Avantage	23
3.2.2 Limite.....	24
3.3 Les méthodes qualitatives dans les essais contrôlés randomisés	24
3.3.1 Le rapport entre théorie du changement et la « boîte noire »	26
3.3.2 Méthodes mixtes : entre le quantitatif et le qualitatif.....	27
3.3.3 Bénéfices des méthodes qualitatives.....	27
4. Études de cas sur la couverture sanitaire en Indonésie et au Nicaragua	31
4.1 Indonésie	31
4.1.1 Contexte.....	31
4.1.2 Méthodologie.....	33
4.1.3 Résultats.....	37
4.2 Nicaragua	41
4.2.1 Contexte.....	41
4.2.2 Méthodologie.....	42
4.2.3 Résultats.....	46
5. Analyse des études de cas	53
5.1 Inférence causale.....	53
5.2 Le contrôle des biais	54
5.3 L'explication des facteurs causals.....	56
5.4 Le rapport à la boîte noire.....	57
6. Discussion	59
7. Recommandations	63
Conclusion	64
Bibliographie	67
Annexes	74

Liste des abréviations

ECR: Essais contrôlés randomisés

EBM: Evidence-based medicine

EBP: Evidence-based policy

EMP: *Empresas Médicas Previsionales*

IMF : Institutions micro-financières

INSS: l'Institut National de Sécurité Sociale Nicaraguayen

MQ: Méthodologie qualitative

Liste des figures

Tableau 1: Apport des méthodes qualitatives au sein des évaluations contrôlés randomisées

Tableau 2: Répartition des groupes de traitements par ville

Tableau 3: Effets des subventions et de l'assistance internet sur l'enrôlement

Tableau 4: Caractéristiques des participants à l'étude

Tableau 5: Les déterminants de l'enrôlement à l'assurance

Tableau 6 : Effets de l'assurance sur les dépenses en santé et l'utilisation des services de santé

Tableau 7: Résumé des deux études de cas

Figure 1: La théorie du changement

Figure 2: Les mécanismes dans une théorie du changement

Graphique 1: Enrôlement par mois depuis le début de l'offre par type de subvention

Graphique 2: Valeurs des réclamations des frais d'hospitalisations par type de subvention

Graphique 3: Nombre de réclamations depuis le début de l'enrôlement par type de subventions

1. Introduction

Les essais contrôlés randomisés (ECR) (*randomized controlled trials* (RCTs)), aussi appelés: études par assignation aléatoire, essais randomisés de terrain, expérimentations sociales, méthode expérimentale aléatoire, ou encore évaluations contrôlées randomisées, ont récemment marqué la scène internationale à travers le laboratoire de recherche *Abdul Latif Jameel Poverty Action Lab* (J-PAL). Les fondateurs et partenaires du J-PAL, par leur démarche expérimentale dans la réduction de la pauvreté au niveau global, ont été honoré en 2019 en recevant le prix Nobel d'économie. Ces chercheurs focalisent leurs méthodes d'évaluation sur les expériences randomisées dans le but d'informer, en se basant sur des évidences scientifiques, les politiciens et les acteurs du développement sur ce qui fonctionne, à petite échelle, dans la lutte contre la pauvreté. Le J-PAL est arrivé à démontrer que l'on pouvait appliquer la randomisation dans le domaine social à travers ses évaluations de programmes de développement (Banerjee et Duflo, 2012).

Les forces clés de l'évaluation contrôlée randomisée porte sur deux points majeurs: la capacité d'établir une inférence causale et celle de produire la meilleure évidence possible à travers le contrôle des biais de sélection. En d'autres termes, la randomisation, en comparant deux groupes, soit un groupe de traitement et un groupe de contrôle, choisis aléatoirement dans une population donnée, dans le cadre d'évaluation d'une politique par exemple, tente de montrer l'effet d'un facteur X (ou d'un ensemble de facteurs) sur un résultat Y. Toute différence observée entre les résultats des deux groupes constitue une évidence de l'efficacité du Facteur X *ceteris paribus*.

Cette évidence de l'effet X sur Y constitue la force de la validité interne des essais aléatoires. De ce fait, ils sont capables de mieux éclairer les politiciens sur ce qui fonctionne dans les efforts de lutte contre la pauvreté. Cependant, une expérience prend place dans un contexte précis avec des variables et conditions propres à ce contexte. Prenant acte de ce fait, la capacité de généralisation des résultats des ECR est minée. En outre, le design d'évaluation des ECR, qu'est la randomisation, bien que se départissant

en général des aspects subjectifs, culturels et contextuels des études, en se focalisant principalement sur les facteurs quantifiables reliant causes et effets, peut être renforcé à partir d'autres méthodes d'enquête.

Face à cette limite minant la validité externe, des chercheurs (Deaton, 2010; Rodrik, 2008; Ravallion 2009; Stern et al., 2012; Gertler et al. 2011; Bamberger et al. 2010) ont porté l'attention sur des options pouvant renforcer cette limite dans les ECR. Parmi ces options, nous avons la méthodologie qualitative (MQ). Cette méthodologie fait intervenir des méthodes de collecte de données telles que: les entrevues, l'observation direct, l'analyse de contenus, les groupes de discussions, les notes de terrain, les études ethnographiques, etc. (ACYF, 2016; Dumez, 2011; Gauthier 2009). Bien que critiqué pour son aspect subjectif (analyse du chercheur) et peu scientifique (modèle d'échantillonnage), les méthodes qualitatives participent pourtant à la connaissance à travers leur dimension épistémologique prenant en compte les contextes sociétaux, et offrant des explications causales (Hammarberg et al., 2016; Drapeau, 2004). En outre, combiner diverses techniques d'enquête permet de récolter des données pouvant mieux contribuer à la compréhension des résultats d'une évaluation, mais surtout à relever les facteurs pouvant expliquer le *pourquoi et comment* les résultats diffèrent (Gertler et al., 2011; Rallis, 2017).

Nous avons trois phases au cours desquelles les méthodes qualitatives peuvent intervenir dans une évaluation expérimentale (Gertler et al., 2011, p.16-17):

- a) **La phase de conception de l'évaluation:** ici, la MQ peut aider au développement d'hypothèses qui serviront à la planification du modèle d'atteinte des objectifs souhaités mais aussi, à soulever les questions de recherche à élucider à travers l'expérimentation.
- b) **La phase intermédiaire:** ici, la MQ permet de collecter des informations pertinentes, sur l'évolution du programme, pouvant éclairer les décideurs, explorer les rétroactions des gestionnaires et des bénéficiaires tout au long de l'évaluation expérimentale.

c) **La phase d'analyse:** ici, la MQ soutient l'explication de certaines données quantitatives en les replaçant dans leur contexte et aide à expliquer les phénomènes ou variations apparaissant durant l'évaluation et pouvant affecter le déroulement de cette dernière.

Les ECR, en général, recourent aux méthodes qualitatives pour collecter des données de base qui serviront au design du programme (phase de conception) et à l'analyse des résultats finaux (phase d'analyse). Cependant, c'est durant la phase intermédiaire ou d'évaluation de processus, que les informations collectées servent à relever les facteurs affectant la réussite ou l'échec d'une intervention.

Question de recherche

Afin de comprendre l'apport des méthodes qualitatives dans les évaluations expérimentales, nous analysons la problématique de l'adhésion des travailleurs du secteur informel à la couverture universelle en Indonésie et au Nicaragua. Pour ce faire, nous nous focalisons principalement sur le design des méthodes d'évaluation de ces deux études de cas. L'Indonésie et le Nicaragua sont deux pays où la population de travailleurs dans le secteur informel occupe une large partie de la population active. Les dirigeants de ces deux pays cherchent à étendre la couverture sociale à ce groupe de leur population afin de les protéger des chocs financiers récurrents face à la nature instable de leur emploi. Ces autoentrepreneurs déboursent des montants importants dans les soins de santé pourtant, peu d'entre eux bénéficient d'une couverture santé nationale. Face aux enjeux liés à l'expansion de la couverture santé dans ce groupe, nous posons la question de savoir comment les méthodes qualitatives, intégrées dans le design des essais contrôlés randomisés, impactent la validité externe ainsi que sa portée? Étude de cas sur l'élargissement de la couverture sanitaire en Indonésie et au Nicaragua.

Ce travail est structuré en six sections principales. La section 2 brosse l'évolution des méthodes d'évaluation randomisées au niveau *gold standard* dans la science médicale et dans les sciences sociales. La section 3 décrit le cadre théorique des essais expérimentaux en relevant leurs forces, faiblesses ainsi

que leur avantage et limite. Nous y analyserons également le lien entre les ECR et les méthodes qualitatives. La section 4 présente les études de cas une après l'autre en portant une attention particulière au design d'évaluation. La section 5 porte sur une analyse de ces études de cas en soulignant leur apport dans la compréhension de la problématique soulevée. La section 6 porte sur une analyse des études de cas en lien avec les politiques sociales. Dans la section 7, il s'agit de proposer des pistes d'amélioration du design des essais randomisés incluant des méthodes qualitatives, avant de passer à la conclusion de cette étude dans la partie finale.

2. Champ historique: les courants de pensées autour du gold standard

2.1 *Evidence-based médecine*: le rapport entre la médecine et les essais randomisés

Définition

L'evidence-based medicine (EBM) est une « approche contemporaine de production et de traitement des savoirs professionnels visant à sélectionner les traitements les plus efficaces et les plus efficaces. » (Perrenoud, 2014, p.133) Le Groupe de travail, de l'université de McMaster au Canada, ayant publié *A New Approach to Teaching the Practice of Medicine* (1992) défend que « l'[EMB] désaccentue l'intuition, les expériences cliniques non-systématique, et les justifications physiopathologiques comme une base suffisante pour la prise de décisions cliniques en défendant une démarche de recherche clinique basée sur l'évidence. » (Guyatt et al., p.2420) Sackett et al. (1996) soutiennent que l'EBM est « l'utilisation consciencieuse, explicite et judicieuse de la meilleure évidence [preuve] dans la prise de décisions concernant l'état de santé d'un patient. » (Sackett et al., p.71) Autrement dit, l'EBM se focalise sur les évidences scientifiques, issue des meilleures preuves cliniques expérimentales dans l'optique d'offrir les meilleurs traitements possibles aux patients. Ces preuves sont souscrites à la rigueur de la méthode de randomisation dans les essais expérimentaux.

Origine de l'evidence-based medicine

Qu'est-ce qui explique l'engouement autour de l'*evidence-based* en médecine? Nous noterons que la découverte de la méthode de randomisation par Arnold A. Fisher et plus tard par Austin B. Hill ont eu un impact majeur sur la médecine contemporaine. Fisher est reconnu pour avoir contribué de manière importante à la conception des procédés expérimentaux. Son livre, *The design of Experiments* de 1935, fait état de sa technique d'assignation aléatoire de variétés de grains aux parcelles agricoles. Cette technique est basée sur la **rigueur comparative** et l'**assignation aléatoire** afin d'isoler l'effet de facteurs: humidité du sol, lumière, vent, sur la qualité de la récolte dans une parcelle de champ (Fisher, 1999, p.113.) Son intérêt fut de maximiser l'efficience de son test en tentant d'induire une inférence valide; c'est-à-dire montrer la relation entre l'assignation des souches et les résultats finaux dans les deux parcelles. La validité de son test est fondée sur le principe « d'estimation des erreurs par réplication », autrement par randomisation (Armitage, 2003, p. 926).

Quant à Bradford Hill, durant sa recherche sur l'épidémiologie dans les années 1930, il va mettre la méthode de la randomisation en pratique dans son essai sur la tuberculose pulmonaire (Armitage, p.926). Son apport fut particulièrement sa capacité de contrôler les biais par la randomisation (Silverman et al., 1992). Durant son étude sur le cancer du poumon, Hill a comparé deux groupes en s'assurant que les similarités en termes de santé, d'âge, de lieu d'hospitalisation, de moment d'hospitalisation des patients furent semblables d'une part, mais surtout, que l'ignorance des agents d'enquête sur l'état des malades fut assurée d'autre part. Par le contrôle de ces facteurs, Hill introduisait le principe de l'évidence causale en santé; soit la relation entre le traitement et l'effet observé en éliminant tous biais pouvant affecter la validité des résultats. Son étude conclue que « fumée est un facteur important dans la production de tissus cancérogènes dans les poumons » (Doll, 1992, p.1521). Plus tard, dans son étude sur la streptomycine pour le traitement de la tuberculose pulmonaire, dans les années 1940, Hill soulève pour la première fois une

caractéristique importante dans l'application de la randomisation. Celle de la considération éthique. Il s'agissait de savoir s'il était valable de refuser à un patient un traitement, disponible en petite quantité, qui pourrait vraisemblablement lui procurer un bénéfice? Le *Medical Research Council* conclu qu'il serait contre-éthique de ne pas tester l'efficacité d'un traitement dans un contexte où l'offre en médicament était limité par rapport au besoin (Doll, 1992, p.1525). On parlera ainsi de l'éthique de la randomisation (Amann, 2012, p.125) dans les expériences, soit donner la chance à chacun de bénéficier du traitement.

Autres aspects fondamentaux qui relèvent de l'éthique pour Hill portaient sur la généralisation des résultats et le consentement du patient. Tout d'abord, reconnaissant la nature spécifique des essais, Hill affirmait que l'éthique de l'approche spécifique (par cas) peut être corrigé par la répétition. La valeur statistique n'éliminant point le hasard attaché à chaque traitement et chaque échantillon, la généralisation implique de ce fait l'évaluation du risque dans chaque essai clinique (Amann, p.127). Sur le consentement, il défendait qu'il ne soit pas nécessaire que le patient ait le choix de consentir ou pas à un test clinique, car le médecin a la responsabilité éthique d'évaluer le risque dans l'essai clinique (Amann, p.131). La vision de Hill était claire. Il fallait être capable, dans la randomisation clinique, d'offrir les meilleures évidences possibles tout en s'assurant que le risque pour le patient est le plus faible possible, au cas contraire, l'essai devrait être arrêté.

La naissance d'un nouveau paradigme

Evidence-based medicine a été considéré comme un nouveau paradigme. Ce « nouveau paradigme », au sens Kuhnien, se comprend comme un alignement « d'ensemble de croyances, de valeurs, de techniques, qui sont partagées au sein d'une communauté scientifique » (Sehon, Stanley, 2003, p.3). Il est également une « manière de voir le monde où les problèmes sont légitimement adressés et la gamme d'évidences de preuves capables de répondre à ses problèmes observés » (Guyatt et al., 1992, p.2420). De ce fait, contrairement à l'ancien paradigme, basé uniquement sur les observations non-systématiques,

l'expertise personnelle du praticien et les ouvrages médicaux, le nouveau paradigme se base sur les critères suivants (Guyatt et al., p.2421):

- 1. Les expériences cliniques:** qui permettent aux cliniciens et aux médecins de s'inspirer d'études de cas précédentes basées sur des évidences de « niveau supérieur » grâce à la technique de randomisation et des observations systématiques.
- 2. La recherche systématique:** qui consiste à la sélection des meilleures études évaluées selon le modèle de randomisation et devant servir aux praticiens.
- 3. Efficacité des traitements:** qui permet d'assigner aléatoirement les groupes, de traitement et de contrôle, aux différents traitements afin de relever l'efficacité de ces derniers.

Limites

Pearce et al. (2015) ainsi que Whitley et al. (2011) soulèvent quelques enjeux sapant la qualité *gold standard* de l'*evidence-based medicine* face à la dimension sociale des enjeux évalués. De ce fait, nous retenons trois défis majeurs en lien avec l'EBM :

- 1. La faible reconnaissance d'autres sources d'évidence et le problème d'extrapolation des résultats:** minimisent la connaissance du praticien, écartent le facteur culturel et contextuel dans les essais et de par ce fait, limitent la portée des évidences dans la vie réelle.
- 2. L'absence d'études qualitatives au sein de l'EBM:** sape les efforts d'explication causale des résultats dans un contexte multiculturel. Il est reconnu que les méthodes qualitatives sont plus aptes à produire des explications complètes des résultats dans les essais randomisés.
- 3. Le problème de biais:** résulterait des intérêts personnels de divers acteurs ayant des attentes face aux résultats des évaluation. Les biais peuvent survenir lors de la publication des résultats, ou durant le test clinique, lorsque les bailleurs détournent l'objet de la recherche sur des enjeux non prioritaires pour la santé publique par exemple.

En somme, bien que *l'evidence-based medicine* soit considérée comme la meilleure approche dans les essais cliniques grâce à la méthode de randomisation, le qualificatif *gold standard* de cette dernière n'est pas statique. Comme le soutient Quine avec sa métaphore de toile d'araignée (*Web of belief*), l'EBM et la science basic (*basic science*) sont complémentaires. Les « [deux] sont inter-reliés dans le réseau de connaissance, formé de périphéries (les faits observables) et d'un centre (les revendications théoriques) » (Sehon, Stanley, p.6). Les faits testent les théories, qui elles offrent une validité à l'existence des faits; de la même manière, la randomisation vise à tester des théories et hypothèses, qui elles lui donnent un sens.

2.2 L'evidence-based policy: le rapport entre évidence scientifique et expertise professionnel

Contexte

L'*evidence-base-policy* (EBP) a pour but de mieux informer les processus décisionnels et de justifier les actions des politiciens et praticiens en « essayant de trouver la plus fiable, la plus objective, la plus pertinente évidence disponible pour en tirer le meilleur bénéfice en fonction des contraintes en place » (Bédard, Ouimet, 2016, p.2). Ces contraintes font référence à la complexité de prise de décisions dans le secteur public : définition des priorités, tensions idéologiques, contraintes budgétaires, mais surtout l'agencement des évidences scientifiques avec les connaissances des experts; le rapport évidence et idéologie. Comme le définit Brian Head (2010) « *The EBP movement represents [...] a political rhetoric seeking to legitimate forms of decision-making that are alternatives to ideological or faith-based policy-making.* » (Head, p.77) L'EBP appelle à un virage où les prises de décisions ne sont plus juste le fruit des opinions et idéologies des politiciens, mais également celui des preuves probantes. Ce virage dans la pratique en politique se traduit d'une part, par un besoin de développer la recherche dans le but d'accéder à plus d'informations probantes, suivant la rigueur de l'évidence, et d'autre part, par le besoin de valoriser les politiques et programmes ayant fait leur preuve.

L'*evidence-based policy*, défend les mêmes valeurs que l'*evidence-based medicine* dans la quête des meilleures évidences disponibles. Avec l'EBP, il s'agit de former et motiver les praticiens politiques, les chercheurs, et les étudiants aux pratiques de revue systématique; de développer des réflexions sur la qualité des informations en procédant à une hiérarchisation des preuves; collecter des articles faisant preuves de corroboration entre les actions et les résultats afin d'informer les décideurs (Laurent et al., 2009, p.856). La Grande-Bretagne et les États-Unis sont deux cas considérables dans l'illustration de l'évolution de l'EBP.

En Grande-Bretagne, la création de la *Cochrane Collaboration* en 1993, une organisation œuvrant dans la recherche systématique médicale a principalement impacté le développement de revues systématiques. Le fondateur, Ian Chalmers (2005), affirme qu'il arrive que les décideurs et praticiens politiques fassent, par mégarde, plus de mal que de bien aux citoyens à travers les décisions qu'ils prennent sur des bases subjectives, et donc biaisées. Ainsi, Chalmers propose le recours aux données empiriques pertinentes, à travers la recherche rigoureuse, transparente, et des évaluations permettant une mise à jour des informations politiques (Chalmers, 2005, p.239). Cependant, lorsque Hammersley (2005) critique cette affirmation en soutenant que les évidences n'échappent ni aux interprétations des experts et ne prévaut non plus sur la gamme de connaissance de ses derniers (Hammersley, 2005, p.88), Chalmers reconnaît que les jugements des experts, lorsque conscient des besoins et des défis, sont essentiels dans l'interprétation des résultats d'évaluation d'un programme et dans la définition du sens des évidences obtenus (Chalmers, p.229). Cependant, la meilleure source d'information reste celle de la randomisation, exception faite lorsqu'on ne peut accéder à des évidences obtenues sous la base de celle-ci.

Le Premier ministre Tony Blair, à partir de 1997, avait fait de l'*evidence-based policy* une approche centrale dans le développement des capacités politiques et la réflexion sur les politiques (Head, 2010, p.79). Comme le montre cette affirmation venant du Cabinet du ministre: « *What counts is what works*

[...] policy decisions should be based on sound evidence. The raw ingredient of evidence is information [...] from variety of sources [as] expert knowledge; existing domestic and international research; existing statistics [...] including the internet» (Head, p.80). Celle-ci montre cette volonté de promouvoir les programmes et politiques qui fonctionnent afin qu'ils servent de base dans les prises de décisions. De surcroît, il importait autant de prendre en compte la variété de sources de connaissances disponibles car, « le système politique ne peut dépendre uniquement des données de recherche objectives » (Head, p.80). Depuis 2013, un réseau d'expertise, *What Works Network*, est mis en place au sein du gouvernement britannique afin de promouvoir la bonne gestion des fonds publics à travers les évidences disponibles¹.

Quant aux américains, entre 1930 et 1970, des évaluations randomisées avaient déjà commencées à prendre forme dans les politiques sociales. John Baron (2018) nous en brosse des cas où les ECR ont servi dans l'évaluation de programmes destinés : aux adolescents à risques de délinquance; à la libération d'accusés de crimes sans caution; à l'offre d'éducation préscolaire aux enfants issues de contexte économique précaires; à l'impact de l'assurance santé sur les dépenses en santé et l'état de santé des bénéficiaires, etc. (Baron, 2018, p.43). Ces évaluations randomisées, à l'époque, importaient plus par leur faisabilité dans le domaine sociale que par les évidences qu'elles offraient, car elles prouvaient que les programmes sociaux pouvaient également être soumis à des évaluations similaires à celles cliniques.

L'héritage de l'EBM se fait sentir lorsqu'en 1962, le congrès américain adopte une législation qui oblige le *Food and Drug Administration* (FDA) à démontrer l'efficacité de nouveaux médicaments à travers des évaluations randomisées (Baron, p.42). Les évidences obtenues lors de ces évaluations servaient aux décideurs politiques dans la répartition efficace des ressources publiques et privées vers des programmes ayant fait leur preuve. Comme mentionné par Daniel Fox (2005), « les politiciens se basaient sur les revues systématiques pour décider des allocations de ressources limitées et pour répondre aux

¹ Voir site internet: <https://www.gov.uk/guidance/what-works-network>. Visité le 24 juin 2020.

plaidoiries sur la couverture de certains médicaments, procédures et dispositifs médicaux » (Fox, 2005, p.1). Suite aux retours positifs de ces revues, l'état de Washington proclamait dans sa législation que: « la meilleure évidence scientifique disponible devrait guider les décisions portant sur la couverture en santé [...] et cette évidence devrait provenir de revues systématiques » (Fox, p.3).

Par ailleurs, les évaluations randomisées sur l'aide sociale et l'emploi dans les années 1980-1990, grandement financé par la fondation Ford, constituaient les premiers ECR à succès sur large échelle. Ces succès motivèrent le support de la maison blanche, sous les présidents H.W Bush et Bill Clinton, au *Health and Human Services* (HHS) dans le développement de nouveaux programmes d'aide sociale visant les travailleurs pauvres (Baron, 2018, p.44). Sous la présidence de Barack Obama, les décisions liées au financement des programmes dépendaient en grande partie des évidences obtenues sous le modèle des évaluations aléatoires. Dans le cadre de l'initiative gouvernementale, *Home Visiting*, portant sur l'éducation, la santé maternelle, infantile, et de la petite enfance, l'administration Obama appliquait les principes de *evidence-based* pour financer les programmes les plus prometteurs. Deux concepts importants motivaient ainsi le financement de ces programmes : **efficacité** et **évaluation rigoureuse**. Comme le montre les auteurs: « *Administration officials greatly increased the likelihood of good evaluations by requiring entities applying for federal grant funds to give detailed information on how they would evaluate their project* » (Haskins, Margolis, 2014, p.3.) Six principes d'évaluation avaient été adoptés afin de mesurer l'impact d'une recherche sur la politique. Ce sont : divulguer les résultats de recherches aux décideurs politiques; garantir que les décideurs et le corps administratif comprennent l'objet de la recherche menée; influencer la manière de voir des décideurs sur l'enjeu posé et ses solutions; utiliser la recherche pour guider les décideurs; faire en sorte que les conclusions de la recherche servent à décréter une nouvelle politique; et s'attendre à un impact positif de la nouvelle politique sur les citoyens (Haskins, Margolis, p.11). Sur le plan global, même si les évaluations rigoureuses n'ont pas eu d'impacts

majeurs, elles produisirent des mini-impacts, à petite échelle, qui contribuèrent à l'affectation efficace des fonds publics aux programmes ayant fait leur preuve.

Les limites de l'evidence-based policy en Politique

Pour se départir de la tentation de réduire le milieu politique qu'aux connaissances testées et approuvées à travers des expériences randomisées, il est important de relever la nature hybride de la production des connaissances (Bédard, Ouimet, 2016, p.3); c'est-à-dire une production basée sur l'apport de la science, donc de l'évidence, et de l'expertise, dont celle des praticiens et décideurs politiques (Pearce, Raman, 2014, p.388). Ainsi, il apparaît là encore évident que les données quantitatives, à elles seules, ne suffisent à informer les politiciens. Il est nécessaire de prendre en compte d'autres sources de connaissances, dont : l'expertise des politiciens et les rapports de pouvoir autour de la structure politique, afin de rendre concrets les résultats des évaluations. Trois limites sont ainsi soulevées ici par Pearce et Raman (2014, p.390-391 et 395-397) :

- 1. Nature variée de l'évidence:** l'évidence prend forme dans un contexte où des hypothèses ont été développées et testées. Vouloir appliquer cette évidence dans un autre contexte oblige à l'identification des limites de celle-ci face au nouvel enjeu adressé. Ainsi, la portée de l'évidence faiblit face à la nouvelle politique recherchée, qui est sujette à des interprétations variées.
- 2. L'interprétation de l'évidence:** l'EBP sollicite l'avis de multiples experts pour interprétation, évaluation, applicabilité. Cette façon d'aborder les enjeux en politique nécessite la prise en compte des différences entre les sciences, mais aussi entre les sciences et le sens public (interprétation publique). Ainsi, écarter les autres types de connaissances c'est appauvrir le processus politique.
- 3. Le rôle des institutions:** face à l'environnement politique qui est sujet à des tentatives de corruptions, il est important que les institutions soient fiables et dignes de confiance dans la gestion des multiples

évidences, leurs interprétations et leur encadrement, tout en reconnaissant les limites du positivisme scientifique dans certains contextes (comme la traduction de l'évidence en politique).

Quant à Laurent et al. (2009, p.860-868), sur les limites de l'EBP dans leurs études sur l'agriculture et protection de la biodiversité, ils soulèvent des difficultés liées à la recherche et l'accessibilité des preuves rigoureusement évaluées; des enjeux techniques et financiers liés au développement d'une revue systématique; des problèmes d'hierarchisation des preuves (scientifiques versus qualitatives par exemple); et l'existence de différents types de preuves (d'existence, de causalité, d'efficacité, d'innocuité). À cela, on peut souligner la remarque pertinente de Donald Campbell lorsqu'il affirme que la méthode quasi-expérimentale (où les variables sont manipulables) dans l'évaluation des politiques, est plus propice que celle expérimentale lorsque le contexte nécessite la prise en compte de données supplémentaires (Campbell,1969, p.411).

Pour les politiques de développement qui sont basées sur diverses sources de connaissances, la prise en compte de l'opposition entre différents niveaux de preuves est centrale. Voyons à présent les enjeux de *l'evidence-based* en développement international afin d'approfondir cette assertion.

2.3 Les évaluations randomisées en développement international

Pour comprendre comment les essais expérimentaux randomisés se sont imposés, notons deux facteurs clés. Premièrement, l'échec du consensus de Washington, dans les années 1980-1990, qui préconisait entre autres des « plans d'ajustement structurels » basés sur la libéralisation des échanges économiques, la déréglementation des marchés et la privatisation de nombreux services sociaux, qui relevaient des fonctions régaliennes de l'État (Jatteau, 2013, p.10). Cette phase fut marquée par la prise de conscience des limites des modèles libéraux, surtout en Amérique Latine et en Afrique, où les politiques préconisées n'avaient pu produire les résultats escomptés (la croissance économique, le développement

de capacité humaine, surtout, sortie de la pauvreté). Cette période fut celle de « la décennie perdue » comme la nommait le Programme des Nations unies pour le développement (PNUD) (Jatteau, p.10).

Deuxièmement, la crise économique de 2007 affecta les capacités des États bailleurs à soutenir financièrement les programmes de développement. Il était primordial pour ces acteurs de répondre aux enjeux au niveau interne: déficit, dette publique, chômage (Jatteau, p.10). À ces facteurs, s'y ajoutent les résultats décevants des évaluations de programmes qui manquaient de rigueur méthodologique et de preuves montrant le lien entre les résultats et les aides financières (Picciotto, 2012). Face à cette crise sur l'efficacité de l'aide, le besoin d'une nouvelle approche pouvant répondre aux besoins *d'evidence-based* dans les programmes de développement devient évident. De là, des opportunités pour des économistes s'ouvrent. Le J-PAL, sous Banerjee, Duflo et leurs partenaires, saisira l'opportunité pour promouvoir les ECR dans la sphère du développement international (Picciotto, 2012, p.215).

Dans cet élan, de Souza Leão et Eyal (2019), nous offrent une explication détaillée de la différence entre la première vague et la deuxième vague d'évaluation des programmes de développement basée sur les ECR, et montrent comment la deuxième vague a réussi à survivre à la résistance politique. Deux facteurs essentiels pour nous éclairer ici: **le réseau d'expert** responsable des évaluations et le **type de programme/projet évalué**. Concernant le réseau d'experts, la différence entre les deux vagues s'observe sur trois critères: **l'affiliation disciplinaire des experts et leur relation avec les acteurs étatiques et non-étatiques; l'origine des fonds de financement des programmes; et les partenaires de recherche sur le terrain**. Dans les années 1980-1990, les experts étaient plus issus de domaines tels que: la sociologie, la psychologie, la démographie, la science statistique, la physique, l'épidémiologie, la politique. Ils étaient affiliés à des agences internationales et gouvernementales, soit en tant que consultants, soit en tant qu'employés et chercheurs. Cette relation de proximité avec les décideurs politiques, mais aussi les bailleurs, comme l'Agence américaine pour le développement international

(USAID), minait l'autonomie d'action et de gestion des fonds de ces évaluateurs de première vague. De ce fait, les capacités réelles pour offrir des évaluations non biaisées avec des évidences plus convaincantes étaient faibles. Concernant le financement, l'USAID, les fondations Ford et Rockefeller, étaient les plus grands bailleurs, en plus de l'apport de la société civile (université, *think-tank*, organisation non-gouvernementales) et des gouvernements locaux qui pouvaient être des partenaires sur le terrain. En outre, les programmes évalués durant cette vague s'étendaient au niveau national en général et s'allongeaient sur de longs laps de temps, soit de 1 à 8 ans. Ce qui rendait presque impossible la mise en application des ECR. De plus, les pays États bénéficiaires avaient des capacités institutionnelles limitées qui minaient l'accès à des données statistiques probantes (de Souza Leão et Eyal, p.395-397).

Par contre, avec la nouvelle vague des évaluations randomisées, illustrée par la montée du J-PAL, début 2000, se dessine un paradigme de ce à quoi doit ressembler le nouveau consensus sur le développement. Non plus celui d'un antagonisme entre les solutions à adopter, soit financer le développement, comme le préconise Jeffrey Sachs, ou laisser les pays pauvres prendre en main leur problèmes, comme défend sciemment William Easterly (Banerjee, Duflo, 2012, p.20-21). En effet, les randomistes remettent en question l'idéologie, l'ignorance et l'inertie (les trois i) (Banerjee, Duflo, p.39) et proposent un consensus sur la méthode à adopter. La randomisation devient ainsi la solution à même de combler l'insatisfaction des bailleurs face aux attentes des résultats escomptés de leur financement (Jatteau, 2013, p.11). Ainsi, cette seconde vague se distingue à travers une cohorte d'experts en économie (80%), détenant en majorité un Doctorat (93%), des jeunes universitaires avec des bases en politiques publiques et en développement international (de Souza Leão et Eyal, p.393-394). Le laboratoire J-PAL à travers son approche, a su attirer l'attention des organisations privées et fondations, comme *Bill et Melinda Gates*, les *fondations McArthur*, et celui de la Banque mondiale, qui est d'ailleurs un des bailleurs importants du J-PAL (de Souza Leão et Eyal p.396). Ici, le mécanisme de fonctionnement est autonome,

car les fonds perçus ne dictent pas nécessairement les types de programmes qui seront évalués et donc, moins de contrôle s'applique de la part des acteurs étatiques, privés et publics, sur les évaluateurs. De plus, les partenaires de recherche du J-PAL comprennent des acteurs locaux², qui ensemble, représentent environ 75% de partenaires d'implémentation des programmes d'évaluations. Ce qui facilite la collecte des données et pallie la résistance politique aux essais randomisés; J-PAL se définit d'ailleurs comme apolitique dans son approche. Cette structure de J-PAL a permis l'attraction de bailleurs importants qui peuvent enfin voir le fruit de leur financement. De plus, les ECR se font sur courte durée, un mois à une année en moyenne (Leão et Eyal, p.398), avec des échantillons assez larges pour être représentatifs d'une population donnée et facilement mesurable.

Ainsi, les ECR se sont vu s'élargir dans le champ du développement grâce aux économistes randomistes. Les évidences qu'ils produisaient à travers leurs évaluations transforma la crainte des bailleurs de fonds et dans la même lancée, le mécanisme d'évaluation des programmes de développement. Les ECR ont pu renouer le lien entre le développement et son financement à travers l'implémentation d'objectifs clairs et mesurables. Identifier les problèmes concrets contribuant à une gestion responsable et efficiente des ressources comme le montre les auteurs Baqir, Banerjee et Yaya dans leur livre *Better Spending for Localizing Global Sustainable Development Goals* (2020).

Dans la section qui suit, nous aborderons de manière claire et succincte les forces et faiblesses des ECR, aussi bien que leurs limites incitant à la prise en compte d'autres méthodes pour les renforcer. Nous verrons également la place des méthodes qualitatives au sein des évaluations randomisée contrôlées.

² Voir site internet: <https://www.povertyactionlab.org/j-pal-offices>. Visité le 24 juin 2020.

3. Le cadre d'analyse des essais contrôlés randomisés

3.1 Les forces et faiblesses de la randomisation

3.1.1 Les forces

Le contrôle des biais

Le principe d'un essai contrôlé randomisé est de comparer deux groupes, aléatoirement sélectionner dans une population donnée, puis, aléatoirement, assigné à un groupe de traitement et un groupe de contrôle (certaines études peuvent avoir un groupe de contrôle pur en plus). La comparaison entre les deux groupes fournira un résultat qui nous indiquera si le traitement que nous voulons tester est efficace ou non. L'aspect aléatoire ici permet d'éviter l'introduction de biais. Ces biais peuvent être dus à différents facteurs (Bickman, Reich, 2017), comme lorsque les évaluateurs de l'étude ont connaissance du traitement ou de la particularité des groupes étudiés (*unmasked assignment*); ou lorsque les groupes soumis à l'expérimentation découvrent la nature du traitement qu'ils reçoivent et décident de se comporter favorablement (ou négligemment) par rapport au traitement; ou encore lorsque les évaluateurs sont soumis à la pression de preuves et de publications (de la part des bailleurs et de la communauté scientifique). Cependant, comparativement à d'autres méthodes (Bouffard, 2012; Laurent et al., 2009), la randomisation à le plus haut niveau de validité de résultats internes grâce à sa méthode de randomisation. Même s'il est reconnu que les biais peuvent être introduits durant le design du programme, la mise en œuvre, ou encore l'étape de l'analyse des résultats, le fait de s'assurer que les groupes soumis à l'évaluation sont aléatoirement équivalents, permet d'éliminer, ou du moins minimiser, le facteur biais (Worall, 2007, p.1007). Cet élément fait des ECR les plus propices à obtenir la preuve d'efficacité d'un traitement.

L'inférence causale: validité interne

Autre élément important des ECR, est que leur structure méthodologique (design) nous permet de déterminer la relation de causalité dans une expérience. Cette relation se situe entre l'élément dont on veut

tester l'efficacité et le résultat observé; autrement dit c'est la relation entre le traitement et l'effet. Lorsqu'une évaluation ressort l'effet de causalité d'une politique, les décideurs politiques seront plus aptes à tester cette politique et ainsi déboursier des fonds pour son application. Ceci entre en compte avec l'idée d'une gestion efficace des dépenses, mais aussi la démonstration du principe même de validité interne très ancré dans les essais expérimentaux. C'est d'ailleurs une des raisons pour lesquelles le J-PAL est aujourd'hui perçu comme un acteur important sur la scène internationale. Les évidences probantes et vérifiables qu'il produit lui manifeste le soutien des grandes fondations et institutions³.

3.1.2 Les faiblesses

L'expansion des résultats : validité externe

On parle de validité externe lorsque les résultats tirés d'une population peuvent être transposés sur une autre population ayant, à peu près, les mêmes caractéristiques. En médecine et en science sociale, la généralisation des résultats reste un des enjeux les plus cités par les critiques, mais également reconnu par les protagonistes de la méthode. Dans le contexte politique, comme le mentionne Arthur Jatteau, les gouvernements sont autant intéressés par les effets à moyen terme d'une politique que par la généralisation de ses effets (Jatteau, 2019, p.96). La généralisation dépend de la capacité des chercheurs à inférer une explication causale solide, relevant les facteurs ayant menés au succès du traitement. Le programme *Progesa*, au Mexique, n'aurait, par exemple, pas pu être autant étendu si certaines conditions n'étaient pas en place dans les autres pays comme: une demande importante en santé et en éducation qui nécessite une capacité d'offre importante; un support politique; l'accès à des écoles et centres de santé pour les bénéficiaires, etc. (Deaton, 2010, p.449). D'où l'importance de prendre en compte le contexte des études d'évaluations et l'apport d'autres méthodes (et même l'expertise des politiciens), comme celles mixtes, dans l'identification des facteurs causals potentiels pouvant affecter les effets de l'étude.

³ Voir site internet: (<https://www.povertyactionlab.org/partners>). Visité le 24 juin 2020.

L'explication des facteurs causals

Comme référé précédemment, l'explication causale dépend des facteurs relevés qui ont permis d'inférer une relation causale. Cette explication permet de comprendre quel(s) élément(s) dans le traitement a causé un changement sur une variable. En isolant cet élément, on saura à même de réfléchir aux moyens de, soit améliorer le traitement (ou la politique), soit imiter la même expérience dans un contexte différent. Spécifiquement, il s'agit d'aller au-delà de la relation causale, en montrant quel élément de cette relation, peut être provoqué de nouveau pour affecter le changement (Bouffard, 2012, p.84). Si l'explication des facteurs causals « contribue à développer et tester des hypothèses ou théories concernant la structure des effets, des causes [...], essentiels au transfert de la relation causale » (Campbell et al., 2002, p.25), alors les méthodes qualitatives ou celles mixtes, intégrés dans le design des ECR, pourraient être des moyens efficaces dans la détection et l'explication de ces hypothèses.

La théorie dans les évaluations contrôlées randomisées

Certains penseurs (Deaton, Cartwright, 2017; Rodrik, 2008; Kyle, 2017; Stame, 2004; Favereau, 2016) ont porté, certes de manière différente, l'attention sur la nécessité d'adopter une approche théorique, *ex-ante*, qui facilitera la généralisation des résultats au sein des ECR. Comme le mentionne Nancy Cartwright, une évaluation sans base théorique a moins de capacité à informer sur le mécanisme ayant mené à obtenir un impact et de ce fait, faille à contribuer à la connaissance général (Deaton, 2010, p.450). Bien que les évaluateurs, comme ceux du J-PAL, adoptent une théorie du changement (*theory of change*), cette théorie est basée sur des assomptions permettant d'esquisser le processus d'atteinte des impacts souhaités. Cependant, ces assomptions, au-delà d'esquisser comment atteindre un résultat, doivent viser le développement d'un cadre permettant de détecter les mêmes facteurs dans divers contextes d'étude (Vogel, 2012, p.26).

3.2 Avantage et limite de la randomisation

3.2.1 Avantage

Le contrefactuel

Les essais aléatoires ont l'avantage, grâce à leur design de randomisation, de pouvoir expliquer ce qui serait arrivé si un traitement n'avait pas été mis en œuvre ou testé. Ceci est le contrefactuel. L'analyse factuelle explique la relation entre le traitement et le résultat qui permet de savoir exactement que le traitement a fonctionné; mais elle ne suffit pas pour autant à établir l'inférence causale ou la relation causale. Pour réaliser parfaitement cette dernière, on a besoin d'un groupe de comparaison, qui ainsi, permettra d'établir le contrefactuel. On peut traduire cette explication par la formule suivante :

$$\alpha = (Y | P = 1) - (Y | P = 0)$$

L'effet causal (α) d'un programme (P) sur un résultat (Y) est la différence entre le résultat (Y) obtenu avec le programme (P = 1) et le même résultat (Y) obtenu sans le programme (P = 0) (Gertler et al., 2011, p.34). Autrement dit, Y | P = 1 est le groupe de traitement pour lequel on sait que le résultat Y a produit un impact. Mais pour vérifier la validité de cet impact, on le compare à Y | P = 0, le groupe de contrôle, qui nous permet de savoir ce qui arrive sans administration du traitement. Cette capacité de comparabilité, en temps réel, dans les expériences aléatoires permet de vérifier la validité des résultats d'un traitement.

3.2.2 Limite

La problématique de la « boîte noire » ou black Box

L'enjeu de la « boîte noire » fait état du besoin d'explicitier les phénomènes se produisant entre les intrants et les résultats. Ceci nécessite une évaluation de processus (opposé à celle d'impact comme la randomisation) à travers laquelle les défis de généralisation des résultats et de recommandations politiques se posent concrètement (le pourquoi et le comment) (Gertler et al., 2011, p.232). Des membres du J-PAL ont récemment mené une étude en Indonésie sur l'impact de la participation communautaire dans le

renforcement de l'offre de services de santé maternelle et néonatale. En optant pour une approche mixte de collecte des données, l'analyse du contexte sociale et de ses effets sur les résultats, ainsi que l'explication des dynamiques de pouvoirs au sein de cette communauté, ont pu contribuer à une meilleure compréhension des enjeux opérationnels, contenus dans la « boîte noire » (Levy et al. 2020). Détecter les éléments ayant une influence majeure sur la participation permet de réactiver ces mêmes éléments, et ainsi améliorer la portée de la validité externe.

3.3 Les méthodes qualitatives dans les essais contrôlés randomisés

Une littérature assez large et diversifiée (Spillane et al., 2010; Bamberger et al. 2010; Flemming et al., 2008; Chatterji, 2005; Murtagh et al., 2007; Stern et al., 2012), décrit l'introduction des méthodes qualitatives dans les essais contrôlés randomisés en mettant en avant leur potentiel dans l'explication des résultats des évaluations. La méthodologie qualitative, dans les évaluations de programmes se centre sur les questions répondants au « comment » et « pourquoi » de tel résultat (ACYF, 2016, p.4). Cette méthodologie facilite le **développement, la corroboration, et la correction d'une théorie du changement; la collecte de données sensibles et leur mesure; l'implémentation du programme d'évaluation; exploration du contenu de la « boîte noire »; l'analyse et le report des résultats de la recherche** (ACYF, p.5-6).

Par ailleurs, on peut noter deux courants face à l'introduction des méthodes qualitatives au sein des essais randomisés. Le premier vise l'amélioration du protocole des essais randomisés (méthode), en misant sur l'explication détaillée des facteurs affectant la réalisation de l'expérience. Le second vise l'amélioration de l'implémentation (intervention/traitement) en misant sur étude du contexte de l'évaluation (Mannell, Davis, 2019). Bien que ces auteurs établissent une distinction entre ces deux courants, on pourrait également établir une corrélation, dans le sens où la question d'intervention, donc

l'enjeu à évaluer, guide le choix d'évaluation, à savoir la meilleure méthode à adopter pour efficacement répondre à la question d'évaluation (Gertler et al., 2011; Bamberger et al., 2010). En outre, cette séparation que fait les auteurs permet de saisir le rapport entre méthode quantitative (validité interne) et qualitative (validité externe). En effet, dans le premier courant, les tenants cherchent à améliorer la rigueur de l'évaluation en renforçant l'évidence des résultats, tandis que dans le second, il s'agit de promouvoir l'expansion de la même expérience en misant sur les données qualitatives.

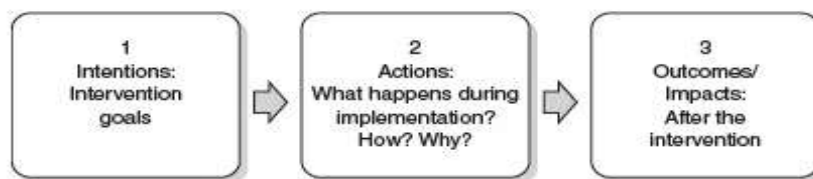
Cette dichotomie entre méthodes fragilise la valeur de la MQ, vu comme moins scientifique. Les auteurs proposent ainsi une approche innovante, qui ne limite plus la MQ aux techniques d'entrevues et de groupes de discussion, mais qui leur permet d'également répondre aux enjeux complexes sur le terrain. Cela se fera en adoptant des méthodes mixtes qui tenteront de répondre aux questions de type: quelles interventions marchent, pour qui, et sous quelles circonstances (Mannell, Davis, p.11). Cette vision des évaluations peut se comprendre selon les réalistes des ECR qui voient les relations causales comme un phénomène complexe nécessitant des approches appropriées. Certains critiquent voient plutôt cette vision comme minant la nature des ECR (à cause de la subjectivité de la MQ) (Van Belle et al., 2016). Sans entrer ici dans le débat théorique entre ces différentes visions, les auteurs montrent, premièrement, que les méthodes qualitatives peuvent contribuer à mieux interpréter les résultats des ECR et expliquer les changements observés; et deuxièmement, qu'une valorisation des apports qualitatives dans les ECR pourrait briser les opinions biaisées sur la qualité des résultats de la MQ (Mannell, Davis, 2019). Cela, en soutenant la publication des résultats des différentes méthodes dans un même article, car nombre d'articles publient séparément les résultats des enquêtes (O'Cathain et al., 2013).

3.3.1 Le rapport entre la théorie du changement et la « boîte noire »

La théorie du changement est le modèle qui établit ce qui doit être fait pour atteindre un résultat et quelles conditions sont nécessaires à l'atteinte de ce résultat (Vogel, 2012, p.9). Cette théorie, développée

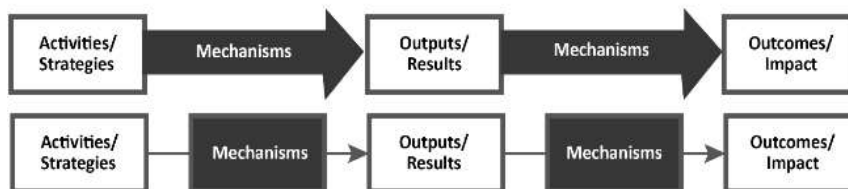
à partir d'outils qualitatifs, guide le projet d'évaluation depuis la conception, car il se définit bien avant la mise en œuvre de ce dernier, et impacte le choix de la question et du design d'évaluation (Gertler et al., 2011; Rallis, 2017; White, 2013). La théorie permet de questionner le contexte; planifier les changements à long-terme; déceler les mécanismes affectant les résultats; développer des assumptions ou hypothèses qui décrivent les étapes plausibles du changement; définir les bénéficiaires et acteurs du programme ainsi que les relations de pouvoir; etc. (Vogel, p.15). La théorie du changement incite à relever les mécanismes dans la « boîte noire » qui dissimulent la relation entre le traitement et le résultat; ou entre les activités, résultats et impacts dans le cadre des évaluations de projets (Dhillon, Vaca, 2018, p.79)

Figure 1: La théorie du changement



Source : Rallis (2017), p.3.

Figure 2 : Les mécanismes dans une théorie du changement



Source : Tiré de Dhillon, Vaca (2018), p.79.

La boîte 1 et la boîte 3, de la figure, 1 se réfèrent à l'explication de l'inférence causale. Les essais aléatoires tirent leur force de cette relation, X produit Y (validité interne), en offrant la meilleure évidence possible (White, 2013, p.63). Cependant, lorsqu'on cherche à comprendre ce qui se passe entre le traitement et le résultat, pourquoi et comment X produit Y (explication causale), dans la boîte 2, nous devons nous tourner vers des méthodes permettant de questionner les phénomènes en cours dans cette

phase. L'explication causale, dans la figure 2, prend source dans les boîtes noires reliant actions et résultats.

3.3.2 Les méthodes mixtes : entre le quantitatif et le qualitatif

Parmi les modèles d'évaluations mixtes (Greene et al., 1989, p.259; Stern et al., 2012, p.32; Aldebert, Rouzies, 2015, p.47) permettant de répondre aux limites des données quantitatives, nous avons ci-après:

- 1. La triangulation:** c'est l'utilisation simultanée de plusieurs méthodes différentes dans le but de contrôler les possibles biais en espérant renforcer la validité des résultats. L'objet est de pallier la faiblesse d'une méthode en contrebalançant par une autre méthode; faire de la corroboration. Le piège est qu'il est important d'évaluer la validité de chaque résultat afin de détecter des potentielles similarités de biais entre méthodes, car certains biais peuvent converger.
- 2. La complémentarité:** c'est l'utilisation simultanée ou séquentielle d'une méthode qualitative et/ou quantitative dans le but de mesurer des variables différentes ou similaires afin de comprendre leur impact sur le résultat observé. Ce modèle de design renforce la compréhension des résultats et l'interprétation des évidences et s'appuyant sur les forces de chaque méthode.
- 3. Exploratoire:** c'est l'utilisation, en général, de la méthode qualitative pour évaluer l'évolution du programme, et de la méthode quantitative pour évaluer les résultats du programme. L'objectif de ce design est de renforcer la portée et la généralisation des résultats.

3.3.3 Les bénéfices des méthodes qualitatives

Bien que l'approche qualitative soit critiquée sur sa nature subjective, son apport dans les expériences aléatoires sont pourtant assez bien documenté (Gertler et al., 2011; Bamberger et al., 2010; Stern et al., 2012; O'Cathain et al. 2014; Murtagh et al., 2007; Davis et al., 2019). Les techniques qualitatives peuvent contribuer aux ECR des façons suivantes (Bamberger et al., p.6-7):

- **Évaluation de processus:** permet de relever des données pouvant expliquer la nature des résultats, un problème de design ou d'implémentation. Ceci peut se faire à travers des techniques d'observations, d'entrevues, de groupes de discussion qui permettront de vérifier la théorie de base, les assomptions des chercheurs, ou encore les évidences littéraires supposées clarifier la relation entre cause et effet.
- **L'accès à des indicateurs additionnels:** les techniques qualitatives permettent de collecter des données dont les indicateurs pourront faciliter une analyse multidimensionnelle. Par exemple, il est facile de comprendre la décision de participation des ménages à un programme d'assurance santé si on détecte les facteurs affectant la décision du chef de ménage par exemple: croyance, culture, système social, revenu, accès à d'autres filets sociaux (aide familial), etc.
- **Le problème d'hétérogénéité:** les méthodes qualitatives sont utiles pour relever les variations de l'effet de traitement sur les groupes étudiés. L'identification des caractéristiques particulières des bénéficiaires permettra de déterminer les effets marginaux, sur le groupe étudié, des effets moyens du traitement, qui porte sur la population visée.
- **La prise en compte du contexte local:** les méthodes quantitatives sont inflexibles lorsqu'il s'agit de prendre en compte des phénomènes contextuels. L'approche qualitative peut compenser cette faiblesse en collectant des données offrant un aperçu des facteurs reliant causes et effets.

O'Cathain et ses collègues observent 5 catégories et 22 sous-catégories sur lesquelles la recherche qualitative agit au sein des essais randomisés en santé. Ces auteurs montrent que la MQ peut aider à relever: les bénéfices et valeurs de l'intervention; la faisabilité et l'acceptabilité de l'intervention en pratique; les mécanismes sous-jacents la théorie du développement (étape de planification); améliorer la rétention; promouvoir la diversité chez les répondants; adapter

l'intervention au contexte local; respecter les principes éthiques; évaluer les variation dans les résultats; tester l'efficacité des outils de mesures (étape des mesure) (O'Cathain et al., 2014, p.15).

Quant à Davis et al. (2019), ils proposent un cadre explorant l'apport de la MQ au-delà des groupes de discussion et des entrevues. Ce cadre explicite comment les techniques qualitatives peuvent contribuer à améliorer six catégories importantes dans le report d'une évaluation randomisée suivant le modèle CONSORT⁴. Cette approche pourrait affecter la nature des ECR en optant pour la participation des répondants dans le design des ECR; appelant ainsi à une tournure vers des « études », au lieu des « essais », qui incitent à faire le choix de la meilleure méthode pour répondre à la question, plutôt que de supposer quelle méthode mène à la meilleure évidence, selon une hiérarchie établie (Davis et al., 2019, p.13).

Tableau 1: Apport des méthodes qualitatives au sein des évaluations contrôlés randomisées

Catégories adaptées à partir de CONSORT	Défis méthodologiques	Menaces au design ou aux résultats	Solutions apportées par la recherche qualitative	Exemples de méthodes qualitatives envisageables
Contexte	Assurer la pertinence du problème dans le contexte	Affecte l'applicabilité, l'acceptabilité, la transférabilité et la durabilité de potentiels résultats positifs	Identifier les facteurs sociaux, culturels, économiques, politiques et de santé qui pourraient affecter	Observation directe approfondie dans le cadre d'une évaluation ethnographique rapide ou d'une enquête à grande échelle
Design de l'intervention et conformité	Assurer l'acceptabilité culturelle and la faisabilité pratique de l'intervention	Affecte l'adhésion et hausse le nombre d'abandons	Adapter l'intervention afin d'hausser la rétention et l'adhésion	Méthodes d'agenda – via des messages vocaux interactifs ou des SMS
Recrutement et enrôlement	Garantir la représentativité de l'échantillon et des méthodes de recrutement efficient	Risques liés à la taille de l'échantillon afin de détecter des effets significatifs	Déterminer la meilleure méthode possible de recrutement afin d'atteindre la population cible	Cartographie communautaire

⁴ Voir site internet: <http://www.consort-statement.org>. Visité le 24 juin 2020.

Catégories adaptées à partir de CONSORT	Défis méthodologiques	Menaces au design ou aux résultats	Solution apportée par la recherche qualitative	Exemples de méthodes qualitatives envisageables
Randomisation et allocation	Garantir une randomisation équilibrée (égale)	Réduit la comparabilité entre les groupes	Identifier des facteurs contextuels qui pourraient affecter l'effet de l'intervention and réduire les différences entre les groupes évalués	Observation/randomisation publique
Participants et suivi	Minimiser le nombre de participants quittant l'étude	Induit une perte inégale de participants entre les groupes, ce qui affecte l'inférence causale	Pour prévenir ou comprendre les raisons de perte dans le suivi et pour améliorer les stratégies de rétention	Tenue d'un agenda, utilisation des messages vocaux interactifs par SMS ou WhatsApp soutien des pairs pour l'adhésion
Collecte des données	Maintenir une objectivité durant le processus de collecte des données	Menace la validité des données collectées et/ou mesurées afin de répondre à l'objectif de l'étude	Évaluer et valider le processus de collecte des données	Recours aux groupes de discussion utilisant des cartes de mémorisation ou des images; des entretiens en binôme et jeux de rôle
Analyses et résultats	Identifier les mécanismes sous-tendant l'effet de l'intervention	Limite une discussion éclairée des résultats (positifs ou négatifs)	Triangler les résultats quantitatifs et identifier les données contextuelles qui pourraient affecter les résultats	Méthode d'analyse participative

Source: adapté de Davis et al. (2019), p.8.

Somme toute, les méthodes qualitatives, introduites dans les essais contrôlés randomisés, aident à valider des résultats quantitatifs, mais aussi la portée externe des évidences en palliant la problématique de la « boîte noire ». Celle-ci nécessite une meilleure compréhension du contexte de l'étude et de l'effet de ce dernier sur les résultats.

Explorons à présent nos deux études de cas et la description de leur design pour mieux comprendre comment la méthode qualitative affecte la portée des résultats des essais randomisés, dans un cadre précis.

4. Études de cas sur la couverture sanitaire en Indonésie et au Nicaragua

Dans cette section, nous présentons deux formes de design d'évaluation expérimentale portant sur la couverture en santé chez les travailleurs du secteur informel en Indonésie et au Nicaragua. En Indonésie, le design de l'expérience s'appuie largement sur des méthodes quantitatives, alors qu'au Nicaragua, des enquêtes qualitatives ont servi à corroborer les résultats quantitatifs. Le but est de comprendre comment les deux types de design affectent la portée des informations censées clarifier les causes d'enrôlement ou de non-enrôlement à l'assurance santé chez les groupes visés. La première section porte sur une présentation du cas d'étude sur l'Indonésie, en décrivant le contexte, la méthodologie (design) d'évaluation et les résultats obtenus. La seconde partie porte sur le Nicaragua pour lequel on présente également les mêmes éléments que l'Indonésie.

4.1 Étude de cas sur l'Indonésie

4.1.1 Contexte

L'Indonésie est le quatrième pays le plus peuplé au monde avec une population de 268.1 millions d'habitants en 2019 (Statistik Indonesia, 2020, p.2). Son produit intérieur brut (PIB) s'élevait à plus d'un milliard en 2018 pour un revenu national brut par habitant de 3840 dollars US (Banque mondiale, 2018). Son indice de Gini, entre 2010-2017, était de 38.1 (0.381), avec un indice de développement humain (IDH) de 0,707 en 2018, le classant ainsi au 111^e rang de la tranche des pays à développement humain élevé (Rapport PNUD, 2019). Le taux de pauvreté était de 9.4 % en 2019 (Banque mondiale), soit 25.2 millions de pauvres suivant le seuil de pauvreté national de 82 cents par jour, ou 302, 735 Rupiah (Rp) (25 dollars US) par mois et par personne (Aji, 2015, p.3). L'institut national indonésien définit la pauvreté sur deux bases : la consommation alimentaire journalière, soit 2,100 calories, et la satisfaction des besoins en habillement, en éducation, en santé et autres services non-alimentaires (Statistik Indonesia, p.158). L'Indonésie comprend 300 groupes ethniques et 720 dialectes (Augustina et al., 2019, p.75) répartis sur

les 16 056 îles (Statistik Indonesia, p.10). Cette diversité géographique et sociale constitue un immense défi sur le plan du système de santé et de l'élargissement de la couverture sanitaire universelle.

Le programme national d'assurance sociale indonésien, *Jaminan Kesehatan Nasional* (JKN), géré par l'agence de sécurité sociale *Badan Penyelenggara Jaminan Sosial* (BPJS), est l'un des plus larges systèmes de couverture dans le monde avec 221 millions de bénéficiaires, soit 83 % de la population en 2019 (Prabhakaran et al., 2019, p.39). Officiellement entré en vigueur en 2014, le programme englobe les premiers systèmes de couverture sociale du pays, à savoir : *Askes* (fonctionnaires d'État), *Jamsostek* (employés du privé), *Jamkesmas* (pauvres et « moyennement pauvres ») et *Jamkesda* (provinciaux) (Prabhakaran et al., p.10; Ly, 2018, p.2). KJN est fondé sur deux modèles de paiement: l'un non-contributif, où les bénéficiaires sont soutenus par l'État et les entreprises, et l'autre contributifs, où les bénéficiaires paient eux-mêmes les primes d'assurance (Dartanto et al., 2016, p.323). Avant JKN, les travailleurs du secteur informel étaient couverts par *Jamskeskmas* sur une base volontaire et contributive. Seul les 40% en dessous du seuil de pauvreté (soit 82 cents/jour/personne) ont droit à des assurances gratuites. Le reste, dit non-pauvres, se prend en charge.

Les travailleurs de l'informel: non-salariés, entrepreneurs, travailleurs occasionnels en agriculture, pêche, services (Ly, 2018, p.24) sont tenus par le second modèle de paiement sous le programme KJN *Mandiri*. La majorité de ces travailleurs de l'informel bénéficient de trois modèles de paiement, en fonction du service voulu, pour accéder à la couverture santé (Prabhakaran et al., 2019, p.34; Ly, p.7-8):

- 25,500 rupiahs (2 dollars US) par personne/mois pour un traitement de classe 3 (5 lits et plus)
- 42,500 rupiahs (3 dollars US) par personne/mois pour un traitement de classe 2 (2 à 5 lits)
- 59,500 rupiahs (4.5 dollars US) par personne/mois pour un traitement de classe 1 (1 à 2 lits)

Avec un ratio de 60% de la main d'œuvre active, les travailleurs du secteur informel, nécessaire à l'achèvement de la couverture universelle en Indonésie, comptaient environ 160.9 millions de personnes en 2014 (Dartanto et al., 2016, p.318). Bien que le taux de travailleurs du secteur informel ait baissé entre 2006 et 2016 passant de 68.9% à 57.6% (OCDE, 2019, p.15), c'est 25 millions étaient enrôlés en 2017 (Prabhakaran et al., 2019, p.39); ce qui signifie qu'environ 135 millions sont sans couverture. Souvent, ceux qui s'enregistrent au programme finissent par ne plus continuer à payer leur couverture. De plus, ce sont les plus vulnérables aux maladies qui s'inscrivent. Ce qui cause deux problèmes: la sélection adverse, seuls les malades sont attirés par l'offre, et la viabilité financière du programme, l'État paie plus de subventions gratuites qu'il ne reçoit en primes de couverture (OCDE, p.19; Prabhakaran et al., p.71)).

S'inscrivant ainsi dans l'objectif d'expansion de la couverture universelle en Indonésie d'ici 2034, l'étude du J-PAL (Banerjee et al., 2019) vise à fournir des évidences sur les facteurs affectant l'enrôlement à la couverture sanitaire des ménages travaillant dans le secteur informel et les effets de ces facteurs sur la solvabilité financière de l'État.

4.1.2 Méthodologie

Question

Quels sont les incitatifs les plus à même de motiver les travailleurs du secteur informel, non-pauvres, à souscrire à la couverture sanitaire nationale indonésienne, *JKN Mandiri* ?

Design du programme

Cette étude expérimentale fut implémentée en coopération avec le gouvernement Indonésien entre février et décembre 2015 dans deux larges villes urbaines, Kota Medan et Kota Bandung. Le programme visait à tester trois (3) types de traitements: les subventions temporaires, l'assistance à l'inscription sur internet et la fourniture d'informations sur l'assurance et les coûts de certains types de maladies. Cette

évaluation est inspirée de théories et études réalisées par d'autres auteurs montrant les effets de ces traitements sur l'enrôlement chez les travailleurs du secteur informel (Banerjee et al. 2019, p.3) L'objectif de l'expérimentation est de relever les facteurs encourageant l'adhésion à la couverture santé chez les ménages non-pauvres couverts par le programme JKN *Mandiri*. Sur 6000 ménages constituant l'estimation de l'intention de traitement (IDT), 5 996, 1446 à Medan et 4550 à Bandung, ménages ont réellement participé au programme, ce qui correspond à l'estimation de l'effet de traitement sur les traités (ETT). Les trois classes de couverture offertes durant l'étude sont divisées selon les coûts suivants: 59.500Rp ou 4.5\$ (classe 1), 42.500Rp ou 3\$ (classe 2) et 25.500Rp ou 2\$ (classe 3), qui était la couverture la plus prisée par les participants (Banerjee et al., p.8). La taille estimée des ménages en 2019 était de 4,3 à Medan et 3,8 à Bandung (Statistik Indonesia, 2020, p.93). Le salaire minimum en 2017 par mois dans la province de Sumatra du Nord (Medan) et Java de l'Ouest (Bandung) était respectivement de 1 961 355 Rp (140.33 US⁵) et 1 420 624 Rp (101.64 US) (Statistik Indonesia, p.138).

Procédure d'échantillonnage

Les ménages choisis pour l'expérimentation furent sélectionnés selon les critères suivants: être éligible au programme JKN *Mandiri*; n'avoir jamais eu d'assurance et ne pas se situer en dessous du seuil de pauvreté des 40 % très pauvres; et représenter la diversité culturelle de l'Indonésie. Deux (2) sous-districts ont été sélectionnés à Medan et huit (8) sous-districts à Bandung en fonction du ratio élevé de travailleurs informels non-pauvres (Banerjee et al., p.8-9). À Medan, douze (12) municipalités sur seize (16) furent choisies aléatoirement, et quatre (4) municipalités furent choisies dans chaque sous-districts de Bandung, pour un total de 32. Ainsi, sur un échantillon de départ de 52 584, 7 629 ménages, soit 14%, répondaient aux critères d'éligibilité. Après avoir croisés les données de sondage avec ceux du

⁵ Estimations du 10 juin sur : <https://themoneyconverter.com/FR/IDR/USD>

gouvernement, 1 633 ménages furent éliminés car ils avaient, sans le savoir, droit à une assurance gratuite offerte par le gouvernement (Banerjee et al., p.9).

Design de l'expérimentation

Tableau 2: Répartition des groupes de traitements par ville

Medan					
	Status quo		Assisted internet registration		Subsidy treatment totals
	Standard information	Extra information	Standard information	Extra information	
No subsidy	37	63	134	237	471
Half subsidy	171	215	26	66	478
Full subsidy	176	54	170	97	497
Registration treatment totals	716		730		

Bandung						
		Status quo		Assisted internet registration		Subsidy treatment totals
		Standard information	Mandate information	Standard information	Mandate information	
No subsidy	Standard information	236	307	241	274	2236
	Waiting period information	232	300	297	349	
Half subsidy	Standard information	160	153	77	82	918
	Waiting period information	141	114	100	91	
Full subsidy	Standard information	85	40	62	54	478
	Waiting period information	63	51	70	53	
Bonus subsidy	Standard information	114	86	170	111	918
	Waiting period information	101	86	131	119	
Registration treatment totals		2269		2281		

Source : Banerjee et al. (2019), p.28.

L'enquête préliminaire des chercheurs a permis d'assigner les ménages aux trois groupes de traitements: les subventions temporaires, le soutien à l'inscription en ligne et la fourniture d'informations (Banerjee et al., 2019, p.9).

Le traitement portant sur les subventions temporaires: porte sur trois (3) groupes dans lesquels les ménages sont assignés de manière aléatoire : un groupe **sans subvention** (groupe contrôle); un groupe avec une subvention pour la moitié de la famille (l'autre moitié devant payer pour bénéficier du traitement) (**demi-subvention**); et un groupe pour lequel toute la famille bénéficie d'une subvention (**subvention**

complète), le tout sur une durée d'un (1) ans. Les familles ont une période de 8 semaines à partir de l'offre de subventions gratuites pour s'inscrire au programme JKN (Banerjee et al., p.10); après cela, l'offre n'est plus valable.

Le traitement portant sur le soutien à l'inscription sur internet: vise à réduire les coûts liés à l'inscription, nécessitant un déplacement dans la capitale du district. La moitié de l'échantillon fut offert un service d'assistance à l'inscription sur internet à partir de chez eux. L'agent aidant détenait un ordinateur avec réseau internet pour aider les familles à remplir les informations et si nécessaire, il retournait sur place pour les aider à continuer la complétion de l'inscription.

Le traitement portant sur la fourniture d'informations supplémentaires: permet de fournir, à tous les ménages, des renseignements sur la couverture JKN *Mandiri*, les différentes primes d'assurance et les procédures d'inscription au programme. À Medan, des ménages choisis au hasard furent en plus briffés sur les coûts financiers des maladies chroniques commun en Indonésie ainsi que les coûts liés à une crise cardiaque. À Bandung, deux informations supplémentaires furent offertes sur: la période d'attente entre l'inscription et l'activation de la couverture santé et l'obligation judiciaire d'inscription au programme au risque d'une amende. Ces derniers ménages bénéficièrent également d'informations sur les coûts des services de santé.

Collecte des données

Les données de l'étude sont issues de **données primaires et secondaire**. Les données primaires sont obtenues à partir d'une **enquête** menée par les chercheurs et une firme indépendante **avant le début de l'évaluation**. L'objectif ici est de relever des informations sur les caractéristiques démographiques des membres de familles, leur état de santé et l'utilisation précédente des services de santé, ainsi que leur connaissance du programme de protection social (JKN *Mandiri*) (Banerjee et al., 2019, p.11). À Bandung uniquement, des données sur le revenu et l'emploi ont été collectées. Les **données secondaires**, à savoir:

l'enregistrement au programme d'assurance, le paiement des primes, les réclamations des bénéficiaires, sont d'origines **gouvernementales**, et sont prélevées entre février 2015 et août 2018. Ces données sont **utilisées à la fin de l'évaluation** pour mesurer : les taux d' enrôlement au programme (inscription à l'assurance durant les huit semaines d'offre), la couverture sociale (avoir payé la prime d'assurance durant le mois prévu) et l'utilisation des services de santé (avoir réclamé les frais de services médicaux); un numéro d'identification national permet d'assortir les répondants aux données administratives (Banerjee et al., 2019, p.12). Il à noter qu'à Medan, l'évaluation était plus axée sur la différence du taux d'inscription à la couverture, tandis qu'à Bandung, le focus était plus sur la détection du taux de réclamations (après utilisation de la couverture).

4.1.3 Résultats

Tableau 3: Effets des subventions et de l'assistance internet sur l' enrôlement

	Decomposition			
	Enrolled within 1 year	Attempted to enroll within 8 weeks of offer date	Enrolled within 8 weeks of offer date	Enrolled after 8 weeks, but within 1 year of offer date
	(1)	(2)	(3)	(4)
Panel A: Main effects				
Full subsidy	0.186*** (0.020)	0.276*** (0.020)	0.209*** (0.018)	-0.023** (0.010)
Half subsidy	0.100*** (0.014)	0.161*** (0.016)	0.114*** (0.013)	-0.014* (0.008)
Assisted internet registration	0.035*** (0.011)	0.237*** (0.011)	0.043*** (0.009)	-0.008 (0.006)
No subsidy mean	0.086	0.099	0.030	0.056
P-value of test of hypothesis				
Half subsidy = full subsidy	0.000	0.000	0.000	0.442
Assisted internet registration = full subsidy	0.000	0.107	0.000	0.266

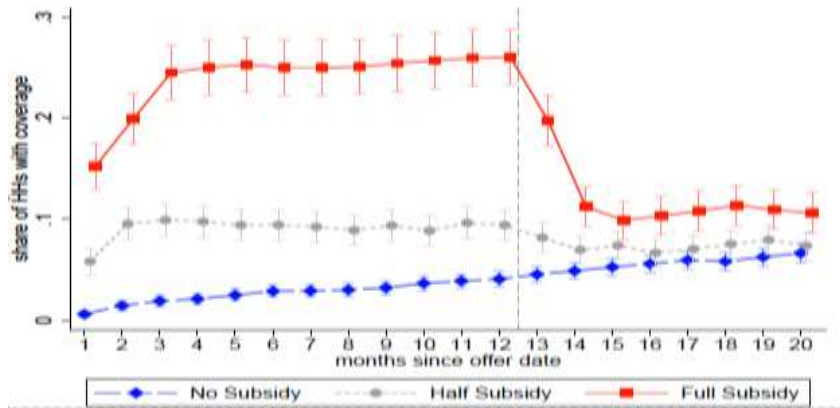
Source: Banerjee et al. (2019), p.28. L'erreur standard est entre parenthèse. **Valeur p** : ***p<0.01 (1%); **p<0.05 (5%); *p<0.1 (10%).

Les subventions complètes: en utilisant le modèle de probabilité linéaire, on peut estimer un taux d'inscription de 27.2 %⁶, soit une hausse à l'inscription de 18.6 points de pourcentage (216%⁷) par rapport au groupe de contrôle qui a une hausse de 0.086 points de pourcentage soit un taux d'inscription de 8.6% durant la période d'offre des subventions de 12 semaines à partir de la date d'offre (ceux qui se sont inscrits après la date d'offre des huit semaines ont quand même payé les primes d'assurance) (Banerjee et al., p.14). Le test de probabilité, valeur p, indique que les subventions totales gratuites et les demi-subventions ont des effets statistiques différents sur l'inscription. La valeur 0.000 (0) signifie qu'il y'a une chance négligeable que les deux traitements aient les mêmes effets. De plus, les ménages de ce groupe ont une probabilité de rester plus longtemps couvert l'année suivant la fin des subventions gratuites que les groupes non-subsventionnés; au 20^{ème} mois, les deux traitements sont presque similaires. Ceci pourrait s'expliquer par l'expérience positive dont le groupe a bénéficié auprès du système de santé, qui accroît la valeur de l'assurance, d'une part, et l'importance des subventions pour maintenir un enrôlement important sur le long terme, d'autre part (Dartanto et al., 2019; Erlangga et al., 2018; Dartanto et al., 2016; Dartanto et al., 2020). Le coût net de couverture par personne pour ce groupe, soit la différence entre les revenus sur les primes d'assurance et les dépenses de l'État, durant la période de l'intervention, était similaire à ceux des non-subsventionnés (3.50\$ versus 9\$, ou encore, un déclin de réclamation de 152,000Rp par rapport à un revenu sur les primes de 71.000Rp) (Banerjee et al., p.21-22). Une des explications est que le groupe totalement subsventionné aurait moins de réclamations (plus de visites - *outpatient*) par rapport au groupe de contrôle (plus d'hospitalisation - *inpatient*) (48 % versus 62 %) (Banerjee et al., p.19), et serait donc potentiellement plus en santé (Prabhakaran et al. 2019, p.63).

⁶ Dans la colonne 1 du tableau de régression, le taux d'inscription sur un an correspond à: $(0.186 + 0.086) * 100$.

⁷ Effet de traitement : indique l'effet de la subvention gratuite sur la hausse du taux d'inscription par rapport au groupe de contrôle. Modèle de calcul: $(27.2 - 8.6 / 8.6) * 100$.

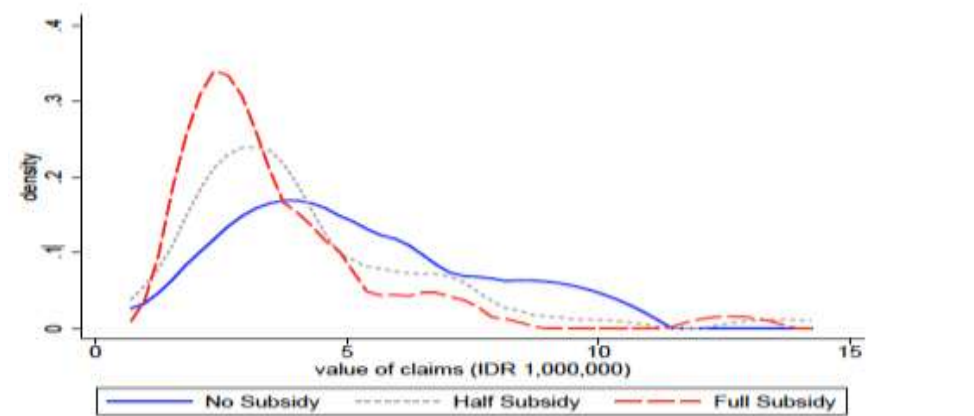
Graphique 1 : Enrôlement par mois depuis le début de l’offre par type de subvention.



Taille de l’échantillon : 5996. Banerjee et al. (2019), p.30.

Les demi-subsidations: ont mené à une hausse de 18.6%, soit une le taux d’inscription de 10 points de pourcentage (116 %), par rapport au groupe de contrôle (Banerjee et al., 2019, p.14). Le taux de couverture de ce groupe se situait entre celui des subventions complètes et celui du groupe contrôle. Leur nombre de réclamations était de surcroît supérieur au groupe ayant reçu une subvention totale, et la valeur des réclamations supérieure à celle du groupe de contrôle (Banerjee et al., p.20).

Graphique 2: Valeurs des réclamations des frais d’hospitalisations par type de subvention.

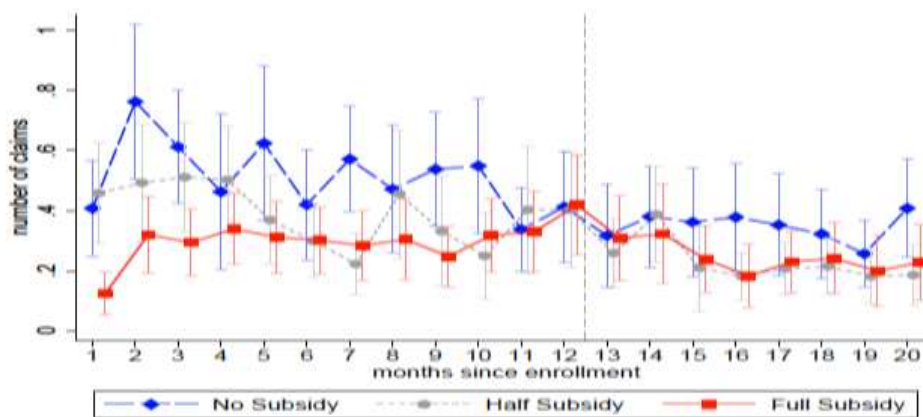


Taille de l’échantillon : 3827. Banerjee et al. (2019), p.30.

Le groupe de contrôle (sans subventions): avait plus de réclamations (hospitalisation) dans les trois premiers mois de l’intervention et abandonnait la couverture (ne payait plus les primes) très tôt par rapport aux autres groupes. L’inscription à seulement haussée de 9 % dans ce groupe sur une période de

12 mois depuis le début de programme. Les ménages semblent tirer avantage d'une faille du système qui permet aux personnes de s'inscrire au besoin et d'abandonner lorsque voulu. En outre, les coûts de réinscription sont élevés pour décourager les personnes à se réinscrire. Les évaluateurs relèvent ainsi que 72 % de ce groupe optent pour l'assurance de classe 3, soit le coût le plus bas de 2\$ (25,500 Rp) (Banerjee et al., 2019, p.8).

Graphique 3: Nombre de réclamations depuis le début de l'inscription par type de subventions.



Taille de l'échantillon : 749. Banerjee et al. (2019), p.30.

L'assistance à l'inscription sur internet: hausse l'inscription de 3.5 points de pourcentage (41 %). En outre, lorsque cette offre était combinée à l'offre de subvention gratuite, 60 % ont tenté de s'inscrire, mais seulement 30% y sont arrivés (Banerjee et al., p.5 et 23). Ces différences de ratio cachent les causes de non-enrôlement de ces ménages. En effet, ce groupe de traitement tend à vite arrêter le paiement des primes d'assurance que ceux n'ayant pas eu d'assistance, reflétant en partie les coûts indirects et non-monétaires liés aux paiements mensuels qui obligent les ménages à se déplacer (Banerjee et al., p.18; Prabhakaran et al., 2019, p.40).

La fourniture d'informations: semble dans cette étude avoir peu d'impact sur l'enrôlement et de ce fait, ne constitue pas une barrière primaire de l'expansion de la couverture dans ce cas-ci (Banerjee et al., 2019, p.15).

En somme, les subventions gratuites pendant une année et l'assistance à l'inscription aux ménages ont des effets positifs sur le taux d'inscription et améliorent la sélection adverse, en attirant des groupes représentant un faible coût en raison de leur faible taux de réclamations (40% inférieur à ceux n'ayant pas bénéficiés des subventions) (Banerjee et al., p.19). Par contre, les évaluateurs observent que, même la meilleure combinaison d'offre est loin de l'objectif global de couverture nationale, avec seulement 30% de hausse sur l'enrôlement par rapport à un taux initial de 8 % avant l'étude (Banerjee et al., p.6). Il semble que les faibles capacités administratives pour soutenir l'enrôlement et prévenir l'abandon de la couverture à court terme freinent l'objectif de couverture nationale. Notons qu'aucune enquête qualitative ne fut entreprise durant cette évaluation.

4.2 Étude de cas sur le Nicaragua

4.2.1 Contexte

Le Nicaragua a fait des progrès considérables dans son effort de lutte contre la pauvreté. Entre 2005 et 2009, le taux de pauvreté est passé de 48.3 % à 42.5 %. Cet effort, marqué par un recul de 5.8 points de pourcentage de la pauvreté, était soutenu par les politiques distributives durant la période de la crise économique, 2007-2008 (IMF, 2011, p.7). Ces succès ne dissimulent pourtant pas le fait que le Nicaragua reste le deuxième pays le plus pauvre de l'Amérique Latine, après Haïti, avec un revenu national brut par habitant de 2 020 dollars US sur une moyenne régionale de 8 695 dollars US. Son produit national brut (PIB) était de 13.064 milliards en 2018. Le taux de pauvreté quant à lui se situait à 24.9 % (Banque mondiale, 2008), sous le seuil de pauvreté national défini comme la consommation de bien alimentaire et non alimentaire en 2016 pour une population de 6.303.974 habitants, soit environ 1.569.689 millions vivant dans la pauvreté (Banque mondiale, 2016). En 2018, le Nicaragua était classé au 126^e rang parmi des pays à développement moyen, au niveau mondial, avec un IDH de 0.65. Son coefficient Gini

correspondait à 46.2 (0.462) entre 2010-2017, soit un niveau d'inégalité assez profond entre les riches et les pauvres (Rapport PNUD, 2019).

Cette étude de cas (Hatt et al., 2009; Thornton et al., 2010) est menée entre 2006 et 2007. Avant 2007, la couverture sociale, offerte via l'Institut National de Sécurité Sociale (INSS), était accessible qu'aux travailleurs du secteur formel et les fonctionnaires publics, qui ensemble, représentaient 18.5 % de la population active (Hatt et al., 2009, p.5). Quant aux travailleurs du secteur informel, c'est en 2007 qu'un programme similaire à celui de l'INSS, appelé *Seguro Facultativo de Salud*, fut mis en place pour leur servir de levier face aux instabilités financiers. Antérieurement à ce nouveau programme, cette catégorie de travailleurs accédait aux services de santé gratuits auprès du Ministère de la santé du Nicaragua (MINSa), ou déboursait de leur poche dans les centres de santé privés. Cependant, peu d'entre eux se tournaient vers le MINSa, car les infrastructures étaient défectueuses et les capacités de prises en charges très précaires (Hatt et al., p.2).

Environ 1.2 million de personnes, opérant en tant que vendeurs, professionnels indépendants, micro-entrepreneurs et contractuels, constitue le secteur informel qui regorgerait d'environ 93% de personnes vulnérables sans couverture sociale (Hatt et al., p.1). Avec plus de 200 institutions micro-financières dont 20 couvrent des ménages sans couverture bancaire (Thornton et al., 2010, p.184), le gouvernement nicaraguayen décida d'étendre le programme à travers des partenaires micro-financiers.

4.2.2 Méthodologie

Question d'évaluation

Quels facteurs influencent la volonté des travailleurs du secteur informel à s'enrôler au programme d'assurance de l'Institut National de Sécurité Sociale nicaraguayen (INSS)?

Design du programme

Cette évaluation était centrée sur un projet pilote, cofinancé par l'USAID (*PSP-one*) et le *Global Development Network* (GDN), portant sur l'extension de la couverture de sécurité sociale aux travailleurs du secteur informel à travers l'INSS, entre 2007 et 2008. L'objectif de l'étude était d'accroître l'accès aux services de santé et la protection contre les risques financiers chez les pauvres. Il s'agissait d'offrir une **gamme de couverture santé**, dont entre autres: des soins curatifs, des soins de préventions, des services primaires et spécialisés, des examens de laboratoires, des hospitalisations et la distribution de médicaments, des soins pré-postnatals, des services de pédiatrie, le planning familial et l'offre de vaccins pour les enfants, prévention et traitement du paludisme, consultation pour le VIH/Sida et les maladies sexuellement transmissibles, etc., qui sera accessible aux travailleurs qui bénéficieront du nouveau programme⁸ (Thornton et al., 2010, p.184). *Seguro Facultativo de Salud* fonctionnait sur une base de participation volontaire avec un paiement de prime mensuel, qui est d'environ 17 dollars US les 2 premiers mois et 15 dollars US les mois suivants (Hatt et al., 2009, p.5). L'étude s'est déroulée dans la ville de Managua où 93% des travailleurs dans l'informel sont sans couverture sociale (Hatt et al., p.1). Trois institutions micro-financières (IMF) contractuelles se chargeaient de fournir les services de couverture aux bénéficiaires. Ceux-ci pourront faire les paiements directement au niveau de leur institution d'assignation.

Design de l'expérimentation

Pour évaluer les facteurs influençant l'enrôlement au programme de couverture sociale et ses effets sur l'utilisation des services de santé, les participants à l'étude ont été assignés au hasard à trois groupes de traitements offrant: **une brochure informationnelle sur les produits d'assurance de l'INSS; une brochure et 6 mois d'assurance subventionnée avec instruction de s'inscrire auprès de l'INSS, et une brochure et 6 mois d'assurance subventionnée avec instruction de s'inscrire auprès d'une des**

⁸ Les coûts de couverture par services ne sont fournis dans aucun des articles cités par les auteurs.

IMF. Le **groupe de contrôle pure** est celui n'ayant bénéficié d'**aucun prix** durant la loterie (Thornton et al., 2010, p.191).

Le processus de randomisation s'est fait au niveau individuel chez les 2 608 répondants. Un groupe de 175 individus, soit 112 qui ont bénéficié de 6 mois de subventions et 63 individus qui ont juste reçu la brochure (Hatt et al., 2009, p.8), ont eu l'opportunité de s'enrôler directement sur leur lieu de travail (au marché) en 2007. Le coût de la subvention de 6 mois valait 96 dollars US (Thornton et al. p.204); la subvention était fournie sous forme de tickets gratuits. Le paiement des assurances subventionnées s'effectue au nom des répondants directement auprès de l'INSS.

Pour comprendre le sens des résultats de l'évaluation et l'opinion des répondants sur le programme, des entrevues et des groupes de discussion furent tenus au printemps 2008 avec ceux ayant donné leur consentement pour y participer. Les discussions furent enregistrées et retranscrites par un professionnel (Thornton et al. p.199).

Caractéristiques des échantillons

Tableau 4 : Caractéristiques des participants à l'étude

	All (N = 2608)		MFI client (N = 1013)		Non-MFI client (N = 1595)		
	Mean	SD	Mean	SD	Mean	SD	Difference
<i>Panel A: Demographic characteristics</i>							
Age	37.79	9.28	39.05	8.57	36.98	9.63	2.07***
Male	0.35	0.48	0.28	0.45	0.39	0.49	-0.11***
Years education	9.26	4.22	9.25	4.15	9.27	4.27	-0.02
Married	0.70	0.46	0.70	0.46	0.70	0.46	0.00
Number of children	2.04	1.52	2.22	1.52	1.93	1.52	0.29***
With under 12 children	0.66	0.47	0.66	0.48	0.67	0.47	-0.01
Pregnant	0.03	0.18	0.03	0.18	0.03	0.17	0.00
	All		MFI client		Non-MFI client		
<i>Panel B: Economic characteristics</i>							
	Mean	SD	Mean	SD	Mean	SD	Difference
Monthly income	273.89	380.34	297.22	404.12	258.77	363.45	38.44***
Proportion with savings	0.29	0.46	0.28	0.45	0.30	0.46	-0.02
Monthly savings	25.11	158.51	23.40	157.89	26.19	158.95	-2.78
Last year's savings Balance	145.73	574.94	141.45	609.66	148.45	551.85	-7.00
owns home	0.76	0.42	0.80	0.40	0.74	0.44	0.06***
Total household health-care costs	84.46	171.62	88.67	174.38	81.78	169.85	6.89
Total health-care costs for respondent	46.50	114.44	51.68	128.77	43.20	104.20	8.48*
Last visit cost for respondent	17.77	59.98	20.79	82.05	15.86	39.95	4.93**
Proportion MFI Clients	0.39	0.49					

Source : Thornton et al. (2019), p.187. **Valeur p :** * signifiant à 10%; **signifiant à 5%; *** signifiant à 1%.

L'étude se limite à l'échantillon de 2 608 répondants représentant 93% (Hatt et al., 2009, p.10) de l'échantillon initial de 2 806 répondants. L'âge moyen des répondants était de 38 ans, avec 9.3 ans d'éducation complétée, et 35% étaient des hommes. 70 % de l'échantillon était marié avec deux enfants en moyenne, dont deux sur trois ont moins de douze ans (tableau 4, colonne 1 panel A, N=2608). Durant l'enquête de base, sur les 2 608 répondants, 72 % était des clients d'une IMF étaient des femmes, avec plus d'enfants, contre 61% qui n'étaient pas clients. Le revenu moyen par mois était de 274 (tableau 4, colonne 1, panel B, N=2608) dollars US, et 422 dollars US durant l'enquête de suivi, avec 25 dollars de revenu disponible, en moyenne, après les dépenses ménagers (Thornton et al., p.186-187). Cependant ces données sur le revenu doivent être prises avec précaution car le revenu mensuel national rapporté par la banque centrale du pays était de 129 dollars US (Hatt et al., p.16). Les clients des IMF gagnent plus en revenu que les autres groupes et dépensent 8 \$ de plus en santé (Thornton et al., 2010, p.187). 80% de l'échantillon a reporté être malade l'année avant l'enquête, mais la majorité sont des cas mineurs: 60% ont souffert de rhume; en outre, 11% ont manqué plus d'une semaine de travail (Thornton et al., p.187)

En termes d'utilisation des services de santé, 73% ont reporté avoir visité des pharmacies pour des prescriptions, et cela malgré la gratuité au sein des infrastructures publiques. 29 % préfère visiter les centres privés qui offrent de meilleurs services malgré les coûts élevés (16.97\$) par rapport à 9% qui vont dans des hôpitaux publics. Les dépenses en santé avant l'enquête étaient de 84.46\$ pour un individu et un mineur de moins de 12 ans, soit 31 % du revenu, et 46.50 dollars US par individu. Face au coût annuel de l'assurance, 186 dollars US, les ménages étaient moins enclins à payer pour la couverture, car ce coût était supérieur à leurs dépenses personnelles en santé (Thornton et al., 2010, p.187-189).

Collecte des données

Le plus grand marché à Managua, *Mercado Oriental*, fut choisi pour l'étude. **Deux types de données** furent collectés : celles **quantitatives** et **qualitatives**. **Deux enquêtes de base** prirent place en

2007: de mars à avril sur 48 unités du marché, et de juin à juillet sur 25 unités (Hatt et al., 2009, p.7) incluant deux petits marchés hors de *Mercado Oriental*, soit *Mercados Huembes* et *Iven Montenegro* (Thornton et al. 2010, p.185). Ces deux premières enquêtes consistaient à relever les caractéristiques sociodémographiques des répondants, leur taux d'utilisation des différents services de santé avant et pendant l'enquête, ainsi que leur niveau de dépense en santé. De plus, les informations sur les cartes d'identité gouvernementales furent recueillies pour pouvoir assortir les personnes à leur assurance. **L'enquête préliminaire de 2006** permit de sélectionner aléatoirement les personnes éligibles au programme selon les critères: d'âge (18 à 54 ans), de détention de carte d'identité gouvernementale, de possession d'un kiosque, et d'absence d'une couverture santé. Entre **mars et juin 2008**, l'**enquête de suivi** permis d'interviewer 2 608 répondants (Hatt et al., p.9). Entre **septembre et octobre 2008**, cette **enquête qualitative**, portant sur des entrevues et des groupes de discussions, fut tenue avec les vendeurs pour discuter entre autres: des raisons d'enrôlement et de non-enrôlement; de leur expérience avec le processus d'enrôlement; de leur perception du rôle des IMF comme agent d'assurance intermédiaire; et de leur propension à payer pour l'assurance santé (Hatt et al., 2009, p.10). Les trois institutions micro-financières, *ACODEP*, *Banco ProCredit* et *Findesa*, choisies pour collecter les primes d'assurance et promouvoir le nouveau programme ont signé un contrat d'un (1) an avec l'INSS en 2006. Ce contrat ne sera pas renouvelé sous le nouveau régime marquant l'arrivée du président Ortega au pouvoir en 2007.

4.2.3 Résultats⁹

Tableau 5: Les déterminants de l'enrôlement à l'assurance

⁹ Noter que toutes les données n'ont pas été inscrites dans les tableaux par les auteurs. Se référer à la référence entre parenthèse.

Dependent variable: signed up for Insurance (0/1)	All		MFI client	Non-MFI client
	(1)	(2)	(3)	(4)
Information only	-0.049** [0.023]	-0.058** [0.028]	-0.051 [0.041]	-0.093** [0.041]
Six-month subsidy – INSS	0.333*** [0.028]	0.336*** [0.033]	0.384*** [0.051]	0.275*** [0.046]
Six-month subsidy – MFI	0.281*** [0.028]	0.277*** [0.033]	0.313*** [0.048]	0.223*** [0.047]
Male	-0.023 [0.015]	-0.024 [0.017]	-0.007 [0.030]	-0.032 [0.021]
Education	0.002 [0.002]	0.002 [0.002]	0.001 [0.003]	0.003 [0.003]
Married	0.014 [0.016]	0.023 [0.018]	-0.002 [0.029]	0.037 [0.023]
MFI client	0.033** [0.015]	0.031* [0.016]		
Number of children	-0.015*** [0.005]	-0.012** [0.006]	-0.032*** [0.010]	0.001 [0.008]
Has children under 12	0.030* [0.017]	0.035* [0.019]	0.111*** [0.030]	-0.015 [0.024]

Source : adapté de Thornton et al. (2010), p.192. **Valeur p** : * signifiant à 10%; **signifiant à 5%; *** signifiant à 1%. L'erreur standard se trouve entre parenthèse. Modèle de la régression linéaire. (1) enquête de base et (2) de suivi.

Sur 2 608 individus¹⁰, 20.3 % ont souscrit à l'assurance (Thornton et al., 2010, p.191). Le fait d'avoir un enfant en dessous de 12 ans (pour les clients des IMF), d'accéder aux subventions, et de s'inscrire sur place, sont des facteurs qui ont eu un impact positif sur l'enrôlement pour les clients des IMF (tableau 5, colonne 3 et 4).

Le traitement portant sur les brochures informationnelles: a été plus efficace lorsque les répondants ont eu l'opportunité de s'inscrire sur place (**au marché**) que dans un des **centres des IMF ou à l'INSS**. Ainsi, ceux ayant pu s'inscrire sur place se sont inscrit à hauteur de 23% (tableau 5 colonne 1) contre 1% (Thornton et al., p.194). Le signe négatif indique une relation opposée entre l'enrôlement et la variable information; donc l'information seule à un impact peu important sur l'enrôlement.

Le traitement portant sur les subventions de 6 mois: encourage plus l'enrôlement par rapport au groupe de contrôle pur, soit une hausse de 33% chez l'INSS et 28% chez les IMF (tableau 5, colonne

¹⁰ La taille des échantillons est de 119 pour le groupe de contrôle pure; 1200 pour le groupe avec les brochures informationnelles; 758 pour les répondants de l'INSS et 730 pour ceux des IMF (Thornton et al., 2010, p.204). Ces tailles sont moins conséquentes pour mesurer l'effet entre différent traitement.

1 à 4) (*Procredit* 47%; *Findesa* 15%; *ACODEP* 37%) (Thornton et al., p.193). **L'inscription au marché** était supérieure de 70% pour les 112 bénéficiaires d'une subvention de 6 mois (Thornton et al., p.193). Ces données reflètent l'absence des coûts liés au déplacement pour l'inscription, à la prise de photos, aux photocopies pour la carte d'identité, et tous les tracas pouvant affecter le processus d'inscription.

En termes d'enrôlement, les variables **sexe, situation matrimoniale et éducation, influencent très peu**, de manière générale, les choix des répondants à s'enrôler (tableau 5).

Tableau 6: Effets de l'assurance sur les dépenses en santé et l'utilisation des services de santé

	Visited		Number of visits		Ln (amount spent)	
	Coeff.	SE	Coeff.	SE	Coeff.	SE
<i>Panel A: all (N = 2608)</i>						
Pharmacy	-0.020	[0.044]	0.468	[0.444]	-0.548	[0.344]
Private doctor	0.029	[0.038]	0.045	[0.121]	0.222	[0.262]
Laboratory	-0.067	[0.054]	-0.155	[0.208]	-1.072***	[0.365]
EMP	0.427***	[0.027]	1.542***	[0.145]	—	—
Private clinic/hospital	-0.089*	[0.050]	-0.473***	[0.181]	-0.565	[0.351]
Public (MINSA) health center	-0.063	[0.047]	-0.424*	[0.247]	0.006	[0.028]
Public (MINSA) hospital	-0.115***	[0.040]	-0.176	[0.152]	-0.025	[0.021]
All/any	-0.001	[0.034]	0.918	[0.749]	-0.522	[0.343]

Source: adapté de Thornton et al. (2010), p.195. **Valeur p :** * significatif à 10%; **signifiant à 5%; *** significatif à 1%. L'erreur standard est entre parenthèse.

Concernant **les effets de l'assurance**, d'une part **le nombre de visites** des centres de santé **n'ont pas substantiellement augmenté**, mais ont provoqué un virage vers les cliniques médicales de l'INSS, les *Empresas Médicas Previsionales* (EMP) avec une hausse de 42.7 (tableau 6, colonne 1) points de pourcentage. Quant aux **visites dans les centres de santé privés et publics, elles ont baissé** (Thornton et al. p.194). Parallèlement, les **dépenses directes en santé** des ménages **ont baissé** substantielle de 55% (Thornton et al., p.195). Cependant, les individus étaient moins enclins à dépenser pour la prime lorsque celle-ci était supérieure à leur dépense personnelle en santé 84.46\$ versus 186\$ (Thornton et al., p. 189).

D'autre part, **l'utilisation des services de santé maternel et l'accès à la planification familiale** n'ont **pas été affectés** par l'accès à l'assurance. Cependant, les femmes couvertes ont été mieux prises en charges au sein des cliniques EMP, soit 28% contre 7% (Thornton et al., p.196). L'absence de corrélation entre l'assurance et la demande en soins de santé s'observe entre les assurés et non-assurés, laissant supposer que la qualité de l'offre de santé est restée la même et que les subventions ont eu peu d'effet.

Concernant **la rétention** des individus après l'expiration des 6 mois de subventions, moins de 10% des 20% enrôlés au départ du programme étaient restés durant l'enquête de suivi de 2008. Ceux ayant reçu les subventions et les informations étaient plus propices à abandonner le programme, suggérant leur faible volonté à payer pour s'enrôler. À l'opposition, ceux n'ayant pas bénéficié des subventions avaient plus de chance de rester dans le programme; similairement, les plus éduqués avaient une forte propension à garder leur couverture. Les clients des IMF et ceux qui ne l'étaient pas avant l'étude montrent les mêmes caractéristiques face à la rétention. Par contre, les clients des IMF, ayant bénéficié de 6 mois de subventions, restait à hauteur de 4.4 points de pourcentage que ceux de l'INSS dans le programme; ceci peut se comprendre par rapport au système de paiement au sein des IMF qui est rapide et plus convenable (Thornton et al., 2010, p.197-198).

L'enquête qualitative

Les **données qualitatives** issues de 9 groupes de discussions et 60 entretiens avec des vendeurs (20 qui se sont enrôlés et 40 qui ne se sont pas enrôlés) et des agents des IMF permettent de mieux expliciter les résultats quantitatifs obtenus, et de ce fait de comprendre les facteurs ayant limité l'enrôlement. Pour ce faire, cinq (5) strates ont été formées en fonction: du statut de client des IMF; de la réception de subventions; et de la preuve d'enrôlement (Thornton et al., p.198). Il en est résulté de ces données qualitatives que **le temps d'inscription**, les enjeux liés à la **collecte des documents officiels** (carte d'identité), le manque de précision sur **le processus d'enrôlement**, **l'inexactitude des**

informations sur les bénéfices de l'assurance, la **faible confiance** portée **aux IMF**, et le **coût de la prime d'assurance**, sont des facteurs importants pouvant expliquer le faible enrôlement et la faible rétention au programme (Thornton et al., p.199).

La confiance face aux IMF a été soulevée durant les groupes de discussions, où les répondants défendent préférer s'enrôler auprès de l'INSS que les IMF, qui sont vulnérables à la faillite et moins dignes de confiance. Un client du marché Oriental soutient que : « L'INNS connaît les bénéfices de l'assurance santé et ce qu'elle couvre, tandis que les IMF sont des institutions de prêts, [et de ce fait], il ne comprend pas le rôle [des IMF] en santé, car ceci n'étant pas leur spécialité » (Thornton et al., p.200).

Quant **au rôle des EMP**, ces cliniques n'ont pas fournis les services de qualité ni les informations essentielles sur l'utilisation de la couverture; elles minent de par ce fait la rétention de la clientèle.

La disposition à payer pour la couverture fut limitée par le coût d'acquisition de cette dernière, malgré que la plupart la trouvât utile (voir Hatt et al., 2009, p.45). Cependant, avoir un enfant en dessous de 12 ans et avoir des maladies chroniques augmentent les chances d'enrôlement. Par contre, certains répondants préfèrent la qualité des services offerts par les pharmacies dont l'efficacité est plus convenable par rapport à l'attente liée au processus d'enrôlement au programme en général.

Le partenariat public-privé qui fut formé eut du mal à faire ses preuves face aux différences de style de gestion et de fourniture des services. Les IMF et l'INSS ont eu du mal à efficacement communiquer entre eux à cause du manque de confiance. L'INSS suspectait les IMF de chercher à se faire des profits sur le programme en s'attirant les clients les plus en santé et les plus aptes à payer leurs prêts (Thornton et al., p.200). À cela s'ajoute les difficultés des IMF à garder la motivation après la décision du gouvernement de ne pas renouveler leur contrat à la fin de l'évaluation. Tous ces différents facteurs, marqué par le manque de professionnalisme des partenaires d'implémentation du programme, minèrent le succès du programme.

En somme, cette section décrivait le design de deux programmes testant les facteurs incitant à l’ enrôlement à l’ assurance santé chez les travailleurs du secteur informel ainsi que leur utilisation des services de santé. L’ étude sur l’ Indonésie a montré que le taux d’ enrôlement était plutôt faible malgré les subventions d’ une année gratuite et l’ assistance à l’ inscription. Les évaluateurs ont indexé les faibles capacités administratives du gouvernement à offrir les services de qualité, mais sans pour autant offrir des explications sur les motivations des répondants à rester plus longtemps dans le programme. C’ est par contre le cas de l’ étude au Nicaragua, qui nous a permis de relever, grâce aux données qualitatives, les raisons ayant minés la bonne réussite du programme et surtout, la motivation des répondants à ne pas payer les primes après les six mois du subventions gratuites.

Tableau 7: Résumé des deux études de cas

Pays	Design de l’ évaluation	Groupes visés	Types de traitements offerts	Observations
Indonésie	Randomisation.	Ménages non-pauvres travaillant dans le secteur informel (aucune caractéristique distincte entre les répondants ne fut présentée par les évaluateurs).	<ul style="list-style-type: none"> -Subventions gratuites sur 1 an. -Demi-subventions sur 1 an. -Assistance à l’ inscription sur internet. -La fourniture de diverses informations. -Zéro subvention (groupe de contrôle). 	<ul style="list-style-type: none"> -Les subventions gratuites jouent un rôle important sur l’ enrôlement. L’ étude montre qu’ il serait plus bénéfique si le gouvernement continu le soutien financier aux ménages tout en améliorant les services de santé. Sur le long terme, une expérience positive de l’ assurance pourrait mener à une forte rétention à l’ assurance. -Par ailleurs, l’ absence des données supplémentaires sur les répondants a failli relever les facteurs potentiels ayant affectés le choix de certains à rester plus longtemps couverts. Une enquête qualitative aurait pu pallier ce problème.

Pays	Design de l'évaluation	Groupes visés	Types de traitements offerts	Observations
Nicaragua	Randomisation plus enquête qualitative.	Vendeurs pauvres travaillant dans le secteur informel (voir tableau 4 pour les caractéristiques).	<p>-Subventions gratuites de 6 mois et inscription à l'INSS.</p> <p>-Subventions gratuites de 6 mois et inscriptions aux différents sièges des IMF.</p> <p>-Brochure informationnelle sur les produits de l'assurance santé.</p> <p>-Zéro prix (subvention pure)</p>	<p>-Comme dans le cas précédent, les subventions ont eu des effets plus positifs mais moindre, sur l'enrôlement. La différence entre l'INSS et les IMF indique deux éléments importants : les répondants préfèrent la qualité efficiente des services offerts par le IMF, mais ils ont moins confiance en elles. Ce qui indique un pourcentage élevé d'inscription chez l'INSS par rapport aux IMF (voir tableau 5).</p> <p>- L'enquête qualitative a montré que la faible coopération entre l'INSS et les IMF a en partie minée la bonne conduite du programme; ce qui affecte l'enrôlement tout comme la rétention.</p> <p>-L'enquête qualitative a également permis de valider les données quantitatives en validant la relation entre ces dernières et les réponses des répondants. Mais son usage fut très limité et moins exploré enfin d'en tirer une leçon constructive sur le design des programmes d'assurances santé.</p>

À présent, nous allons analyser, en fonction de quatre concepts, le lien entre le design de chaque étude et la portée des évidences fournies grâce au design d'évaluation de chaque étude.

5. Analyse des études de cas

Ces deux études de cas, bien que n'ayant pas été étendues à d'autres pays ou même à d'autres contextes internes dans les pays choisis, permettent d'explorer cependant des essais randomisés visant la généralisation de la couverture santé aux travailleurs du secteur informel. Ces deux cas servent d'exemples de comparaison de design d'évaluation pure et mixte. Rappelons que l'objet de l'étude est de comprendre comment le design des évaluations expérimentales, intégrant des méthodes qualitatives, impacte le niveau informationnel offert par la randomisation. Nous verrons que, dépendamment du modèle d'intégration des méthodes qualitatives dans l'évaluation, souvent les données qualitatives viennent renforcer les données quantitatives, et peu d'informations nouvelles sont tirées de cette mixité. Voyons le rapport de chaque étude par rapport à : la validité interne, le contrôle des biais, l'explication causale et le rapport à la « boîte noire » (*black box*).

5.1 Inférence causale: validité interne

En Indonésie, l'évaluation par randomisation a permis d'établir la relation entre les différents traitements et leurs effets sur le taux d' enrôlement à la couverture. Le tableau 3, avec la valeur-p, montre que chaque traitement a une signifiante statistique distincte. Une année d'accès à des subventions gratuites à plus d'impacts positifs sur la rétention (graphique 1), après la fin des subventions grâce à l'expérience positive liée à l'utilisation de l'assurance (Dartanto et al., 2019; Erlangga et al., 2018; Dartanto et al., 2016; Dartanto et al., 2020). Par contre, malgré le fait d'offrir un an de subvention gratuite et une assistance à domicile pour l'inscription au programme *JKN Mandiri*, l' enrôlement ne haussait que de 30%. Ce qui correspond à un faible taux si l'on veut atteindre tous les travailleurs du secteur informel. Outre ce taux, certains dans le groupe des subventions complètes ont dû payer les primes après avoir dépassé la date d'échéance de l'offre. Le fait de ne pas clairement distinguer ces groupes, mine l'estimation réelle des effets des subventions gratuites sur la rétention. Par ailleurs, on peut déduire qu'offrir des subventions

gratuites temporaires n'est pas autant efficace si l'on souhaite étendre la couverture à toute la population. Le gouvernement en continuant ses dépenses publiques, supporte de nouveaux revenus chez les ménages, mais aussi pour lui-même grâce à la rétention qui est source de revenus sur les primes.

Au **Nicaragua**, l'assignation des individus aux différents traitements s'est également faite de manière aléatoire. Chaque groupe de traitement a permis d'observer des résultats spécifiques. Comme dans le cas de l'Indonésie, les subventions de 6 mois et l'inscription sur place (**au marché**) ont eu un impact positif sur le taux d' enrôlement, tandis que l'absence de subventions ou la fourniture d'information uniquement, ont eu de très faible impact sur l' enrôlement. La valeur-p, montre que les subventions de 6 mois ont un effet important sur l' enrôlement auprès des IMF et de l'INSS comparativement au groupe de contrôle; tout comme le fait d'avoir un enfant est plus important chez les clients des IMF que les « non-clients » en termes d' enrôlement (tableau 5). Cependant, la taille des échantillons n'était pas aussi importante pour observer un impact réel des subventions sur l'inscription. On peut l'observer à travers le faible écart entre groupes de répondants au départ (1) et durant le suivi (2) (tableau 5). Par ailleurs, l'étude a montré que, lorsque les primes valent plus que les dépenses en santé des ménages, ces derniers sont moins tentés de payer pour les primes d'assurance (Mathaeur et al., 2008; Dartanto et al. 2016).

5.2 Contrôle des biais

En **Indonésie**, les évaluateurs se sont assurés que les ménages aient les caractéristiques sociaux et économiques semblables. Dans les zones choisies pour sélectionner les ménages, il n'avait pas encore eu d'offres d'assurances gratuites¹¹. Les localités avec des universités, des usines, et de larges marchés furent exclues de l'étude pour éviter d'assigner des non-citoyens. La collecte des données primaires s'est faite à travers une entreprise indépendante de sondage, tandis que les données gouvernementales et les cartes

¹¹ Voir note de bas de page, Banerjee et al, 2019, p.9.

nationales, indiquant les membres de familles, ont permises de faire correspondre les répondants à leurs données d'inscription à l'assurance. Par contre, le fait que les données familiales de certains ménages n'étaient pas à jour, des erreurs liées à la validité des informations auraient pu biaiser la validité des résultats sur les bénéficiaires réels des subventions. En outre, les chercheurs se sont penchés sur les zones urbaines afin de contrer la possibilité de faibles demandes en assurance en zone rurale. Ce choix pourrait avoir eu un impact sur les résultats positifs des bénéficiaires des subventions complètes. Ces derniers ayant sans doute une capacité financière plus importante et un accès plus facile aux services de santé par rapport aux zones rurales faisant souvent face à des services précaires; surtout qu'en Indonésie, les inégalités sont réparties selon les zones d'habitations (Statistik Indonesia, 2020). Par ailleurs, même si les ménages ont été choisis aléatoirement, les chercheurs ne précisent pas si ces ménages (malgré qu'ils soient tous des urbains) ont des différences de revenus; si les têtes de ménages sont en moyenne des femmes ou des hommes; s'il existe des systèmes traditionnels de filets sociaux ou encore des services de santé à proximité des localités; si le type d'activités de ces ménages diffèrent, etc. Ces informations manquantes peuvent constituer des sources de biais lorsque vient l'interprétation des résultats de l'évaluation expérimentale, car ces informations (sur le contexte sociale, économique, culturel, et même politique) peuvent cacher des explications comportementales des répondants ou même des évaluateurs (Dartanto et al., 2016; Ly, 2018).

Au **Nicaragua**, la majorité des vendeurs furent choisis dans le grand marché *Mercado Oriental*. Seul deux petits marchés, *Huembes et Iven Montenegro* ne faisaient pas partir du grand marché. Les mêmes caractéristiques sont observables entre les différents répondants (tableau 4); les chercheurs s'étant assurés de mener une enquête d'éligibilité à l'étude, en 2006, avant le début de l'évaluation en 2007. Cependant, le fait d'avoir choisi une localité hors du marché *Oriental* a permis d'observer une différence de niveau d'enrôlement entre les vendeurs. À *Huembes* par exemple, le taux d'enrôlement était plus significatif qu'à *Mercado Oriental* en raison de l'environnement plus sécuritaire, de la clientèle plus

nantie, du nombre de tourisme, et du niveau de formalité plus important que dans le marché Oriental (Hatt et al., 2009, p.25). Les auteurs ont constaté que ceux qui n'avaient pas pu participer à l'enquête de suivi ne possédaient pas de carte d'identité (15.1%) ou encore n'étaient pas disponible sur place (22%), et certains ne voulaient pas participer à l'enquête (7.7%) (Thornton et al., 2010, p.186). Ces données peuvent affecter la valeur des résultats lorsqu'on tente de les généraliser car il se peut que ceux absents soient plus en besoin d'assurance que les participants. Les biais liés à la sélection adverse est aussi potentiel dans une étude portant sur l'offre d'assurance. Lorsque les plus vulnérables se sentent attirés par la valeur de l'offre, leur inscription à l'assurance pourrait affecter les raisons réelles d'enrôlement. Par contre, la méthode de randomisation minimise ce problème en faisant en sorte que les groupes comparables soient autant similaires que possible (Thornton et al., 189).

5.3 Explication des facteurs causals

En Indonésie, l'évaluation randomisée a permis de montrer que les subventions permettaient d'hausser l'enrôlement de manière temporaire (durant l'offre de subventions) avec des effets sur le long terme (effet de l'expérience positif de l'assurance). Ce qui permet de montrer que les différents incitatifs, à travers chaque coefficient (voir section 4.1.3), ont une importance significative soutenant la validité interne de l'évaluation. Par contre, on n'a aucune idée précise de ce qui a motivé les bénéficiaires des subventions gratuites de rester plus longtemps après l'expiration de cette offre. Comment définir l'expérience positive des ménages lorsque les avis peuvent diverger d'un ménage à un autre? On ne peut, de manière confidentielle, déceler les éléments qui ont réellement influencé le choix des ménages. Nous voyons ici une inférence causale, basée sur la chaîne causale, de type X produit Y, mais sans expliquer ce qui a conduit à Y. Il se peut que les ménages aient été impacté par le corps médical ou encore par leurs pairs (Dartanto et al., 2016). L'absence de ces explications limite le pouvoir décisionnel des politiciens, qui ont besoin de savoir quels facteurs déclencher, dans un programme similaire, pour avoir de potentiels

retours positifs chez les répondants, afin d'accroître leur rétention par exemple ou minimiser la sélection adverse. Une enquête qualitative durant le processus aurait pu faciliter cette analyse et ainsi accroître la validité externe de l'évaluation (Dartanto et al. 2019).

Au **Nicaragua**, le fait que les chercheurs aient voulu comprendre pourquoi certains se sont inscrits à l'assurance et d'autres non, malgré qu'ils aient bénéficié des offres de subventions de 6 mois, permet de mieux clarifier les données quantitatives. Par exemple, ceux qui ont souscrits à la couverture avec l'offre de subventions, visaient également à protéger leur plus jeune enfant, se protéger en cas d'urgence médical et réduire les coûts financiers liés aux visites médicales (Thornton et al, 2019, p.200). Le refus de s'inscrire est en majorité lié aux coûts des primes d'assurances. Dans ce cas, les données qualitatives ont majoritairement servi à corroborer des données quantitatives, en permettant de mieux comprendre l'opinion des assurés sur le rôle des IMF en tant qu'agent d'assurance santé, en soulevant quelques facteurs relationnels ayant miné l'efficacité de l'implémentation des traitements. Ceci peut aider les acteurs responsables de l'implémentation à améliorer leur image, clarifier leur rôle (surtout pour les IMF) et établir les bases d'une bonne coopération pour des futures évaluations similaires.

5.4 Le rapport à la « boîte noire »

Comme mentionné dans l'introduction, les évaluations expérimentales font faible intégration des techniques qualitatives dans leurs évaluations. C'est souvent durant la phase pré-évaluation, pour définir les objectifs, tester le protocole d'application, s'assurer de la faisabilité de l'étude, et dans la phase d'analyse des résultats, que les chercheurs font valorisation de la méthode qualitative. D'ailleurs, l'étude sur le Nicaragua montre cette utilisation de complémentarité de la méthode qualitative dans l'analyse des résultats. Les entrevues et groupes de discussion tenus, permirent de renforcer la validité des données quantitatives. Certes, ceci est important dans le cas où nous souhaitons renforcer la validité des résultats

quantitatifs face aux réalités des répondants. Par contre, cette utilisation des méthodes qualitatives ne nous permet pas d'efficacement répondre à la problématique de la « boîte noire ». Une problématique qui nécessite de questionner, par exemple, les mécanismes sous-jacents le rapport subventions gratuites et rétention des travailleurs (comme en Indonésie), ou encore la faible confiance des répondants aux institutions micro-financières et leur décision de ne pas s'y inscrire en majorité (comme au Nicaragua). Ces mécanismes permettront, si décelés et analysés, de soit améliorer le design du programme pour des futures recherches ou soit améliorer le design de l'évaluation améliorer la portée des résultats.

Pour le moment, nous constatons, dans les évaluations en sciences humaines, une faible prise en compte des potentialités des outils qualitatifs, et cela malgré les études disponibles faisant état de l'avantage des méthodes mixtes au sein des ECR (Bamberger et al., 2010; Gertler et al., 2011; Mannell, Davis, 2019). Il est important que les chercheurs, trouvent les moyens d'efficacement incorporer la MQ afin de soulever des questions importantes aux yeux des politiciens et acteurs de développement; ce qui est d'expliquer pourquoi et comment, par exemple, nous avons un faible taux de réponse face aux meilleures combinaison de traitements offerts (Indonésie) ou encore, d'analyser, de manière poussée, le rapport entre la confiance des répondants aux IMF et l'offre de services d'assurance par ces derniers (Nicaragua). Le design d'évaluation de ces deux études de cas montre à quel point les méthodes adoptées lors d'une évaluation peuvent affecter la compréhension des effets des politiques, incitatifs, mis en œuvre dans un programme. Les méthodes mixtes comme dans le cas nicaraguayen ont non seulement permis de relever l'effet de chaque traitement sur le taux d' enrôlement, mais d'explorer en plus, le lien explicatif de la nature des résultats quantitatifs.

Naila Kabeer, professeure en développement et genre à la *London School of Economics*, a fait une étude sur le projet pilot BRAC, visant à encourager l'entrepreneuriat chez les femmes vivant dans

l'extrême pauvreté dans deux villes au Bangladesh en montrant, à partir d'une étude qualitative basée sur la littérature, comment l'évaluation randomisée a failli montrer la relation de causalité entre les résultats observés et le contexte de l'étude (Kabeer, 2019). Cette étude qualitative est un exemple concret montrant comment les techniques d'enquête permettent d'aller au-delà de savoir que certaines femmes ont mieux réussi dans le projet que d'autres, en montrant quels facteurs ont également contribué aux différences observées. Cette étude parallèle au projet expérimental a nécessité une revue de la littérature pour mieux comprendre le contexte social des communautés étudiées et pouvoir en tirer un cadre d'analyse. L'analyse contextuel facilite, dans le cadre même d'une étude, la détection de variables importants pouvant être analysé ou confronté à travers la littérature.

En somme, nous retiendrons qu'en termes de validité interne et de contrôle des biais les deux études, en procédant à la randomisation sont arrivés à prouver la relation entre les différents traitements et les résultats. Cependant, au Nicaragua, les enquêtes qualitatives ont permis de mieux comprendre les différences des résultats quantitatifs. Par contre, la faible utilisation de la MQ n'a pas nullement permis d'explorer le rapport de ces résultats face aux défis liés à l'extension de la couverture au groupe visé.

6. Discussion

L'analyse politique que nous pouvons tirer à travers ces deux cas, est que pour atteindre une couverture universelle, inclusive des groupes vulnérables comme les travailleurs du secteur informel, pauvres et non-pauvres, il est essentiel que les gouvernements offrent des subventions gratuites, sur un long terme, accompagnées des services de soins de qualité, de l'accès à des infrastructures capables de répondre aux besoins des assurés et de capacités de prise en charges par le corps médical. Dans le cas indonésien et nicaraguayen, nous avons pu constater que les subventions gratuites avaient plus d'impacts positifs sur le taux d' enrôlement comparativement aux groupes de contrôle. De plus, sur le long terme, les

bénéficiaires en Indonésie semblaient rester plus longtemps couverts grâce à l'expérience positive tirée de l'avantage de posséder une assurance santé. Cependant, dans les deux cas, la rétention tend à retourner à la baisse lorsque l'offre des subventions gratuites prend fin.

L'Organisation mondiale de la santé (OMS) défend que les soins de santé sont un bien commun et non un produit commercial, (OMS, 2008, p.12) dont l'objet est la promotion de la justice sociale. En 2012, l'ancienne directrice générale de l'OMS soutenait que «la couverture universelle en matière de santé constitue le concept le plus efficace que la santé publique puisse offrir » (Nauleau et al., 2013, p.130). Ce qui signifie que, offrir l'accès à un système de couverture sanitaire universel pour tous, indépendamment de la situation économique, sociale, géographique, et en prenant en compte les déterminants sociaux de la santé, pourraient améliorer les conditions de milliers de personnes. Dans les pays aux revenus faibles et moyens, les dépenses en santé constituent un poids énorme pour les populations. Environ 150 millions de personnes dépensent 40% de leur revenu en santé pour (Jamison et al., 2013, p.1911). Sachant qu'environ 100 millions de personnes sont dans la pauvreté en raison des dépenses directes en santé (OMS, p.12), les gouvernements visant la réduction des inégalités et le bien-être des individus devraient investir plus dans les dépenses publiques en santé. Ces dépenses publiques contribueront à baisser les dépenses directes des ménages et parallèlement, réduire les coûts psychologiques et financiers de ces derniers, tout en améliorant d'autres aspects de leur vie (revenu par exemple). En soutenant que la santé affecte la performance des travailleurs, l'améliorer pourrait également affecter leur revenu et dans le même élan contribuer à hausser le capital humain dans la production économique et donc, dans le revenu national par habitant, *ceteris paribus* (Jamison et al., p.1912).

En termes de couverture sanitaire universelle, il n'existe pas une voie unique dans la réalisation de cette dernière. Nombre de pays avancés ont pu parvenir à des progrès considérables dans la généralisation

de la couverture à différents segments de la population (Suède, Finlande, Japon, Vietnam, France) tout en prenant en compte leur contexte particulier (Nauleau et al., p.134). Dans la Recommandation 202 de l'Organisation internationale du travail (OIT), et comme défendu par l'OMS, la couverture sociale devrait viser à réduire les dépenses directes en santé des populations et leur niveau de pauvreté. Pour ce faire, deux modèles de financement sont adoptés selon les pays: le financement public, à partir des recettes fiscales générales, et le modèle d'assurance-maladie universelle, pour lequel des caisses de cotisation obligatoires des travailleurs, entreprises et pouvoirs publics sont instaurées (Nauleau et al., p.143).

Des pays comme la Norvège, la Suède, et la Finlande (Abrahamson, 2005) ont adopté des politiques universalistes de protection sociale, sous le modèle d'État-providence et d'économie de marché (avec un rôle important du privé), en s'assurant que tous les individus aient un accès équitable aux services de santé grâce à des politiques budgétaires expansionnistes qui visent à accroître l'activité économique, mais qui nécessitent également des mesures fiscales importantes avec la hausse de la dette publique. Ces pays très dépensiers en programmes sociaux tirent leur avantage des impôts et politiques fiscales ainsi que de la mise en place de programmes d'emplois et de gestion du chômage très efficace. Par ailleurs, ces modèles, presque parfaits, ne sont pas faciles à promouvoir dans des pays où les politiques fiscales et budgétaires sont très faibles et où l'économie informelle prévaut.

D'autres pays, comme le Japon, la France, la Thaïlande, le Vietnam, le Ghana, le Pérou, ont atteint ou sont presque arrivés à atteindre les groupes exclus et vulnérables grâce à des mesures d'allocations et de redistribution des ressources, de subventions importantes de programmes visant certaines catégories sociales (Maeda, et al., 2015, p.49). Le gouvernement nippon a, par exemple, grandement subventionné les programmes municipaux qui géraient les assurances maladies citoyennes qui ont mené à l'atteinte de la couverture universelle en 1961 (Maeda et al., p.24). Les efforts de croissance économique, visant à

accroître le revenu des ménages ont rendu abordable le paiement des cotisations au système d'assurance (par les ménages), et l'accroissement des fonds alloués à la santé (par l'État) a contribué à l'expansion de la couverture santé. Par contre, l'atteinte de la CSU au Japon a nécessité une intervention importante du gouvernement à travers l'harmonisation des avantages de l'assurance pour les différents groupes couverts, une redistribution des ressources financières entre différents régimes, principalement vers les plus vulnérables, la création d'un système unique de paiement de la couverture et des barèmes tarifaires établis par le gouvernement et soumis à tous les acteurs (privés inclus) de prestations des services médicaux (Maeda et al., p.25-26).

Bien qu'aucun État ne puisse couvrir à 100% sa population ni même assurer la totalité des coûts des services et besoins, car devant clarifier ces trois objectifs : la part de population à couvrir; les types de services à offrir; et la part des coûts à assurer (Nauleau et al., p.145), l'État a un rôle central à jouer. En effet, espérer une couverture globale nécessitera que les pouvoirs publics mettent en place des mesures de régulations des prix des biens, des services et des primes d'assurances; des mesures de paiement du corps médical; de gestion des infrastructures; des mesures fiscales qui permettront de soutenir les recettes publiques et donc les dépenses publiques importantes vers le secteur de la santé, en prenant en compte les groupes les plus vulnérables. Ces dépenses devront être soutenues sur le long terme et accompagnées de l'accès aux services de soins primaires, des soins curatifs et de prévention des maladies. Au plan international, la santé étant une problématique globale et de développement, les organisations internationales, les partenaires privés et fondations devront apporter leurs soutiens autant financiers que techniques dans l'atteinte de la couverture sanitaire universelle. Notons que la prise en compte des déterminants sociaux de la santé servira à un meilleur design des programmes d'assurance pour les pauvres (Wagstaff, 2002), car prenant en compte le contexte et les besoins des bénéficiaires, comme le montre l'analyse au Nicaragua.

7. Recommandations

Suite à ces observations sur nos études de cas, les recommandations suivantes, inspirés de travaux sur les essais expérimentaux en santé, peuvent être considérée pour guider l'adoption de futures évaluations expérimentales incorporant des méthodes qualitatives:

Dans la phase de planification et de design:

- Considérer les méthodes qualitatives afin qu'elles permettent d'explorer les défis contextuels et techniques, que l'essai expérimental ne pourra pas déceler tout seul, et qui peuvent être des plausibles biais affectant l'évaluation et la validité externe.
- Identifier clairement la théorie sous-jacente l'évaluation en cours et son importance pour celle-ci.

Dans la phase intermédiaire

- Faire recours à des évaluations de processus, basées sur des techniques qualitatives de collecte et d'analyse des données, qui serviront à détecter les mécanismes causals entre le traitement et les résultats dans le but d'améliorer le processus d'évaluation, mais aussi développer un cadre analytique qui servira dans des évaluations futures similaires.
- Promouvoir sciemment d'autres outils qualitatifs, au-delà des entrevues, des questionnaires et des groupes de discussion, en explorant leur apport potentiel pour une meilleure compréhension des phénomènes pouvant affecter les résultats des évaluations.

Dans la phase d'évaluation

- Produire des analyses qualitatives qui décrivent comment les résultats des évaluations peuvent servir aux politiciens (et chercheurs) voulant étendre cette étude dans un contexte différent en analysant le contexte de l'étude et les mécanismes sous-jacents un traitement et ses effets.

- Discuter des limites de l'étude et montrant comment les méthodes qualitatives ont contribué à l'évaluation expérimentale et les défis à prendre en compte dans le cadre de ce type d'évaluation.

8. Conclusion

Dans cette étude, nous avons posé la question de savoir comment les méthodes qualitatives impactent la validité externe des essais contrôlés randomisés ainsi que leur portée? Pour répondre à cette question, nous sommes partis de l'hypothèse que les évaluations contrôlées randomisées font face à des limites lorsque vient le besoin d'expliquer les causes ayant menées à un résultat précis. Tout d'abord, nous avons essayé de comprendre la nature *gold standard* des évaluations expérimentales randomisées en analysant le concept de randomisation sous la bannière de *evidence-based medicine* et de *evidence-based policy*, ainsi que dans le cadre du développement international. Un virage fut constaté dans la pratique en médecine, dans la formulation et l'adoption des politiques, ainsi que dans l'évaluation des programmes de développement, où dans chaque cas, la recherche de la meilleure évidence était devenue centrale. Cependant, ce virage propulsé par la randomisation n'est pas arrivé à résoudre la complexité des enjeux intéressant les décideurs. En effet, la randomisation, malgré sa capacité à induire une forte validité interne et à inférer des relations causales, faille à efficacement guider les décideurs politiques et les chercheurs lorsque vient la nécessité de comprendre comment extrapoler les résultats d'une étude dans un contexte différent; d'où le besoin d'un cadre théorique et des méthodes explorant les facteurs contextuels de l'évaluation en cours (Deaton, Cartwright, 2017; Bamberger et al., 2010; Gertler et al. 2011).

Face aux limites soulevées, les méthodes qualitatives, fortes de leur capacité à inférer des explications causales permettant de mieux comprendre la nature des résultats expérimentaux ainsi que leur rapport au contexte d'évaluation, paraissent comme solution dans le design des essais randomisés. En effet, ces méthodes permettent de savoir pourquoi les répondants d'une étude font certains choix; décrivent

les éléments contextuels pouvant affecter les résultats d'une évaluation: aléa politique ou rapport conflictuel entre les acteurs d'implémentation d'un programme; et relèvent les mécanismes liant une cause à un effet.

À travers nos études de cas, nous avons pu observer la relation causale entre des subventions gratuites et le taux d' enrôlement à l'assurance santé. En Indonésie, par exemple, offrir une année de subvention gratuite d'un an et une assistance technique à l'inscription en ligne ont mené à une hausse de 30% de l' enrôlement avec une rétention importante chez ceux ayant reçu ces subventions gratuites. Par contre, cette étude n'a pas explicité les facteurs qui ont incité les bénéficiaires à garder plus longtemps leur couverture; surtout qu'ils paraissent avoir les plus faibles taux de réclamations sur le long terme. Par ailleurs, aucun élément de l'étude ne fait référence aux facteurs économiques, sociaux, culturels, ni même politiques pouvant influencer sur le programme en cours, ou, clarifier les différences d' enrôlement et de rétention entre les différents groupes de traitements.

Au Nicaragua, en plus que la randomisation ait montré la relation entre les subventions, l'inscription sur place sur et l' enrôlement, les enquêtes qualitatives ont permis de mieux comprendre les données quantitatives. En effet, des relations conflictuelles, entre les partenaires d'implémentation, et la faible confiance de certains répondants vis-à-vis des institutions micro-financières, ont miné la promotion du programme et du coup, le taux d' enrôlement. Par contre, les méthodes qualitatives, en étant simplement un complément aux résultats quantitatives, ont manqué d'offrir une analyse détaillée du rapport entre la faible adhérence à la couverture et la faible rétention à l'assurance; ceci aurait pu être une opportunité de proposer des solutions d'amélioration de la couverture santé pour les travailleurs du secteur informel, comme les marchands de cette étude.

Par ailleurs, en termes politiques, l'atteinte de la couverture universelle à toutes les catégories sociales, spécialement les travailleurs du secteur informel ici, nécessitera des subventions importantes sur le long terme. Les gouvernements devront hausser leur dépense publique en santé pour également assurer l'accès équitable aux infrastructures de santé, la promotion des soins de qualité, la disponibilité d'un personnel de santé bien formé et suffisant pour répondre à la demande, le financement suffisant et équilibré des biens et services médicaux. Ils doivent également s'assurer d'adopter des mesures flexibles et adaptables aux contextes de chaque zone visée par les programmes d'assurances santé; cela doit inclure une approche basée sur les déterminants sociaux de la santé détectée à travers des études utilisant des approches mixtes et efficaces dans l'identification de facteurs pouvant affecter le design des programmes d'assurance.

En définitive, on peut affirmer que les méthodes qualitatives offrent une explication claire des facteurs sous-jacents des résultats quantitatives des évaluations randomisées en corroborant les données des deux types de méthodes. Ce qui contribue à l'extrapolation de la validité externe. Cependant, la faible valorisation des méthodes qualitatives n'a pas permis d'explorer efficacement le contenu de la « boîte noire » (*black box*). Cette limite reflète la complexité liée à la mixité des méthodes quantitatives et qualitatives dans le design des essais contrôlés randomisés. Résoudre celle-ci, comme le montre Mannell et Davis (2019), nécessitera une objectivité des évaluateurs face aux valeurs *gold standard*. Cela, en reconnaissant les avantages d'autres méthodes, dans l'offre d'explications causales entre traitements et effets, renforçant la validité externe, et dans le même sens, soutenant la formulation de politiques concrètes, basées sur des solutions précises. Il ne s'agit pas de sacrifier la validité interne des évaluations randomisées, mais de plutôt tendre vers un renforcement des deux types de validité, interne et externe, à travers des designs plus inclusifs des avantages de méthodes d'évaluations variées.

Bibliographie

Abrahamson, P. (2005). La fin du modèle scandinave ? La réforme de la protection sociale dans les pays nordiques. *Revue française des affaires sociales*, p.105-127. doi:10.3917/rfas.053.0105.

Administration on Children, Youth & Families (ACYF). (2016). Qualitative Research Methods in Program Evaluation: Considerations for Federal Staff. *Office of Data, Analysis, Research & Evaluation*, p.1-38.

Aji, P. (2015). Summary of Indonesia's Poverty Analysis. *ADB Papers On Indonesia*, 4, p.2-6.

Aldebert, B., Rouzies, A. (2015). Quelle place pour les méthodes mixtes dans la recherche francophone en management ? *Management international*, 19(1), p.43-60.

Amann, J. (2012). Chapitre 5. La philosophie de l'essai clinique selon Austin Bradford Hill. Dans : Anne Fagot-Largeault éd., *L'émergence de la médecine scientifique* (pp. 113-132). Paris: Editions Matériologiques. doi:10.3917/edmat.fagot.2012.01.0113.

Armitage, P. (2003). Fisher, Bradford Hill, and randomization. *International Journal of Epidemiology*, 32(6), p.925-928.

Augustina, R., et al., (2019). Universal health coverage in Indonesia: concept, progress, and challenges. *Lancet*, 393, p.75-102.

Bamberget, M., et al., (2010). Using Mixed Methods in Monitoring and Evaluation. Experiences from International Development. *Policy Research Working Paper*, 5245, p.1-28.

Banerjee, A, Duflo, E. (2012). *Repenser la pauvreté*. France : Édition du seuil.

Banerjee, Vinayak A., et al. (2016). The Influence of Randomized Controlled Trials on Development Economics Research and on Development Policy. *The state of Economics, the States of the World Conference at the World Bank*, p.1-76.

Banerjee, A., et al. (2019). The challenges of universal health insurance in developing countries: Evidence from a large-scale randomized experiment in Indonesia. *J-PAL*, p.1-41.

https://www.povertyactionlab.org/sites/default/files/2020-05/Subsidies-and-the-Dynamics-of-Selection_Indonesia_Banerjee-et-al_April2020.pdf (visité le 16 juin 2020).

Banque Mondiale (2008). Nicaragua Poverty Assessment. Volume 1, numéro 39736. <http://documents.worldbank.org/curated/en/235491468297893170/pdf/397360ESW0vol110gray0cover01PUBLIC1.pdf> (visité le 16 juin 2020).

Banque mondiale (2018). Contexte en Indonésie. <https://donnees.banquemondiale.org/pays/indonesie> (visité le 16 juin 2020).

Banque mondiale (2016). Contexte Nicaragua. <https://donnees.banquemondiale.org/pays/nicaragua> (visité le 16 juin 2020).

Baqir, F. et al. (2020). *Better Spending for Localizing Global Sustainable Development Goals. Examples From the Field*. New York: Routledge.

Baron, J. (2018). A Brief history of Evidence-Based Policy. *AAPSS*, 678, p.40-50.

Bédard, P-O., Ouimet, M. (2016). Persistent misunderstandings about evidence-based (sorry: informed!) policy-making. *Archives of Public Health*, 74(31), p.1-6.

Bickman, L., Reich, Stephanie M. (2017). Bickman, L. & Reich, S. (2015). Randomized controlled trials. In Donaldson, S., Christie, C., & Mark, M. *Credible and actionable evidence* (pp. 83-113). Thousand Oaks, CA: SAGE Publications, Inc doi: 10.4135/9781483385839.

Bouffard, M. (2012). Pour une pratique fondée sur des données probantes “Evidence Based Practice“ (EBP) en activité physique adaptée : perspectives et enjeux. *La nouvelle revue de l'adaptation et de la scolarisation*, 58(2), p.77-88.

Campbell, D., et al. (2002). Experiments and Generalized Causal Inference & A Critical Assessment of Our Assumption. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference* (pp. 1-81). New York, Boston: Houghton Mifflin Company.

Campbell, Donald T. (1969). Reforms as experiments. *American Psychologist*, 24(4), p.409-429.

Chalmers, I. (2005). If evidence-informed policy works in practice, does it matter if it doesn't work in theory? *Policy Press*, 1(2), p.227-242.

Chatterji, M. (2005). Evidence on “What Works”: An Argument for Extended-Term Mixed-Method (ETMM) Evaluation Designs. *Educational Researcher*, 34(5), p. 14-24.

Davis, K., et al. (2019). Beyond interviews and focus groups: a framework for integrating innovative qualitative methods into randomised controlled trials of complex public health interventions. *Trials*, 20(329), p.1-16.

Dartanto, T. et al., (2016). Participation of Informal Sector Workers in Indonesia's National Health Insurance System. *Journal of Southeast Asian Economies*, 33(3), p.317-342.

Dartanto, T. et al. (2019). Enrolment of Informal Sector Workers on the National Health Insurance System in Indonesia: A Qualitative Analysis. *LPEM-FEBUI Working Paper*, 033, p.1-14.

Dartanto, T. et al. (2020). Why Do Informal Sector Workers Not Pay the Premium Regularly? Evidence from the National Health Insurance System in Indonesia. *Health Economics and Health Policy*, 18, p.81-96.

Deaton, A. (2010). Instruments, Randomization, and Learning about Development. *Journal of Economic Literature*, 48(2), p.424-455.

- Deaton, A., Cartwright, N. (2017). Understanding and Misunderstanding Randomized Controlled Trials. *NBER Working Paper Series*, 22595, p.1-70.
- De Souza Leão, L., Eyal, G. (2019). The rise of randomized controlled trials (RCTs) in international development in historical perspective. *Theory and Society*, 48, p.383-418.
- Dhillon, L., Vaca, S. (2018). Refining Theories of change. *Journal of MultiDisciplinary Evaluation*, 14(30), p.65-87.
- Doll, R. (1992). Sir Austin Bradford Hill and the progress of medical science. *British Medical Journal*, 305, p.1521-1526.
- Drapeau, M. (2004). Réflexion épistémologique sur la recherche qualitative et la psychanalyse : refaire une place au rêve et à l'imaginaire. *Le Coq-héron*, no 177(2), 124-129. doi:10.3917/cohe.177.0124.
- Drisko, James, W. (1997). Strengthening Qualitative Studies and Reports: Standards to Promote Academic Integrity. *Journal of Social Work Education*, 33(1), p.185-197.
- Dumez, H. (2011). Qu'est-ce que la recherche qualitative ? *AEGIS Le libellio*, 7(4), p.47-58.
- Erlangga, D., et al. (2018). The impact of public health insurance on healthcare utilisation in Indonesia: evidence from panel data. *International Journal of public Health*, 64, p.603-613.
- Favereau, J. (2016). On the analogy between field experiments in economics and clinical trials in medicine. *Journal of Economic Methodology*, 23(2), p.203-222.
- Fisher, Lloyd D. (1999). Advances in Clinical Trials in The Twentieth Century. *Public Health*, 20, p.109-124.
- Flemming, K., et al. (2008). Improving the effectiveness of interventions in palliative care: the potential role of qualitative research in enhancing evidence from randomized controlled trials. *Palliative Medicine*, 22, p.123-131.
- Fox, D. (2005). Evidence of Evidence-Based Health Policy: The Politics Of Systematic Reviews In Coverage Decisions. *Health Affairs*, 24(1), p.1-9.
- Gauthier, B. (2009). *Recherche Sociale : de la problématique à la collecte des données* (5^e éd.) Québec : Presses de l'Université du Québec.
- Gertler, Paul J., et al. (2011). L'évaluation d'impact en pratique. *Banque internationale pour la reconstruction et le développement/Banque mondiale*, p.1-223.
- Guyatt, G. et al. (1992). Evidence-Based Medicine. A New Approach to Teaching the Practice of Medicine. *JAMA*, 268(17), p.2420-2425.

Greene, Jennifer C., et al. (1989). Toward a Conceptual Framework for Mixed-Method Evaluation Designs. *SAGE Journals*, 11(3), p.255-274.

Hammarberg, K., et al. (2015). Qualitative research methods: when to use them and how to judge them. *Human Reproduction*, 31(3), p.498-501.

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. *The policy Press*, 1(1), p.85-100.

Haskins, R., Margolis, G. (2014). Introduction: The Obama Strategy for Attacking Social Problems. In: Haskins, R., Margolis. *Show me the Evidence Obama's Fight for Rigor and Results in Social Policy* (p.1-30). Brookings Institution Press: Project MUSE. muse.jhu.edu/book/36692.

Hatt, L., et al., (2009). Extending Social Insurance to Informal Sector Workers in Nicaragua via Microfinance Institutions: Results from a Randomized Evaluation. *Abt Associates Inc*, p.1-63. https://www.abtassociates.com/sites/default/files/migrated_files/4ce3e5d2-0545-4f2e-a8af-efa9cad11d3a.pdf (visité le 16 juin 2020).

Head, Brian W. (2010). Reconsidering evidence-based policy: Key issues and challenges. *Policy and Society*, 29(2), p.77-94.

International Monetary Fund. (2011). Nicaragua: Poverty Reduction Strategy Paper— Progress Report on National Human Development Plan as of 2010. <https://www.imf.org/external/pubs/ft/scr/2011/cr11323.pdf> (visité le 16 juin 2020)

Jamison, Dean T., et al. (2013). Global health 20135: a world converging within a generation. *The Lancet*, 382(9908), p.1898-1955.

Jatteau, A. (2013). Expérimenter le développement ? *Geneses*, 93(4), p.8-28.

Jatteau, A. (2019). Les essais contrôlés randomisés. Une comparaison entre la médecine et l'économie. *Philosophia Scientiae*, 23(2), p.85-110.

Kabeer, N. 2019. Randomized Control Trials and Qualitative Evaluations of a Multifaceted Programme for Women in Extreme Poverty: Empirical Findings and Methodological Reflections. *Journal of Human Development and Capabilities*, 20(2), p.197-217.

Laurent, C., et al. (2009). Pourquoi s'intéresser à la notion d' « evidence-based policy » ? *Revue Tiers Monde*, 200(4), p.853-873.

Levy, D., et al. (2020). Insights from Transparency and Accountability Action Plans in Indonesia and Tanzania. *ASH Center for Democratic Governance and Innovation*, p.1-79.

Ly, C. (2018). Essays On Universal Health Coverage In Indonesia. *Publicly Accessible Penn Dissertations*, p.1-137.

Maeda A., et al. (2015). *Une couverture sanitaire universelle pour un développement durable inclusif : Une synthèse de 11 études de cas pays*. Directions du développement. Washington, DC : La Banque mondiale. doi : 10.1596/978-1-4648-0592-9.

Mannell, J, Davis, K. (2019). Evaluating complex health interventions with randomized controlled trials: How do we improve the use of qualitative methods? *Qualitative Health Research*, 20(5), p.623-631.

Mathaeur, I. et al. (2008). Extending social health insurance to the informal sector in Kenya. An assessment of factors affecting demand. *Int J Health Plann Mgmt*, 23, p.51-68.

Murtagh, M.J, et al. (2007). Qualitative methods in a randomised controlled trial: the role of an integrated qualitative process evaluation in providing evidence to discontinue the intervention in one arm of a trial of a decision support tool. *Qual Saf Health Care*, 16, p.224-229.

Nauleau, M et al. (2013). En chemin vers la couverture sanitaire universelle : Les enjeux de l'intégration des oeuvres aux systèmes de santé. *Revue tiers Monde*, 215, p.129-148.

O'Cathain, A., et al. (2014). Maximising the value of combining qualitative research and randomised controlled trials in health research: the Qualitative Research in Trials (QUART) study – a mixed methods study. *ISSN*, 18(38), p.1-123.

OECD (2019), Social Protection System Review of Indonesia, OECD Development Pathways, OECD Publishing, Paris. <https://doi.org/10.1787/788e9d71-en>. (visité le 16 juin 2020).

Onwuegbuzie, Anthony J., et Collins, Kathleen M.T. (2007). A Typology of Mixed Methods Sampling Designs in Social Science Research. *The qualitative Report*, 12(2), p.281-316.

Organisation mondiale de la santé (2008). Comblent le fossé en une génération. Instauration l'équité en santé en agissant sur les déterminants sociaux de la santé. *Commission des déterminants sociaux de la santé*. https://www.who.int/social_determinants/final_report/csdh_finalreport_2008_execsumm_fr.pdf (visité le 03 juillet 2020).

Pearce, W., et al. (2015). Randomised trials in context: practical problems and social aspects of evidence-based medicine and policy. *BioMed Central*, 16(1), p.1-7.

Pearce, W., Raman, S. (2014). The new randomised controlled trials (RCT) movement in public policy: challenges of epistemic governance. *Policy Sci*, 47, p.387-402. DOI 10.1007/s11077-014-9208-3

Perrenoud, P. (2014). *Naissance et évolution des pratiques : entre Evidence-Based Medicine, expérience et intuition*. Dans : Claudine Burton-Jeangros éd., *Accompagner la naissance: Terrains socio-anthropologiques en Suisse romande* (pp. 133-154). Bangkok, Thailand: BSN Press.

Picciotto, R. (2012). Experimentalism and development evaluation: Will the bubble burst? *Sage*, 18(2), p.213-229.

Prabhakaran, S., et al., (2019). Financial Sustainability of Indonesia's Jaminan Kesehatan Nasional: Performance, Prospects, and Policy Options. *Washington, DC: Palladium, Health Policy Plus, and Jakarta, Indonesia: Tim Nasional Percepatan Penanggulangan Kemiskinan (TNP2K)*, p.1.85.

Rallis, Sharon, F. (2017). When and How Qualitative Methods Provide Credible and Actionable Evidence: Reasoning With Rigor, Probity, and Transparency. In: Donaldson, Stewart I, et al. *Evaluations* (second Edition), *Credible and Actionable Evidence: The Foundation for Rigorous and Influential* (p.1-17), Sage Research Methods.

Ravallion, M. (2009). Evaluation in the Practice of Development. *The World Bank Research Observer*, 24(1), p.29-53.

Rodrik, D. (2008). The Development Economics: We Shall Experiment, but How Shall We Learn? *HKS Working Paper No.RWP08-055*, p.1-38.

Sackett, David L. et al. (1996). Evidence based medicine: what it is and what it isn't. *British Medical Journal*, 312(7023), p.71-72.

Sehon, Scott R., Stanley, Donald E. (2003). A philosophical analysis of the evidence-based medicine debate. *BMC Health Services Research*, 3(1), p.1-10.

Silverman, William A., et al. (1992). Sir Bradford Hill: An Appreciation. *Controlled clinical Trials*, 13, p.100-105.

Spillane, James P., et al. (2010). Mixing methods in randomized controlled trials (RCTs): Validation, contextualization, triangulation, and control. *Educational Assessment, Evaluation and Accountability*, 22(1), p.5-28.

Stame, N. (2004). Theory-based Evaluation and Types of Complexity. *Sage Publications*, 10(1), p.58-76.

Statistik Indonesia. (2020). <https://www.bps.go.id/publication.html?page=4> (visité le 16 juin 2020)

Stern, E., et al., (2012). Broadening The Range Of Designs And Methods For Impact Evaluations. Report of Department for International Development (DFID), *Working paper*, 38, p.1-87.

Thornton, Rebecca L., et al., (2010). Social Security Health Insurance For The Informal Sector In Nicaragua: A Randomized Evaluation. *Wiley InterScience*, 19, p.181-206.

United Nations Programme for Development (2019). Beyond income, beyond averages, beyond today: Inequalities in human development in the 21st century. UNDP, p.1-352. <https://reliefweb.int/sites/reliefweb.int/files/resources/hdr2019.pdf> (visité le 16 juin 2020).

Vogel, I. (2012). Review of the use of 'Theory of Change' in international development. Department for International Development (DFID), p.2-82.

Wagstaff, A. (2002). Pauvreté et inégalités dans le secteur de la santé. *Bulletin of the World Health Organization*, 80(2): 100–108.

White, H. (2013). The Use of Mixed Methods in Randomized Control Trials. *Wiley Periodicals*, 138, p. 61-73.

Whitley, R., et al. (2011). Evidence-Based Medicine: Opportunities and Challenges in a Diverse Society. *The Canadian Journal of Psychiatry*, 56(9), p514-522.

Worall, J (2007). Evidence in Medicine and Evidence-Based Medicine. *Philosophy Compass*, 2(7), p.981-1022.

Annexes

Tableau 1: Apport des méthodes qualitatives au sein des évaluations contrôlés randomisées

Catégories adaptées à partir de CONSORT	Défis méthodologiques	Menaces au design ou aux résultats	Solutions apportées par la recherche qualitative	Exemples de méthodes qualitatives envisageables
Contexte	Assurer la pertinence du problème dans le contexte	Affecte l'applicabilité, l'acceptabilité, la transférabilité et la durabilité de potentiels résultats positifs	Identifier les facteurs sociaux, culturels, économiques, politiques et de santé qui pourraient affecter	Observation directe approfondie dans le cadre d'une évaluation ethnographique rapide ou d'une enquête à grande échelle
Design de l'intervention et conformité	Assurer l'acceptabilité culturelle and la faisabilité pratique de l'intervention	Affecte l'adhésion et hausse le nombre d'abandons	Adapter l'intervention afin d'hausser la rétention et l'adhésion	Méthodes d'agenda – via des messages vocaux interactifs ou des SMS
Recrutement et enrôlement	Garantir la représentativité de l'échantillon et des méthodes de recrutement efficient	Risques liés à la taille de l'échantillon afin de détecter des effets significatifs	Déterminer la meilleure méthode possible de recrutement afin d'atteindre la population cible	Cartographie communautaire
Randomisation et allocation	Garantir une randomisation équilibrée (égale)	Réduit la comparabilité entre les groupes	Identifier des facteurs contextuels qui pourraient affecter l'effet de l'intervention and réduire les différences entre les groupes évalués	Observation/randomisation publique
Participants et suivi	Minimiser le nombre de participants quittant l'étude	Induit une perte inégale de participants entre les groupes, ce qui affecte l'inférence causale	Pour prévenir ou comprendre les raisons de perte dans le suivi et pour améliorer les stratégies de rétention	Tenue d'un agenda, utilisation des messages vocaux interactifs par SMS ou WhatsApp soutien des pairs pour l'adhésion
Collecte des données	Maintenir une objectivité durant le processus de collecte des données	Menace la validité des données collectées et/ou mesurées afin de répondre à l'objectif de l'étude	Évaluer et valider le processus de collecte des données	Recours aux groupes de discussion utilisant des cartes de mémorisation ou des images; des entretiens en binôme et jeux de rôle
Analyses et résultats	Identifier les mécanismes sous-tendant l'effet de l'intervention	Limite une discussion éclairée des résultats (positifs ou négatifs)	Triangler les résultats quantitatifs et identifier les données contextuelles qui pourraient affecter les résultats	Méthode d'analyse participative

Tableau 2: Répartition des groupes de traitements par ville

Medan					
	Status quo		Assisted internet registration		Subsidy treatment totals
	Standard information	Extra information	Standard information	Extra information	
No subsidy	37	63	134	237	471
Half subsidy	171	215	26	66	478
Full subsidy	176	54	170	97	497
Registration treatment totals	716		730		

Bandung						
		Status quo		Assisted internet registration		Subsidy treatment totals
		Standard information	Mandate information	Standard information	Mandate information	
No subsidy	Standard information	236	307	241	274	2236
	Waiting period information	232	300	297	349	
Half subsidy	Standard information	160	153	77	82	918
	Waiting period information	141	114	100	91	
Full subsidy	Standard information	85	40	62	54	478
	Waiting period information	63	51	70	53	
Bonus subsidy	Standard information	114	86	170	111	918
	Waiting period information	101	86	131	119	
Registration treatment totals		2269		2281		

Tableau 3: Effets des subventions et de l'assistance internet sur l'enrôlement

	Decomposition			
	Enrolled within 1 year	Attempted to enroll within 8 weeks of offer date	Enrolled within 8 weeks of offer date	Enrolled after 8 weeks, but within 1 year of offer date
	(1)	(2)	(3)	(4)
Panel A: Main effects				
Full subsidy	0.186*** (0.020)	0.276*** (0.020)	0.209*** (0.018)	-0.023** (0.010)
Half subsidy	0.100*** (0.014)	0.161*** (0.016)	0.114*** (0.013)	-0.014* (0.008)
Assisted internet registration	0.035*** (0.011)	0.237*** (0.011)	0.043*** (0.009)	-0.008 (0.006)
No subsidy mean	0.086	0.099	0.030	0.056
P-value of test of hypothesis				
Half subsidy = full subsidy	0.000	0.000	0.000	0.442
Assisted internet registration = full subsidy	0.000	0.107	0.000	0.266

Tableau 4: Caractéristiques des participants à l'étude

	All (N = 2608)		MFI client (N = 1013)		Non-MFI client (N = 1595)		Difference
	Mean	SD	Mean	SD	Mean	SD	
<i>Panel A: Demographic characteristics</i>							
Age	37.79	9.28	39.05	8.57	36.98	9.63	2.07***
Male	0.35	0.48	0.28	0.45	0.39	0.49	-0.11***
Years education	9.26	4.22	9.25	4.15	9.27	4.27	-0.02
Married	0.70	0.46	0.70	0.46	0.70	0.46	0.00
Number of children	2.04	1.52	2.22	1.52	1.93	1.52	0.29***
With under 12 children	0.66	0.47	0.66	0.48	0.67	0.47	-0.01
Pregnant	0.03	0.18	0.03	0.18	0.03	0.17	0.00
	All		MFI client		Non-MFI client		
<i>Panel B: Economic characteristics</i>							
Monthly income	273.89	380.34	297.22	404.12	258.77	363.45	38.44***
Proportion with savings	0.29	0.46	0.28	0.45	0.30	0.46	-0.02
Monthly savings	25.11	158.51	23.40	157.89	26.19	158.95	-2.78
Last year's savings Balance	145.73	574.94	141.45	609.66	148.45	551.85	-7.00
owns home	0.76	0.42	0.80	0.40	0.74	0.44	0.06***
Total household health-care costs	84.46	171.62	88.67	174.38	81.78	169.85	6.89
Total health-care costs for respondent	46.50	114.44	51.68	128.77	43.20	104.20	8.48*
Last visit cost for respondent	17.77	59.98	20.79	82.05	15.86	39.95	4.93**
Proportion MFI Clients	0.39	0.49					

Tableau 5: Les déterminants de l'inscription à l'assurance

Dependent variable: signed up for Insurance (0/1)	All		MFI client	Non-MFI client
	(1)	(2)	(3)	(4)
Information only	-0.049** [0.023]	-0.058** [0.028]	-0.051 [0.041]	-0.093** [0.041]
Six-month subsidy – INSS	0.333*** [0.028]	0.336*** [0.033]	0.384*** [0.051]	0.275*** [0.046]
Six-month subsidy – MFI	0.281*** [0.028]	0.277*** [0.033]	0.313*** [0.048]	0.223*** [0.047]
Male	-0.023 [0.015]	-0.024 [0.017]	-0.007 [0.030]	-0.032 [0.021]
Education	0.002 [0.002]	0.002 [0.002]	0.001 [0.003]	0.003 [0.003]
Married	0.014 [0.016]	0.023 [0.018]	-0.002 [0.029]	0.037 [0.023]
MFI client	0.033** [0.015]	0.031* [0.016]		
Number of children	-0.015*** [0.005]	-0.012** [0.006]	-0.032*** [0.010]	0.001 [0.008]
Has children under 12	0.030* [0.017]	0.035* [0.019]	0.111*** [0.030]	-0.015 [0.024]

Tableau 6: Effets de l'assurance sur les dépenses en santé et l'utilisation des services de santé

	Visited		Number of visits		Ln (amount spent)	
	Coeff.	SE	Coeff.	SE	Coeff.	SE
<i>Panel A: all (N = 2608)</i>						
Pharmacy	-0.020	[0.044]	0.468	[0.444]	-0.548	[0.344]
Private doctor	0.029	[0.038]	0.045	[0.121]	0.222	[0.262]
Laboratory	-0.067	[0.054]	-0.155	[0.208]	-1.072***	[0.365]
EMP	0.427***	[0.027]	1.542***	[0.145]	—	—
Private clinic/hospital	-0.089*	[0.050]	-0.473***	[0.181]	-0.565	[0.351]
Public (MINSA) health center	-0.063	[0.047]	-0.424*	[0.247]	0.006	[0.028]
Public (MINSA) hospital	-0.115***	[0.040]	-0.176	[0.152]	-0.025	[0.021]
All/any	-0.001	[0.034]	0.918	[0.749]	-0.522	[0.343]

Tableau 7: Résumé des deux études de cas

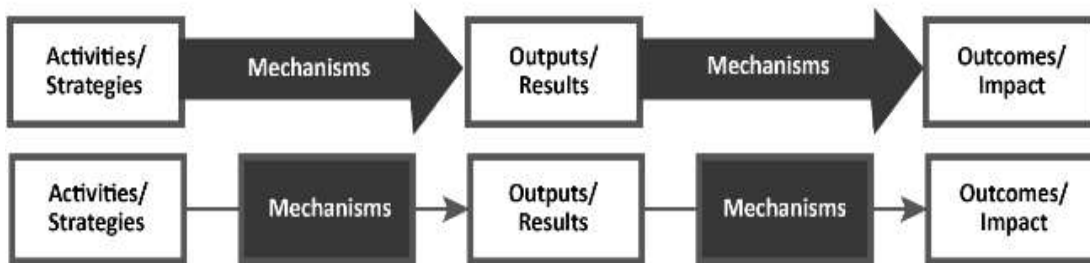
Pays	Design de l'évaluation	Groupes visés	Types de traitements offerts	Observations
Indonésie	Randomisation.	Ménages non-pauvres travaillant dans le secteur informel (aucune caractéristique distincte entre les répondants ne fut présentée par les évaluateurs).	<ul style="list-style-type: none"> -Subventions gratuites sur 1 an. -Demi-subventions sur 1 an. -Assistance à l'inscription sur internet. -La fourniture de diverses informations. -Zéro subvention (groupe de contrôle). 	<p>-Les subventions gratuites jouent un rôle important sur l'enrôlement. L'étude montre qu'il serait plus bénéfique si le gouvernement continu le soutien financier aux ménages tout en améliorant les services de santé. Sur le long terme, une expérience positive de l'assurance pourrait mener à une forte rétention à l'assurance.</p> <p>-Par ailleurs, l'absence des données supplémentaires sur les répondants a failli relever les facteurs potentiels ayant affectés le choix de certains à rester plus longtemps couverts. Une enquête qualitative aurait pu pallier ce problème.</p>

Pays	Design de l'évaluation	Groupes visés	Types de traitements offerts	Observations
Nicaragua	Randomisation plus enquête qualitative.	Vendeurs pauvres travaillant dans le secteur informel (voir tableau 4 pour les caractéristiques).	<p>-Subventions gratuites de 6 mois et inscription à l'INSS.</p> <p>-Subventions gratuites de 6 mois et inscriptions aux différents sièges des IMF.</p> <p>-Brochure informationnelle sur les produits de l'assurance santé.</p> <p>-Zéro prix (subvention pure)</p>	<p>-Comme dans le cas précédent, les subventions ont eu des effets plus positifs mais moindre, sur l'enrôlement. La différence entre l'INSS et les IMF indique deux éléments importants : les répondants préfèrent la qualité efficiente des services offerts par le IMF, mais ils ont moins confiance en elles. Ce qui indique un pourcentage élevé d'inscription chez l'INSS par rapport aux IMF (voir tableau 5).</p> <p>- L'enquête qualitative a montré que la faible coopération entre l'INSS et les IMF a en partie minée la bonne conduite du programme; ce qui affecte l'enrôlement tout comme la rétention.</p> <p>-L'enquête qualitative a également permis de valider les données quantitatives en validant la relation entre ces dernières et les réponses des répondants. Mais son usage fut très limité et moins exploré enfin d'en tirer une leçon constructive sur le design des programmes d'assurances santé.</p>

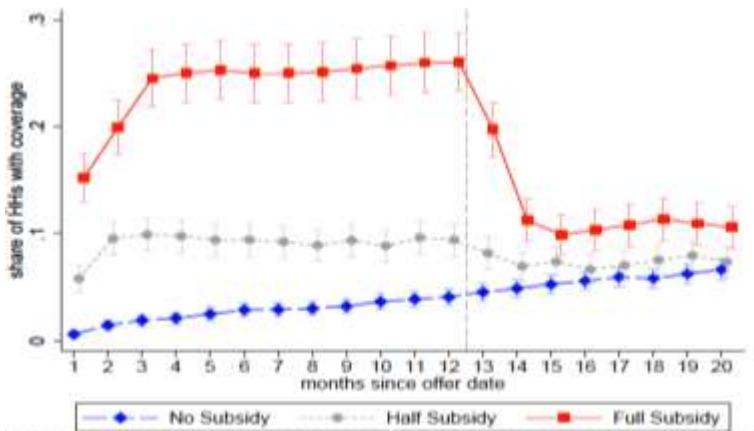
Figure 1: La théorie du changement



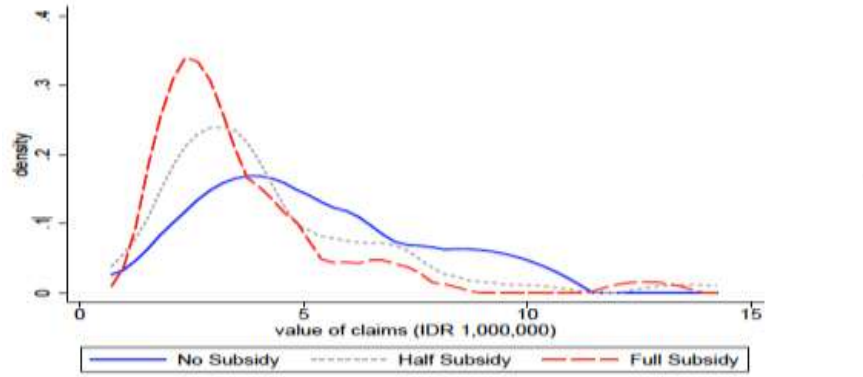
Figure 2: Les mécanismes dans une théorie du changement



Graphique 1 : Enrôlement par mois depuis le début de l'offre par type de subvention



Graphique 2: Valeurs des réclamations des frais d'hospitalisations par type de subvention



Graphique 3: Nombre de réclamations depuis le début de l'enrôlement par type de subventions

