

# ÉQUITÉ, DIVERSITÉ ET INCLUSION DANS LES SYSTÈMES DE SANTÉ APPRENANTS

Anicet Nzeukoc, MSc<sup>1,2</sup>, Rostand Tchouakam, MSc<sup>1,3</sup>

1. Université de Montréal (UdeM)
2. Institute for data valorisation (IVADO)
3. SCIKOOP Inc

rostand.tchouakam.mbouendeu@umontreal.ca  
anicet.nzeukoc@umontreal.ca  
info@scikoop.com



## Résumé

La représentativité est le degré auquel un échantillon de données est représentatif d'une population plus large. Dans le domaine de la santé, la représentativité est essentielle pour s'assurer que les résultats de la recherche sont généralisables à la population cible. Des données collectées non-représentatives peuvent avoir des conséquences néfastes pour la qualité de l'apprentissage des systèmes de santé et les prises de décision en matière de soins. Afin de mesurer la représentativité de leurs données par rapport à une population cible, les personnes chercheuses ont à leur disposition des outils tels que des graphiques et des calculs de proportions pour les statistiques descriptives, des tests de comparaison de proportions pour la statistique inférentielle et des métriques d'équité. À l'aide de ces outils, les personnes chercheuses garantissent que leurs données sont pertinentes pour les populations étudiées et permettent de prendre des décisions éclairées pour améliorer les résultats en santé.

## Introduction

L'introduction de biais dans les données de la recherche en santé causent des erreurs systématiques qui favorisent certains résultats au dépend d'autres et mènent à des erreurs de traitements ou de politique de santé publique. Pour des résultats fiables, il est donc crucial de comprendre les conséquences des biais de représentation et de maîtriser des outils statistiques permettant d'en mesurer l'ampleur et de minimiser l'impact de ces biais sur les conclusions de la recherche.

## Objectifs

Dans le cadre de l'amélioration du système de santé apprenant (SSA), il est essentiel de veiller à la diversité dans les données collectées. Notre travail de transfert de connaissances vers l'entreprise SCIKOOP consistait à :

- Effectuer une revue de littérature pour clarifier les conséquences d'un manque de représentativité dans les données de recherche;
- Identifier des méthodes statistiques pour mesurer et minimiser le biais de représentativité dans les données.

## Méthodologie

Notre approche consiste en une revue de la littérature pour identifier les travaux qui ont exploré l'importance de la représentativité des données dans le domaine de la santé. Cette analyse vise à synthétiser les résultats des études précédentes et de clarifier les conséquences potentielles d'un manque de représentativité dans les données de recherche en santé. D'autre part la revue de la littérature d'un ensemble de méthodes statistiques permettant aux équipes de recherche de mesurer le degré de représentativité des données par rapport à la population cible.

He et al. (2015) " Simulation-based Evaluation of the Generalizability Index for Study Traits ". AMIA Annual Symposium Proceedings

Okunlola et al. (2022) "Pulse oximeter performance, racial inequity, and the work ahead." Respiratory care 67.2 (2022): 252-257.

Qi et al. (2021) "Quantifying representativeness in randomized clinical trials using machine learning fairness metrics ". JAMIA Open 4(3)

Ruzycki et al. (2022) "Equity, diversity and inclusion are foundational research skills." Nature Human Behaviour 6.7 : 910-912.

Sardar et al. (2014) "Underrepresentation of women, elderly patients, and racial minorities in the randomized trials used for cardiovascular guidelines". JAMA Intern Med. Nov;174(11):1868-70.

Sedrak et al. (2021) "Older adult participation in cancer clinical trials: A systematic review of barriers and interventions". CA Cancer J Clin. ;71(1):78-92.

Tipton, Elizabeth. (2014). " How Generalizable Is Your Experiment? An Index for Comparing Experimental Samples and Populations ". Journal of Educational and Behavioral Statistics 39(6):478-501.

Weng et al. (2014) "A distribution-based method for assessing the differences between clinical trial target populations and patient populations in electronic health records." Applied clinical informatics 5.02 : 463-479.

## Tableau 1: Conséquences d'un manque de représentativité dans les données

Étude de représentativité	Biais et conséquences
Cas de la COVID-19	Sous-représentation des femmes enceintes dans les essais cliniques de la vaccination contre la COVID19  Faible taux de vaccination des femmes enceinte dans le personnel soignant
Cas des maladies respiratoires	Sous-représentation des personnes à la peau foncée dans les essais cliniques  Erreur dans les taux d'oxygénation sanguine occultée chez ce groupe de patients
Cas des cancers	Sous-représentation des personnes seniors dans les essais cliniques des traitements pour les cancers  Absence de pratiques adaptées à ce groupe de patients
Cas des maladies cardio-vasculaires	Surreprésentation des hommes dans la recherche biomédicale sur les maladies cardio-vasculaires  Diagnostics et traitements inappropriés chez les femmes
Cas du diabète de type 2	Biais d'échantillonnage des essais cliniques sur le diabète de type 2  Résultats non-généralisables en ce qui concerne les niveaux d'hémoglobine HBA1c et valeurs moyennes de l'hémoglobine HbA1c inadaptée à la population cible de New-York

## Tableau 2: Outils de mesure de la représentativité des données

Outil de mesure	Description
Statistiques descriptives	Histogrammes, diagrammes circulaires (camemberts), tableaux, moyennes, proportions
Indice de représentativité ML	$I = \frac{P(y' = 1   g = 1)}{P(y' = 1   g = 0)}$
Indice de comparaison population-échantillon	$\beta = \sum_{g=1}^k \sqrt{p_g p'_g}$
Indice Gist de représentativité des études cliniques	$GIST = \sum_{i=1}^n \frac{\sum_{j=1}^T I(i_l \leq w_j \leq i_h)}{T} \cdot \frac{\sum_{k=1}^P I(i_l \leq y_k \leq i_h)}{P}$
Statistique du test de comparaison de deux proportions	$Z = \frac{p'_g - p_g}{\sqrt{\frac{p_g(1 - p_g)}{n}}}$

y'=1 : signifie qu'une personne est recrutée g=1: signifie qu'un individu doté de l'attributs protégés ait partie du sous-groupe

## Discussion

Le système de santé apprenant, s'appuie sur une recherche participative qui vise à impliquer les patient.e.s, les personnels soignants et les familles conjointement dans le processus de recherche pour répondre à leurs priorités respectives et améliorer les soins et services de santé. Afin de garantir la représentativité des participant.e.s et capturer une diversité des savoirs expérimentiels dans le processus d'apprentissage des systèmes de soin, il est primordial d'identifier tous les facteurs ou sources de biais de représentation et de les corriger.

Cette étude a permis d'identifier des facteurs qui contribuent à la création d'algorithmes pour évaluer la diversité des participant.e.s à la recherche et leur meilleure inclusion dans les projets de recherche axés sur les patient.e.s. Ces algorithmes intégrés dans les solutions numériques de l'entreprise SCIKOOP Inc visent à favoriser de meilleures pratiques d'équité, diversité et d'inclusion dans le partenariat entre patients et chercheurs, dans la recherche axée sur les priorités des patients, dans l'accès équitable aux systèmes de santé et dans l'amélioration de la santé de tous. De telles solutions contribueront à réduire les biais de représentation dans la recherche en santé et à promouvoir l'équité dans les systèmes de soins.