

## Randomization 2: Designing for common constraints & threats | *Randomisation 2 : conception face aux contraintes & menaces*

Alyssa René HEINZE

28-06-2023

Randomization and design: What we've learned, what's next | *Randomisation et conception : Ce que nous avons appris, prochaines étapes*

Common constraints, more complex designs | *Contraintes courantes, conceptions plus complexes*

Common threats, anticipatory designs | *Menaces courantes, conceptions anticipatrices*

Final takeaway: Keep things simple | *Dernier point : Gardez les choses simples*

## Key points for this lecture | *Points clés*

- ▶ Randomization and design: What we've learned, what's next
- ▶ Common constraints, more complex designs
- ▶ Common threats, anticipatory designs
- ▶ Final takeaway: Keep things simple
- ▶ Randomisation et conception : ce que nous avons appris, prochaines étapes
- ▶ Contraintes courantes, conceptions plus complexes
- ▶ Menaces courantes, conceptions anticipatrices
- ▶ Dernier point : gardons les choses simples

Randomization and design: What we've learned, what's next |  
*Randomisation et conception : Ce que nous avons appris,  
prochaines étapes*

# Randomization and design: What we've learned | Randomisation et conception : Ce que nous avons appris

- ▶ What we've covered:
  - ▶ 4 types of randomization: Simple, complete, block, and cluster
  - ▶ And we can combine them: block-cluster, factorial designs
- ▶ Most of the time these designs suffice.
- ▶ Ce que nous avons couvert :
  - ▶ 4 types de randomisation : simple, complète, par bloc et par grappe
  - ▶ Et on peut les combiner : par bloc et par grappe, conceptions factorielles
- ▶ La plupart du temps, ces conceptions suffisent.

## Randomization and design: What's next | Randomisation et conception : prochaines étapes

- ▶ *BUT* we often face:
  - ▶ Practical constraints in our context
  - ▶ Common threats to inference
- ▶ *So* we need to take these into account in the *design* of our experiment.
  - ▶ Designing to take into account context-specific constraints:
    - ▶ Waitlist designs, encouragement designs
  - ▶ Designing to anticipate problems:
    - ▶ Attrition, spillover
- ▶ *MAIS* nous sommes souvent confrontés à :
  - ▶ Contraintes pratiques dans notre contexte
  - ▶ Menaces courantes à l'inférence
- ▶ *Donc* nous devons en tenir compte dans la *conception* de notre expérience.
  - ▶ Concevoir en tenant compte des contraintes spécifiques au contexte :
    - ▶ Conceptions avec liste d'attente, conceptions incitatives
  - ▶ Concevoir pour anticiper les problèmes :
    - ▶ Attrition, contamination

Common constraints, more complex designs | *Contraintes courantes, conceptions plus complexes*

## Common constraints | *Contraintes courantes*

1. Constrained in how many units can be treated at one time
2. You can't force people to take your treatment

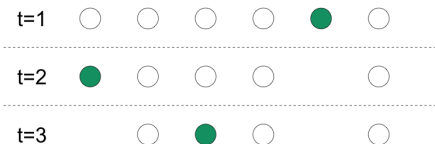
1. Limite du nombre d'unités pouvant être traitées en même temps
2. Vous ne pouvez pas forcer les gens à suivre votre traitement



## More complex designs (1): Waitlist design | *Conceptions plus complexes (1): liste d'attente*

- ▶ Constraint: Only certain number of units can be treated at a time
- ▶ Solution: *Waitlist* design
- ▶ Contrainte: Seul un certain nombre d'unités peuvent être traitées en même temps
- ▶ Solution: conception avec liste d'attente

## 1. Delayed access (Phase-in or wait list) | Accès différé (Accès graduel ou liste d'attente)



- ▶ When an intervention can be or must be rolled out in stages, you can randomize the order (*timing*) in which units are treated.
- ▶ Your control group are the as-yet untreated units.
- ▶ Lorsqu'une intervention peut ou doit être déployée par étapes, vous pouvez randomiser l'ordre (*timing*) de traitement des unités.
- ▶ Votre groupe de contrôle sont les unités pas encore traitées.

## 1. Delayed access (Phase-in or wait list) | Accès différé (Accès graduel ou liste d'attente)

- ▶ Be careful: the probability of assignment to treatment will vary over time because units that are assigned to treatment in earlier stages are not eligible to be assigned to treatment in later stages.
- ▶ Attention : la probabilité d'assignation au traitement variera dans le temps car les unités assignées au traitement à des stades antérieurs ne sont plus éligibles pour être assignées à un traitement aux stades ultérieurs.

## More complex designs (2): Encouragement design | *Conceptions complexes (2): conception incitative*

- ▶ Constraint: You can't force people to take (receive) your treatment
  - ▶ Solution: **Encouragement** design
- ▶ Contrainte : Vous ne pouvez pas forcer les gens à prendre (recevoir) votre traitement
  - ▶ Solution : conception **incitative**

## 2. Encouragement (planning for non-compliance) | *Incitations*

- ▶ Randomize **encouragement** to take the treatment, such as a request to drink coffee or offering a subsidy to participate in a program.
- ▶ Useful when you cannot force a subject to take the assigned treatment.
- ▶ We can learn the average effect of the encouragement to take the treatment.
- ▶ Randomisez l'**incitation** (l'encouragement) à suivre le traitement, en demandant par exemple aux individus de boire un café ou en offrant une subvention pour participer à un programme.
- ▶ Utile lorsque vous ne pouvez pas forcer un sujet à participer.
- ▶ Nous pouvons connaître l'effet moyen de l'encouragement à suivre le traitement.

## 2. Encouragement (planning for non-compliance) | *Incitations*

- ▶ We can also learn the average effect of the taking the treatment *for those subjects that would take the treatment when assigned to treatment and not take the treatment when assigned to control (compliers)*.
- ▶ But we need an additional assumption (exclusion restriction) and estimation is more complicated (instrumental variables).
- ▶ Nous pouvons également connaître l'effet moyen de la prise du traitement *pour les sujets qui prendraient le traitement lorsqu'ils sont assignés au traitement et qui ne prendraient pas le traitement lorsqu'ils sont assignés au contrôle (les conformistes)*.
- ▶ Mais nous avons besoin d'une hypothèse supplémentaire (restriction d'exclusion) et l'estimation est plus compliquée avec l'assignation comme instrument.

## Résumé :

- ▶ Nous pouvons être confrontés à des contraintes qui sont souvent d'ordre pratique et spécifiques au contexte de notre étude
- ▶ Ici, nous avons parlé de deux types de contraintes :
  1. Limitation du nombre d'unités auxquelles nous pouvons administrer le traitement en même temps
  2. Avoir un traitement qui ne peut pas être imposé aux gens
- ▶ Nous avons également abordé les solutions à ces problèmes, que nous pouvons intégrer dans nos conceptions de recherche
  1. Randomiser le *timing* où les unités recevront le traitement (liste d'attente)
  2. Randomiser *l'encouragement* à suivre le traitement (conception incitative)

## Quiz

Graeme wants to study the impact of drinking coffee (treatment) on LD participants' level of enthusiasm in the first lecture of the day. He however cannot force LD participants to drink coffee. He receives 3 pieces of advice from different LD instructors on how to incorporate this into his design. Which piece of advice is best given the specific constraint he's facing?

- A. Nahomi says to randomize the timing at which participants are assigned to drink coffee or not.
- B. Lauren says to randomize the *encouragement* to drink coffee.
- C. Natalia recommends block randomizing according to nationality, and whether or not an LD participant likes coffee.



## Quiz

Graeme souhaite étudier l'impact de la consommation de café (traitement) sur le niveau d'enthousiasme des participants lors du premier cours de la journée. Il ne peut cependant pas forcer les participants à boire du café. Il reçoit 3 conseils de différents instructeurs sur la façon d'intégrer cela dans sa conception de recherche. Quel conseil est le meilleur compte tenu de la contrainte spécifique à laquelle il est confronté ?

- A. Nahomi dit de randomiser le moment auquel les participants sont assignés à boire du café ou non.
- B. Lauren dit de randomiser *l'encouragement* à boire du café.
- C. Natalia recommande la randomisation par bloc en fonction de la nationalité et du fait qu'un participant aime ou non le café.

Common threats, anticipatory designs | *Menaces courantes,  
conceptions anticipatrices*

## Common threats to inference | *Menaces courantes pour l'inférence*

1. Attrition: People do not respond to follow up surveys or are hard to find
  2. Spillover: The effects of a treatment may spill over from people treated to their neighbors, friends, or family members, making it hard to nail down the effect on any one person
1. Attrition : les personnes ne répondent pas aux enquêtes de suivi ou sont difficiles à trouver
  2. Contamination : les effets d'un traitement peuvent se propager des personnes traitées à leurs voisins, amis ou membres de la famille, ce qui rend difficile la détermination de l'effet sur une personne en particulier.

## Anticipating threats via design (1): Attrition | *Anticiper les menaces par la conception (1) : Attrition*

- ▶ Challenge: Attrition
- ▶ Attention to this *important*
- ▶ Solution(s):
  - ▶ Pilot your approach to maximize chances of response
  - ▶ Plan (budget) to reach all subjects at endline
  - ▶ Plan (budget) for double-sampling to follow-up
  - ▶ Blind people to treatment status
  - ▶ Deliver treatment to control group post-research
- ▶ Défi : Attrition
- ▶ Attention, c'est *important*
- ▶ Solution(s) :
  - ▶ Pilotez votre approche pour maximiser les chances de réponse
  - ▶ Planifiez (le budget) pour atteindre tous les sujets à la fin
  - ▶ Planifiez (le budget) pour le double-échantillonnage de suivi
  - ▶ Personnes aveugles à leur statut de traitement
  - ▶ Administrer le traitement au groupe témoin après la recherche

## 1. Attrition: Why it matters | *Attrition : pourquoi est-ce important?*

- ▶ Missing outcome data is problematic when:
  - ▶ Missingness is systematically associated with certain *types* of people, or with units' treatment assignment status
- ▶ Les données manquantes sur les résultats sont problématiques lorsque :
  - ▶ L'absence est systématiquement associée à certains *types* de personnes, ou au statut de traitement des unités

## 1. Attrition: Why it matters | *Attrition : pourquoi est-ce important?*

- ▶ The problem is this:

- ▶ Without outcome data, can't analyze impact of treatment on outcomes
- ▶ Can't just ignore the missing data
- ▶ Treatment and control groups are no longer comparable
- ▶ This breaks the "analyze as you randomize" principle

- ▶ Le problème est le suivant :

- ▶ Sans données sur les résultats, impossible d'analyser l'impact du traitement sur les résultats
- ▶ On ne peut simplement pas ignorer les données manquantes
- ▶ Les groupes de traitement et de contrôle ne sont plus comparables
- ▶ Cela enfreint le principe "analyser selon le niveau de randomisation"

## 1. Attrition: When it can happen | *Attrition : Quand ?*

- ▶ Kids who anticipate failing test are absent the day of the test
- ▶ Researcher effort: Control units more likely to be missing because researcher engaged less with them
- ▶ Les enfants qui anticipent un échec au test sont absents le jour du test.
- ▶ Effort du chercheur: les unités de contrôle sont plus susceptibles d'être manquantes parce que le chercheur s'est moins engagé avec elles

## 1. Attrition: When it can happen | *Attrition : Quand ?*

- ▶ Data problems: elections results not reported for a certain region, people move away and they're more difficult to track, people change phone numbers
- ▶ Responding to surveys is annoying to some people, or they're busy when you approached them
- ▶ Problèmes de données: résultats des élections non communiqués pour une certaine région, les gens déménagent et sont plus difficiles à suivre, les gens changent de numéro de téléphone
- ▶ Répondre à des sondages est ennuyeux pour certaines personnes, ou elles sont occupées lorsque vous les approchez



# 1. Attrition: Anticipating it in design | *Attrition : l'anticiper lors de la conception*

## 1. Pilot your approach: in pilot, figure out what works best to get responses

- ▶ Different types of incentives?
- ▶ Call versus in-person?
- ▶ Best days or timings?

## 1. Pilotez votre étude : dans l'étude pilote, déterminez ce qui fonctionne le mieux pour obtenir des réponses

- ▶ Différents types d'incitations (récompenses) ?
- ▶ Collecter des données par téléphone ou en personne ?
- ▶ Y a-t-il des jours ou des horaires où les gens sont plus disponibles ?

# 1. Attrition: Anticipating it in design | *Attrition: l'anticiper lors de la conception*

## 2. Plan to reach all subjects at endline

- ▶ Budget for strong efforts (time, resources) to get responses

## 3. Plan for double-sampling

- ▶ Budget for *intensive* follow-up with random sample of attriters

## 2. Prévoyez d'atteindre tous les sujets à la fin

- ▶ Préparer votre budget pour des efforts importants (temps, ressources) pour obtenir des réponses

## 3. Prévoyez un double échantillonnage

- ▶ Prendre en compte dans votre budget la possibilité d'un suivi *intensif* avec un échantillon aléatoire de ceux qui n'ont pas répondu à l'enquête finale

# 1. Attrition: Anticipating it in design | *Attrition : l'anticiper lors de la conception*

## 4. Blind people to their treatment status

- ▶ Idea: people won't avoid follow-up if they know they're in control

## 5. Promise to deliver treatment to control post-research

- ▶ Idea: people will remain engaged in the study even if in control

## 4. Les personnes aveugles quant à leur statut de traitement

- ▶ Idée : les gens n'éviteront pas l'enquête de suivi s'ils ne savent pas qu'ils sont dans le group de contrôle

## 5. S'engager à fournir le traitement aux personnes qui faisaient partie du groupe de contrôle une fois l'étude terminée

- ▶ Idée : les gens resteront engagés dans l'étude même s'ils sont dans le group de contrôle

## 1. Attrition: Anticipating it in design | *Attrition : l'anticiper lors de la conception*

- ▶ NOTE: Keeping good records is crucial!
  - ▶ Treatment assignment status
  - ▶ Outcome variable as NA if missing
  - ▶ (By the way, here we're not talking about covariates)
- ▶ REMARQUE : Tenir de bons registres est crucial !
  - ▶ Statut d'assignation au traitement
  - ▶ Variable de résultat comme NA si manquante
  - ▶ (Au fait, ici on ne parle pas de covariables de base)

## Anticipating threats via design (2): Spillover | *Anticiper les menaces lors de la conception (2) : contamination*

- ▶ Challenge: Spillover
- ▶ Attention to this *important*
- ▶ Why?
  - ▶ Violates core assumption for causal inference
- ▶ Solution:
  - ▶ Keep units far apart

- ▶ Défi : Contamination
- ▶ Attention, c'est *important*
- ▶ Pourquoi ?
  - ▶ Il s'agit d'une violation d'une hypothèse fondamentale de l'inférence causale
- ▶ Solution :
  - ▶ Gardez les unités éloignées les unes des autres

## 2. Spillover: Anticipating it in design | *Contamination : l'anticiper lors de la conception*

- ▶ *Minimize* spillover by sampling units far enough apart geographically
- ▶ Idea: avoid that a treated unit *contaminate* another unit
- ▶ NOTE: possible to *study* spillover effects as the focus, but difficult (lots of caveats in order)
  - ▶ Often better to try to avoid spillover altogether
- ▶ *Minimiser* la contamination en échantillonnant des unités suffisamment éloignées géographiquement
- ▶ Idée : éviter qu'une unité traitée *contamine* une autre unité
- ▶ REMARQUE : il est possible d'avoir pour objectif d'*étudier* les effets de contamination, mais difficile
  - ▶ Il est souvent préférable d'essayer d'éviter les effets de la contamination

## Résumé :

- ▶ Certains problèmes sont assez courants dans les expériences aléatoires. Ces problèmes menacent la validité interne d'une étude.
- ▶ Nous avons parlé de deux menaces courantes pour la validité interne d'une expérience aléatoire :
  1. Lorsque nous manquons de données sur les résultats de certaines unités
  2. Lorsque le statut de traitement d'une unité peut avoir un impact sur une autre unité, entraînant des effets de "contamination".
- ▶ Nous avons discuté des différentes méthodes à prévoir au stade de la conception d'une étude pour anticiper ces problèmes très courants
  1. Faites de votre mieux pour éviter les données manquantes. Pilotez vos approches de collecte de données pour maximiser les chances de réponse. Prévoyez un budget pour des efforts intenses afin d'atteindre les sujets lors de l'enquête finale, etc.
  2. Éviter toute possibilité de contamination entre les unités en les maintenant éloignées les unes des autres.

## Quiz

Yabo wants to study the impact of information about voting rights on voter turnout. She wants to assign treatment at the level of individual citizens. However, the context where she works has small villages where people frequently interact. She is worried about this threatening the internal validity of her study. She receives 3 pieces of advice from different LD instructors on how to anticipate such threats in her design. Which piece of advice is best?

- A. Christelle advises that she run her baseline survey in person rather than on the phone, in order to make sure more people respond.
- B. Nahomi suggests that she incentivize responses by paying citizens to take part in the study.
- C. Horace recommends ensuring that the units in the study are far enough apart to avoid contamination. He recommends randomizing at the level of the village, rather than at the individual level within villages.



## Quiz

Yabo souhaite étudier l'impact de l'information sur les droits de vote sur la participation électorale. Elle veut assigner un traitement au niveau du citoyen. Cependant, le contexte dans lequel elle travaille comporte de petits villages où les gens interagissent fréquemment. Elle craint que cela ne menace la validité interne de son étude. Elle reçoit 3 conseils de différents instructeurs sur la façon d'anticiper de telles menaces dans sa conception de recherche. Quel est le meilleur conseil ?

A. Christelle conseille d'effectuer son enquête en personne plutôt qu'au téléphone, afin de s'assurer que davantage de personnes répondent.

B. Nahomi suggère qu'elle encourage les réponses en payant les citoyens pour qu'ils participent à l'étude.

C. Horace recommande de s'assurer que les unités de l'étude sont suffisamment éloignées pour éviter toute contamination. Il recommande la randomisation au niveau du village, plutôt qu'au niveau individuel au sein des villages.

Final takeaway: Keep things simple | *Dernier point : Gardez les choses simples*

## Complex $\neq$ better! | *Complexe $\neq$ mieux !*

- ▶ Complexity is possible, but simplicity is always preferred.
  - ▶ People mistakenly think
    - ▶ Complex = sophisticated  $\rightarrow$  better
  - ▶ In fact, straightforward designs  $\rightarrow$  better
- ▶ La complexité est possible, mais la simplicité est toujours souhaitable.
  - ▶ Les gens pensent à tort
    - ▶ Complexe = sophistiqué  $\rightarrow$  mieux
  - ▶ En réalité, conception simple  $\rightarrow$  mieux

## Keep things simple | *Gardons les choses simples*

- ▶ You now have more complex designs in your toolkit, but:
  - ▶ If you don't need them, don't use them
- ▶ Why? With simplicity:
  - ▶ Less can go wrong
  - ▶ More likely to notice when something does
  - ▶ Easier to understand design properties
  - ▶ Interpretability and communication of findings
- ▶ Vous avez maintenant des conceptions plus complexes dans votre boîte à outils, mais :
  - ▶ Si vous n'en avez pas besoin, ne les utilisez pas
- ▶ Pourquoi ? Avec une conception de recherche *simple* :
  - ▶ Moins de choses peuvent aller de travers
  - ▶ Plus facile de remarquer que quelque chose ne se passe pas comme prévu
  - ▶ Plus facile à comprendre les propriétés de la conception
  - ▶ Interprétabilité et communication des résultats

- ▶ EGAP Methods Guide on Randomization  
(<https://egap.org/resource/10-things-to-know-about-randomization/>)
- ▶ EGAP Methods Guide on Missing Data  
(<https://egap.org/resource/10-things-to-know-about-missing-data/>)
- ▶ Guide des méthodes EGAP sur la randomisation  
(<https://egap.org/fr/resource/10-choses-a-savoir-sur-la-randomisation/>)
- ▶ Guide des méthodes EGAP sur les données manquantes  
(<https://egap.org/resource/10-things-to-know-about-missing-data/>)