

Contents

Résumé	1
1 Qu'est-ce qu'un plan de pré-analyse (pre-analysis plan, PAP)	1
2 Qu'y a-t-il dans un PAP? Aperçu du design de l'étude	2
3 Qu'y a-t-il dans un PAP? Hypothèses	2
4 Qu'y a-t-il dans un PAP? Mesures et construction d'indices	3
5 Qu'y a-t-il dans un PAP? Procédure d'estimation	4
6 Qu'y a-t-il dans un PAP? Critères d'inférence	4
7 Qu'y a-t-il dans un PAP? Procédures pour les problèmes de données	5
8 Qu'y a-t-il dans un PAP? Analyse de puissance statistique	6
9 Pourquoi faire un PAP ?	6
10 Quand et où enregistrer un PAP ?	8
Références	8

Résumé

Ce guide présente l'idée du plan de pré-analyse (pre-analysis plan, PAP), propose un modèle et des questions d'orientation pour rédiger des plans de pré-analyse pour vos études et explique les utilisations d'un plan de pré-analyse. Des liens vers les registres de PAP avec des exemples de plans sont fournis à la fin de ce document.

1 Qu'est-ce qu'un plan de pré-analyse (pre-analysis plan, PAP)

Un PAP est un document qui formalise et décrit le design et le plan d'analyse de votre étude. Il est rédigé avant la réalisation de l'analyse et est généralement enregistré sur un site web tiers.¹

Les objectifs du PAP sont d'améliorer les choix de design de recherche, d'accroître la transparence de la recherche et de permettre à d'autres chercheurs de reproduire votre analyse. Par conséquent, nous recommandons de concentrer le PAP sur les détails analytiques qui vous aideront à analyser votre étude et permettront à d'autres chercheurs de reproduire votre analyse. Une brève section sur la théorie devrait être incluse dans la mesure où elle aide à formuler les hypothèses, mais une analyse détaillée de la théorie et de la littérature n'a pas besoin d'être incluse. Le PAP n'a pas besoin d'inclure le début d'un article académique si ces sections n'aident pas à réfléchir sur votre analyse ou à aider les lecteurs à reproduire votre analyse.

¹ Les PAP sont encouragés dans le cadre des directives pour la promotion de la transparence et de l'ouverture (Transparency and Openness Promotion, TOP) (Nosek et al. 2015), publiées dans Science, avec les principales revues scientifiques en sciences sociales s'engageant à mettre en œuvre les directives TOP.

Dans les sections suivantes, nous fournissons des directives pour les détails que vous devez inclure dans les PAP, y compris un exemple de texte. Nous vous recommandons également d'inclure autant de code et d'analyse de données simulées que possible.² De nombreux PAP ne pourront pas inclure toute notre liste, mais un PAP devrait, au minimum, inclure la liste complète des hypothèses que vous avez l'intention de tester, comment vous allez mesurer les variables pertinentes pour ces hypothèses et un horodatage vérifiable.

2 Qu'y a-t-il dans un PAP? Aperçu du design de l'étude

La première section d'un PAP doit fournir un bref aperçu du design de votre étude. Si cette étude est une expérience, décrivez la procédure de randomisation et l'intervention ou la procédure expérimentale. Si l'étude n'est pas une expérience, décrivez les données. Ces descriptions doivent inclure : (1) l'unité d'analyse, la population et les critères d'inclusion/exclusion, (2) la méthode (observationnelle, expérimentale, quasi-expérimentale), (3) l'intervention expérimentale ou la variable explicative, et (4) les résultats d'intérêt.

Exemple: Prenons une version simplifiée d'une intervention au Malawi comme exemple. Nous utilisons une expérience randomisée par bloc pour évaluer l'effet de *la prestation d'informations et de services publics* (variable explicative) sur la *conformité fiscale* (résultat) au Malawi, où la prestation de services est faible et la non-conformité fiscale élevée. Notre unité d'analyse est un ménage propriétaire de son logement dans une ville du Malawi. Nous excluons les locataires car seuls les propriétaires paient les taxes d'habitation pertinentes pour notre étude. Nous mettons les ménages dans 80 groupes (grappes/clusters) au niveau du quartier d'environ 20 ménages chacun. Chaque quartier est également un bloc tel que 10 ménages dans chaque quartier sont assignés au traitement et 10 ménages sont assignés au contrôle. Notre intervention est la prestation d'informations et de services publics. Nous mettons en œuvre cette intervention en offrant deux collectes de déchets gratuites, une visite d'un solliciteur pour discuter de la façon dont les paiements d'impôts financent les services municipaux comme la collecte des déchets, et une brochure avec plus d'informations sur le paiement des impôts. Notre principal résultat est la conformité fiscale, que nous étudions en tant que paiement des taxes municipales. Nous menons une enquête et collectons des données administratives sur le paiement des impôts avant et après l'intervention.

3 Qu'y a-t-il dans un PAP? Hypothèses

Le PAP doit spécifier vos hypothèses - la ou les relations que vous vous attendez à observer entre les variables. La formulation d'une hypothèse doit indiquer clairement si elle implique des tests unilatéraux ou bilatéraux (c'est-à-dire prédire une augmentation, une diminution ou un changement de la variable de résultat).

Il existe deux types d'hypothèses qu'il faut envisager d'inclure dans votre PAP : confirmatoire et exploratoire.

Les hypothèses confirmatoires sont au cœur de la plupart des études ; ce sont les hypothèses que votre étude est conçue pour tester. Vos analyses pour ces hypothèses seront généralement puissantes et vous disposerez généralement d'une théorie solide menant à ces hypothèses *a priori*.

Les hypothèses exploratoires sont des hypothèses que vous souhaitez peut-être tester mais qui ne sont pas l'objectif principal de votre étude. Ce sont souvent des hypothèses secondaires sur des mécanismes, des sous-groupes, des effets hétérogènes ou des résultats en aval. Les analyses guidées par ces hypothèses peuvent ne pas être suffisamment puissantes et votre théorie peut ne pas se concentrer sur ces effets, mais l'analyse de ces hypothèses peut conduire à des découvertes surprenantes.

Certaines personnes préfèrent énumérer peu d'hypothèses et d'autres préfèrent en énumérer beaucoup. En règle générale, vous devez inclure autant d'hypothèses que celles liées à votre théorie ou intervention.³ Il

²Des ressources telles que le projet DeclareDesign peuvent vous aider à simuler et à analyser de fausses données qui imitent les données réelles que votre projet collectera. L'analyse de puissance statistique est un élément important d'un PAP et nécessite des données simulées.

³Comme R.A. Fisher a dit, souvent cité par la suite : "élaborez vos théories ... lors de la construction d'une hypothèse causale, il faut envisager autant de conséquences différentes de la vérité que possible," (Cochran, 1965; cité dans Rosenbaum

peut s'agir d'un résultat unique, mais si votre intervention ou théorie expérimentale fait des prédictions sur 8 résultats, énumérez les hypothèses pour ces 8 résultats. Si votre intervention ou théorie expérimentale suppose des mécanismes spécifiques par lesquels la variable explicative affecte un résultat ou des résultats, ces mécanismes doivent être clairs dans l'hypothèse.

Notez qu'avec plus d'une hypothèse, vous devrez spécifier une procédure pour gérer plusieurs hypothèses dans la section des critères d'inférence de votre PAP, soit en corrigeant pour plusieurs tests, soit en agrégeant toutes les hypothèses dans un indice ou un test omnibus.⁴ Si vous utilisez un test omnibus, vous pouvez lister tous les résultats sous une seule hypothèse. Consultez notre section sur les critères d'inférence pour en savoir plus sur la correction de plusieurs tests.

Exemple: Hypothèses Nous spécifions notre hypothèse confirmatoire comme :

H : Le traitement augmentera la conformité fiscale parmi les ménages de traitement par rapport aux ménages de contrôle.

En plus de l'amélioration globale des attitudes envers le gouvernement, nous nous attendons à ce que les résidents modifient leurs attitudes sur des questions spécifiques. Nous aimerions explorer quels problèmes peuvent être à l'origine des changements d'attitude globaux avec les hypothèses exploratoires suivantes :

H : Le traitement augmentera l'opinion selon laquelle la ville mérite de percevoir des impôts parmi les ménages de traitement par rapport aux ménages de contrôle.

H : Le traitement augmentera les croyances dans la capacité de la ville à fournir des services parmi les ménages de traitement par rapport aux ménages de contrôle.

H : Le traitement augmentera les croyances dans la capacité de la ville à faire appliquer les règles parmi les ménages de traitement par rapport aux ménages de contrôle.

4 Qu'y a-t-il dans un PAP? Mesures et construction d'indices

Le PAP doit spécifier la manière dont vous mesurez ou opérationnalisez les variables d'intérêt, y compris les résultats, les covariables et les variables explicatives. Cette opérationnalisation peut avoir sa propre partie ou être incluse après l'hypothèse pour laquelle chaque variable est pertinente.

Pour chaque variable, le PAP doit énumérer la manière dont la variable est mesurée (comme les questions de l'enquête ou de l'entretien, les données administratives, les mesures comportementales/observationnelles, etc.) et à quelle(s) hypothèse(s) la variable se rapporte. Les détails de ces mesures, tels que la formulation précise des questions de l'enquête, doivent être inclus soit dans cette section, soit dans une annexe. Si vous utilisez des indices ou des facteurs, ou si vous combinez des résultats d'une autre manière, précisez comment les résultats combinés seront construits. Si vous manipulez ou transformez les résultats d'une manière ou d'une autre (comme l'enregistrement d'une variable), décrivez le processus de manipulation ou de transformation.

Nous vous recommandons d'inclure du code dans votre PAP pour montrer comment vous prévoyez d'exécuter toutes les transformations de données.

Exemple : Mesures et construction d'indices Nous mesurons notre résultat principal, la *conformité fiscale*, à l'aide de données administratives sur les paiements d'impôts par les citoyens. La mesure de conformité fiscale prend la valeur 0 si le ménage n'a pas payé d'impôts et 1 si le ménage a payé des impôts. Notre variable explicative est l'assignation du traitement, où les individus assignés au traitement reçoivent 1 et les individus assignés au contrôle reçoivent 0. Nous créons également un indice "attitude envers le gouvernement" en utilisant les poids de covariance inverse (inverse covariance weighting, ICW) de 6 questions d'enquête où des valeurs plus élevées signifient des attitudes plus positives envers le gouvernement. L'indice ICW pondère

(2010), pp. 327). Bien que cela ait été dit à propos de la détermination de la causalité dans les études d'observation, la logique s'applique également aux études expérimentales.

⁴Pour un exemple de test omnibus, voir Caughey, Dafoe, and Seawright (2017).

les questions de base par la covariance des réponses dans le groupe de contrôle pré-traitement et pondère les questions post-traitement par la covariance des réponses dans le groupe de contrôle post-traitement. Nous standardisons ensuite l'échelle ICW par l'écart type de sorte qu'on puisse comparer les réponses pré- et post-traitement.

5 Qu'y a-t-il dans un PAP? Procédure d'estimation

Maintenant que vous avez décrit votre design d'étude, vos hypothèses et vos variables, vous êtes prêt à discuter de vos procédures de test et d'estimation.

Cette section doit clairement spécifier ce que vous estimez (c'est-à-dire l'estimande) et comment vous allez l'estimer (c'est-à-dire l'estimateur). Par exemple, de nombreuses études estiment l'effet moyen du traitement d'une intervention expérimentale en utilisant la régression linéaire MCO (méthode des moindres carrés ordinaires) comme estimateur.⁵ Décrivez clairement la spécification de votre modèle, y compris vos résultats, les variables explicatives, les covariables, ainsi que vos statistiques de test.

Nous recommandons d'inclure le code du modèle statistique ou la forme fonctionnelle du modèle statistique dans le PAP.

Exemple : Procédure d'estimation Nous estimons l'effet de la campagne d'information et du service de collecte des déchets sur le paiement des taxes d'habitation chez les résidents avec une analyse d'intention de traiter. Notre estimation est l'effet moyen du traitement. Si nous avons un équilibre entre les résultats de base et finaux, nous utiliserons l'estimateur suivant pour estimer l'effet moyen du traitement : $Y_{i,j} = \beta_0 + \beta_1 Z_{i,j} + X_{i,j} + \epsilon_{i,j}$ où i est l'individu dans le quartier j , Z est l'indicateur de traitement et Y est le résultat. X est le résultat de base pour l'individu i . Nous utiliserons des poids de régression proportionnels à la taille des quartiers j . Si les résultats de base et de fin ne sont pas équilibrés, nous utiliserons le score de changement, $Y_i = Y_{i,fin} - Y_{i,base}$ et nous n'utiliserons pas X .

6 Qu'y a-t-il dans un PAP? Critères d'inférence

Les critères d'inférence sont des règles de décision pour déterminer la détectabilité des effets (c'est-à-dire si une variable explicative affecte réellement une variable de résultat). L'établissement de critères d'inférence nécessite de faire plusieurs choix pour croire que l'effet estimé est "statistiquement significativement différent" de l'hypothèse nulle. Nous vous recommandons de préciser et de justifier clairement les réponses aux questions suivantes :

1. Quel type d'erreur type utiliserez-vous ? Pourquoi utilisez-vous ce type d'erreur type ?⁶
2. Prévoyez-vous de faire un test unilatéral ou bilatéral ?
3. À quel niveau α rejetterez-vous l'hypothèse nulle d'une p -valeur ?
4. Avez-vous plusieurs hypothèses ? Si oui, quelle procédure utiliserez-vous pour ajuster les comparaisons multiples ?

Vous pouvez choisir d'utiliser plusieurs procédures pour les critères d'inférence. Par exemple, vous pouvez utiliser des ajustements de taux d'erreurs par famille (family-wise error rate, FWER) et de taux de fausse découverte (false discovery rate, FDR) pour plusieurs comparaisons et comparer les deux. Ou vous pouvez utiliser les p -valeurs de l'inférence de randomisation pour vérifier les p -valeurs d'une distribution nulle supposée normale. Si vous choisissez d'utiliser plusieurs procédures, vous devez spécifier comment vous interpréterez les résultats si différentes procédures aboutissent à des conclusions différentes.

⁵Vous pouvez également être intéressé par l'estimation d'autres types d'effets. Voir ce guide sur les types d'effets de traitement pour plus d'informations sur les types d'effet.

⁶Par exemple, vous pouvez calculer les erreurs types et les p -valeurs en utilisant l'inférence de randomisation basée sur la permutation. Ou vous pouvez approcher précisément les erreurs types et les p -valeurs en utilisant des méthodes analytiques (Samii and Aronow 2012)

Exemple : Nous utilisons les erreurs types HC2 avec notre expérience randomisée par bloc car cela équivaut à l'estimateur de variance de Neyman basé sur la randomisation (Samii and Aronow 2012). Nous nous attendons à ce que le groupe de traitement paie plus de taxes municipales que le groupe de contrôle, et nous utilisons donc un test unilatéral où $H_1 > H_0$. Nous définissons $\alpha = 0,05$ et rejetons l'hypothèse nulle lorsque la p -valeur est inférieure à 0,05. Comme nous n'avons qu'une seule hypothèse confirmatoire, nous n'ajustons pas pour des comparaisons multiples. Pour vérifier les erreurs types HC2, nous calculons également les p -valeurs directement en utilisant l'inférence de randomisation, avec la différence des moyennes comme statistique de test.

7 Qu'y a-t-il dans un PAP? Procédures pour les problèmes de données

Les expériences et la collecte de données ne se déroulent souvent pas comme on s'y attend ou ce qu'on espère. Le PAP vous donne l'occasion de réfléchir à ce que pourraient être ces problèmes et de spécifier comment vous envisagez de les résoudre.

Deux problèmes de données courants sont (1) les données extrêmes et (2) les données manquantes.

Les données extrêmes peuvent représenter une vraie valeur aberrante - une unité avec des résultats beaucoup plus grands ou plus petits que d'autres unités - ou peuvent se produire en raison d'erreurs de collecte de données. Les tablettes de sondage, les outils d'enregistrement, les web scrapers et d'autres outils de collecte de données peuvent enregistrer des données extrêmes en raison de problèmes techniques. Il peut être difficile de savoir si les données extrêmes représentent de vraies données collectées ou s'ils sont dus à des erreurs de saisie de données, mais il est important de spécifier dans le PAP quand vous vous attendez à voir des erreurs de collecte de données et les procédures pour traiter les données extrêmes.

Les données manquantes peuvent se présenter sous deux formes : covariables et/ou résultats manquants. Il est également important de spécifier quand vous vous attendez à voir des covariables et/ou résultats manquants et les procédures pour les traiter dans votre PAP.

Les données extrêmes et/ou manquantes qui sont aléatoires seront moins problématiques que les données extrêmes et/ou manquantes qui semblent avoir un motif. Les procédures courantes pour traiter les données extrêmes ou manquantes sont 1) l'analyse des bornes ; 2) l'imputation; 3) l'abandon des observations. Nous vous recommandons de considérer les questions suivantes lors de la détermination de la procédure que vous souhaitez utiliser :

- Quels problèmes peuvent être à l'origine de ces données extrêmes ou manquantes ? Que pouvez-vous faire à l'avance pour atténuer ces problèmes de données ?
- Comment jugeriez-vous si les données extrêmes ou manquantes sont vraisemblablement aléatoires ? (i.e. les données extrêmes ou manquantes sont-elles corrélées au traitement, à des covariables/sous-groupes spécifiques ou aux résultats ?)
- Quelles procédures utiliserez-vous pour traiter les données extrêmes ou manquantes et comment justifiez-vous l'utilisation de votre procédure ?
 - Si les données extrêmes/manquantes ne sont pas aléatoires, nous vous recommandons d'inclure une sorte d'analyse des bornes pour déterminer les bornes de votre estimation avec des données extrêmes/manquantes. Par exemple, les bornes des valeurs extrêmes peuvent vous aider à déterminer la plage de l'effet moyen du traitement en attribuant pour tous les résultats manquants dans le groupe de traitement le plus grand résultat possible et en attribuant pour tous les résultats manquants dans le groupe de contrôle le plus petit résultat possible (Gerber and Green 2012).
 - Si les données extrêmes/manquantes sont aléatoires, nous recommandons d'imputer les données extrêmes ou manquantes. Vous pouvez également supprimer ces données si les données extrêmes/manquantes sont aléatoires, et vous devez spécifier les conditions dans lesquelles vous

choisissez de supprimer les données. Par exemple, envisagez de définir un seuil pour une covariable manquante de sorte que si le pourcentage de données manquantes de la covariable dépasse le seuil, vous supprimez la covariable.

Nous vous recommandons d'inclure le code utilisé dans vos procédures pour résoudre les problèmes de données.

Exemple : Résultats manquants pour les questions d'une enquête Certains répondants ne répondront pas à une ou plusieurs questions qui mesurent un résultat. Si nous remarquons que les taux de non-réponse aux questions semblent élevés ($\geq 10\%$) lors de la session pilote, nous demanderons des explications aux répondants et aux enquêteurs afin de pouvoir modifier les questions.

Lorsque les personnes interrogées ne répondent pas à une ou plusieurs questions mesurant un résultat pendant que nous menons notre enquête, notre procédure est la suivante :

Nous évaluerons la relation entre les résultats manquants et l'assignation du traitement à l'aide d'un test d'hypothèse et rapporterons ces résultats.

Si $p < 0,05$ pour l'évaluation de la relation entre le traitement et les résultats manquants, nous ajouterons une analyse des bornes pour ces valeurs extrêmes dans laquelle nous fixons tous les résultats manquants pour le traitement au maximum du bloc et tous les résultats manquants pour contrôle au minimum du bloc.

Si $p \geq 0.5$ pour l'évaluation de la relation entre le traitement et les résultats manquants, nous imputerons les résultats manquants en utilisant la moyenne de la sous-catégorie d'assignation par bloc.

Si nous avons encore des questions sur les résultats avec un taux de non-réponse particulièrement élevé de plus de 10 %, nous allons :

Décrire la relation entre la non-réponse à cette question et d'autres données sur les personnes via des explorations bivariées de graphiques et/ou de tableaux

une sélection de variables à l'aide d'un modèle de réseau élastique adaptatif avec des paramètres de réglage choisis par une validation croisée répétée 10 fois au sein du sous-groupe du design.

Envisager d'abandonner ces résultats d'enquête et répertorier la raison de notre décision.

8 Qu'y a-t-il dans un PAP? Analyse de puissance statistique

Une analyse de puissance statistique est généralement incluse dans les PAP. La puissance statistique est la probabilité que votre étude détecte un effet, s'il y a un effet à détecter.

Il existe des outils pour vous aider à calculer la puissance, mais vous pouvez également produire votre propre analyse de puissance programmatiquement. L'un des avantages de faire l'analyse de la puissance programmatiquement est qu'elle sert également de code d'analyse finale, ou au minimum de base pour le code d'analyse finale.

À bien des égards, l'analyse de la puissance de calcul implémente toutes les spécifications que vous avez faites dans votre code de PAP. En fait, vous pouvez combiner le code de l'analyse de puissance avec celui écrit pour d'autres sections du PAP et créer un PAP de calcul. Un tel outil, qui peut être implémenté avec un logiciel comme DeclareDesign, peut alors vous aider à diagnostiquer des problèmes potentiels, vous en apprendre davantage sur votre design et renforcer votre PAP.

9 Pourquoi faire un PAP ?

Vous avez peut-être entendu des arguments pour et contre les PAP. Cette discussion propose quelques réflexions sur les avantages et les coûts des PAP.

1. Les PAP aident à contrer la crise de la reproductibilité.

Un manque de transparence de la recherche a conduit à plusieurs problèmes, dont l'un est la crise de la reproductibilité.

La crise de la reproductibilité est que de nombreuses études universitaires sont difficiles ou impossibles à reproduire. Les PAP réduisent le nombre d'études qui ne sont pas reproductibles en raison de mauvaises pratiques en matière de données, telles que le data mining jusqu'à trouver des données montrant une relation causale fallacieuse ou le trucage involontaire des p -valeurs (p-hacking). Les PAP augmentent également le nombre d'études qui élaborent les procédures nécessaires à la reproductibilité, ce qui permet de tenter de reproduire ces études.

2. Les PAP aident à contrer les biais de publication.

Un manque de transparence de la recherche contribue également au biais de publication.

Un biais de publication se produit lorsque les revues publient des études basées sur les résultats de l'étude plutôt que sur la qualité de la recherche. Ce biais peut conduire à des croyances erronées sur un lien entre les variables X et Y car les revues ne publient que des études qui montrent que X affecte Y et ne publient pas d'études où X n'affecte pas Y , même si ces études sont plus nombreuses.

Les registres PAP fonctionnent comme des référentiels de tentatives d'études, à la fois publiées et non publiées. Ces registres permettent aux universitaires et aux praticiens d'identifier si les effets publiés sur un sujet représentent avec précision les effets trouvés dans les études non publiées, ou si les effets publiés diffèrent de ceux des études non publiées.

Le pré-enregistrement des études confirmatoires ou exploratoires permet en outre aux chercheurs de savoir si les recherches futures devraient s'appuyer sur l'étude ou si les recherches futures devraient corroborer l'étude. *Il n'y a rien de mal avec la recherche exploratoire*, et de nombreuses relations importantes mais inconnues peuvent être découvertes grâce à des analyses exploratoires. Il est important de reconnaître quand les résultats sont exploratoires et doivent être confirmés dans d'autres études. Les détails dans les registres PAP permettent aux universitaires de le faire.

3. Les PAP encouragent la recherche de qualité.

La création d'un PAP oblige le chercheur à élaborer les nombreuses décisions de design qui devront être prises lors de la réalisation d'une étude. Dans le cas des études d'observation, la modification ultérieure de ces décisions de design n'est pas un problème. Mais pour les études expérimentales ou d'autres études utilisant la collecte de données originales, le chercheur a une seule opportunité pour collecter les données nécessaires à l'analyse. Les PAP aident à garantir que le chercheur ait réfléchi à toutes les décisions et collecte les bonnes données.

4. Les PAP encouragent la recherche impactante.

Les PAP augmentent la transparence de la recherche, et la recherche transparente devrait être plus facilement fiable et utilisée que la recherche non transparente, car les décisions de l'étude et les raisons de ces décisions sont prises avant que les résultats de l'étude ne soient connus. La transparence garantit aux universitaires, aux politiques et aux autres communautés que les résultats de la recherche peuvent être utilisés comme base pour davantage de recherches, pour des programmes politiques et pour d'autres applications dans le monde réel.

5. Les PAP peuvent raccourcir le processus de publication.

Les PAP nécessitent plus d'investissement en temps avant la recherche, mais raccourcissent considérablement le temps d'analyse après la recherche car les décisions analytiques et le code sont écrits à l'avance. Les PAP *devraient* également raccourcir le processus d'examen qui mène à la publication. Les revues scientifiques nécessitent souvent de nombreux contrôles de robustesse avant d'accepter les résultats d'une étude. Les examinateurs peuvent demander moins d'analyses alternatives à partir de travaux prédéfinis, car il est clair pour les examinateurs que les décisions d'analyse prédéfinies n'ont pas été influencées par les résultats de l'étude. Ceci est particulièrement utile lorsque le chercheur souhaite utiliser un test statistique non

conventionnel mais plus puissant qui peut sembler suspect sans spécification préalable, comme des tests unilatéraux ou des statistiques de test autres que la différence des moyennes.

Pourquoi ne pas faire un PAP ?

1. La recherche est imprévisible et les PAP rendent la recherche inflexible.

Certaines personnes soutiennent qu'un PAP verrouille l'étude dans une stratégie de design, d'intervention et d'estimation particulière bien que les détails de la stratégie de design, d'intervention et d'estimation puissent changer pendant la réalisation d'une étude. Dans les études expérimentales, des difficultés imprévues modifient souvent les aspects de la randomisation ou de l'intervention, ou une nouvelle mesure de résultat peut échouer aux tests de validité. Et dans les études d'observation, une réflexion future sur la façon dont une théorie s'applique à vos données peut révéler le besoin de nouvelles variables de contrôle, de médiation et/ou de modération.

Les chercheurs doivent se rappeler que *tout plan de pré-analyse peut être révisé* ! Votre premier PAP ne vous enferme pas dans un design de recherche, une variable de résultat ou une spécification de modèle spécifique. *Le PAP rend le processus de recherche transparent et non rigide*. Des révisions peuvent être apportées soit en soumettant un nouveau PAP, soit via un amendement décrivant les changements par rapport au PAP précédent. Les analyses exploratoires sont acceptables ! Le but d'un PAP n'est pas d'empêcher ces analyses imprévues, mais de formaliser et d'expliquer le processus qui a conduit aux analyses.

Pour plus de détails sur les avantages et les inconvénients des PAP, consultez le chapitre 19 de Rosenbaum (2010) et Olken (2015).

10 Quand et où enregistrer un PAP ?

Votre PAP est maintenant terminé et vous êtes prêt à l'enregistrer ! Pour les études expérimentales, vous devez enregistrer le PAP au plus tard avant la collecte des données finales. Pour les études d'observation, vous devez enregistrer le PAP au plus tard avant toute analyse, y compris l'examen des statistiques descriptives. Vous pouvez réviser les PAP par le biais d'un avenant à tout moment.

Il existe plusieurs sites tiers sur lesquels vous pouvez enregistrer votre PAP. Nous listons ci-dessous les sites communs pour les PAP en sciences sociales. Vous pouvez répertorier votre PAP sur plusieurs sites, et certaines revues exigent d'enregistrer un PAP sur un site spécifique pour potentiellement publier un article.⁷

- Registre EGAP
- Registre AEA (pour les essais randomisés contrôlés seulement)
- Registre OSF

Références

- Caughey, Devin, Allan Dafoe, and Jason Seawright. 2017. "Nonparametric Combination: A Framework for Testing Elaborate Theories." *The Journal of Politics* 79 (2): 688–701.
- Gerber, Alan S, and Donald P Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. WW Norton.
- Nosek, B. A., G. Alter, G. C. Banks, D. Borsboom, S. D. Bowman, S. J. Breckler, S. Buck, et al. 2015. "Promoting an Open Research Culture." *Science* 348 (6242): 1422–25.
- Olken, Benjamin A. 2015. "Promises and Perils of Pre Analysis Plans." *Journal of Economic Perspectives* 29 (3): 61–80.

⁷Notez que certaines organisations n'utilisent pas de sites tiers. Par exemple, le U.S. General Services Administration Office of Evaluation Sciences Process utilise Github, qui permet de vérifier que le PAP a été créé avant la réalisation de l'analyse.

- Rosenbaum, Paul R. 2010. *Design of Observational Studies*. Springer Series in Statistics. Springer.
- Samii, Cyrus, and Peter M Aronow. 2012. “On Equivalencies Between Design Based and Regression Based Variance Estimators for Randomized Experiments.” *Statistics and Probability Letters* 82 (2): 365–70.