The feasibility of the Platform Trial design for the development of a mobile health application for improving treatment adherence among persons living with HIV.

Asma Aqib

Department of Family Medicine

McGill University, Montreal

April 2020.

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science in Family Medicine.

Abstract

Background: The I-Score study was conducted in Canada and France to develop and validate a Patient-Reported Outcome Measure (PROM), identifying barriers to improve Anti-Retroviral Treatment adherence (ART) in People Living with HIV (PLHIV). The conceptual framework of the PROM is based on qualitative studies conducted on PLHIV in high-income countries. Subsequently, stakeholders (PLHIV and HIV clinicians) were engaged for the development and implementation of the PROM. The aim was to deliver the validated PROM to patients and healthcare providers via a mobile health application ('mHealth app') while continuing to develop various aspects and design features of the mHealth app.

To enable an effective and unbiased approach towards evaluating different developer versions of the mHealth app, an adaptive trial design called Platform Trial (PT) has been proposed. As this is a novel trial design, not all challenges and practicalities relevant for proper trial conduct are understood. Therefore, before commencing the PT, it is important to explore all the ethical, pragmatic or technical challenges associated with the design that has been reported in the literature. The I-score study also aims to accommodate patient diversity and to contextualize, personalize and generalize the mHealth app to a diverse population of PLHIV in Montreal. Therefore, for ensuring proper planning and set up of the trial and for ensuring the external validity of the outcomes, it is important to characterize the study population.

The purpose of this thesis was to 1) inform the I-score study about the ethical, pragmatic, and technical challenges of PT reported in literature 2) to characterize a study population, the 'Cohorte de Montreal' (CM), composed of PLHIV from 4 HIV clinics in Montreal, 3) and to identify similarities and differences in the characteristics of PLHIV among these different clinics.

Methods: Phase 1: I conducted a descriptive literature review to identify ethical, pragmatic, and technical issues related to the conduct of a PT. The literature review was conducted using two databases (Medline via Pubmed & EMBASE via Ovid). Phase 2: Through a descriptive cross-sectional study of the Cohorte de Montreal (CM) data, I characterized the target population and explored commonalities and differences in the demographic and clinical characteristics of PLHIV in the four clinics, with a focus on the PLHIV from the McGill University Health Network.

Results: Phase 1: 459 articles were screened by title and abstract; 164 full-text articles were assessed for eligibility and 27 were eventually included in the synthesis. I identified several ethical issues related to informed consent, equipoise, and justice, as well as various pragmatic or technical issues including biases, logistical or statistical matters that may challenge the integrity, and validity of a PT. **Phase 2**: PLHIV followed at MUHC differed slightly from the other clinics on various demographic characteristics. Namely, a higher proportion of women, larger ethnic diversity and a higher proportion of heterosexual individuals from endemic countries were observed in the MUHC clinics compared to other clinics. However, the clinical characteristics of PLHIV at MUHC (mean CD4 counts and the proportion of PLHIV with undetectable viral loads) were comparable across all clinics.

Conclusion: Several factors may impede the proper roll-out of a PT in the context of developing a mobile health app in larger patient populations. However, with adequate preventive measures, pre-planning, and involvement of all stakeholders, these limiting factors may be controlled. PT design enables contextualization, personalization, and generalization of the results if rolled out in large and diverse study populations. Finally, PT has promising utility in the seamless evaluation of interventions that undergo continuous development, including the mHealth app.

Résumé

Contexte: Le but de l'étude I-Score est de développer et valider une mesure-patient permettant aux personnes vivant avec le VIH (PVVIH) d'identifier les obstacles au maintien de l'adhésion aux traitements antirétroviraux. Le cadre conceptuel de cette mesure-patient a été développé à partir d'une synthèse d'études qualitatives menées avec les PVVIH dans les pays à revenu élevé. Ensuite, des PVVIH et cliniciens spécialisés ont été engagés dans le développement et la réalisation de cette mesure-patient. L'objectif était de valider la mesure-patient, et de s'assurer de son utilité sur une application de santé mobile ("l'app."). En parallèle, les différents aspects et caractéristiques de conception de l'app. supportant la mesure-patient ont aussi été travaillés. Pour permettre une approche efficace et impartiale de l'évaluation des différentes versions de l'app., un essai de type platetorme (EP) a été proposé. L'EP est un type d'essai novateur qui présente plusieurs défis et dont les aspects pratiques ne sont pas encore pleinement compris, ce qui peut impacter son bon déroulement. Pour cette raison, il est important d'explorer les défis éthiques, pragmatiques et techniques des EP rapportés dans la littérature. L'étude I-Score vise également à tenir compte de la diversité des PVVIH à Montréal et à contextualiser, personnaliser et généraliser l'app. Donc, pour une planification et une mise en place adéquates de l'EP et pour assurer la validité externe des résultats, il est important de caractériser la population étudiée. Mes objectives était 1) d'informer l'étude I-Score des défis éthiques, pragmatiques et techniques des EP et autres essais adaptatifs rapportés dans la littérature, 2) de caractériser la population de l'étude, à partir de la Cohorte de Montréal (CM), et 3) d'identifier les points communs et les différences dans les caractéristiques des PVVIH inclus dans la CM.

Méthodologiquement : dans un premier temps, j'ai effectué une analyse documentaire descriptive à partir de deux bases de données dans le but de cerner les défis éthiques,

pragmatiques et techniques liées à l'exécution d'un EP. Dans un deuxième temps, par une étude transversale descriptive de la CM, j'ai caractérisé la population cible de l'étude I-Score et j'ai exploré les caractéristiques démographiques et cliniques des PVVIH de quatre cliniques inclues dans la CM, tout en gardant un focus particulier sur l'unité VIH du Centre Universitaire de Santé McGill (CUSM).

Résultats: Lors de la phase 1, 459 articles individuels ont été examinés et 27 ont été inclus dans la synthèse. J'ai identifié plusieurs questions éthiques liées au consentement éclairé, à l'incertitude et à la justice, ainsi que diverses questions pragmatiques ou techniques liés aux préjugés et aux questions logistiques ou statistiques qui peuvent remettre en question l'intégrité et la validité d'un EP. Lors de la deuxième phase, les PVVIH suivies au CUSM présentaient des caractéristiques démographiques légèrement différentes de celles des autres cliniques, avec des proportions plus élevées de femmes, une plus grande diversité ethnique et une proportion plus élevée d'individus hétérosexuels issus de pays endémiques. Cependant, les caractéristiques cliniques des PVVIH au CUSM étaient comparables à celles des patients des autres cliniques.

Conclusion: Plusieurs facteurs peuvent entraver le bon déroulement d'un EP dans le contexte du développement d'une app. visant de grandes populations de patients. Cependant, avec des mesures préventives adéquates, une planification préalable et la participation de tous les intervenants, ces facteurs limitatifs peuvent être contrôlés. La conception de l'EP permet la contextualisation, la personnalisation et la généralisation des résultats s'ils sont déployés dans des populations vastes et diversifiées. Enfin, l'EP a une utilité prometteuse dans l'évaluation intégrée des interventions qui font l'objet d'un développement continu, ce qui inclut l'app.

Table of Contents

List of Abbreviations

ART: Anti-Retroviral Treatment

CM: Cohorte de Montreale

CUSM: Centre Universitaire de Santé McGill

CHUM: Centre Hospitalier de l'Université de Montréal

CMQL: Clinique Medicale Quartier Latin

MUHC: McGill University Health Center

CVIS: Chronic Viral Illness Services

IDU: Injection Drug Use

PLHIV: People Living with HIV

PVVIH: Personnes vivant avec le VIH

List of Figures:

List of Tables:

Acknowledgements

I first like to acknowledge my principal supervisor, Dr. Tibor Schuster, who has provided guidance, support, encouragement, and constant reassurance. Your mentoring enabled me to understand the subject, satisfied my intellectual curiosity, and made me an avid learner and researcher. Vielen Dank! Secondly, I am grateful to my co-supervisor Dr. Bertrand Lebouché for his significant support and guidance on my research project. As a clinician myself, I deeply appreciate how you generously welcomed me into your clinic. This experience helped me truly understand the challenges faced by PLHIV, which further fueled my interest in platform trials. Merci beaucoup!

I am thankful to my thesis committee members (Dr.Tibor Schuster, Dr. Bertrand Lebouché and Dr. Alexendra De Pokomandy) for critically reviewing my project and providing me with constructive feedback that helped me improve my work. I like to thank Genèvieve Gore, for helping me develop my literature search strategy and identifying the relevant articles. In addition, special thanks to Stephanie Long, who edited my writing and suggested improvements. Grateful to genevieve.arsenault-lapierre who helped me with the French translation of the abstract. Thanks to Dr. Tibor Schuster and Dr. Bertrand Lebouché for the stipends and to Mitacs accelerate and Merck Canada (IT05857) for the grant that funded my studies I must acknowledge the travel awards that I received from CHIR and the Department of Family Medicine, McGill for the presentation of my research at international conferences.

I have deep gratitude for the moral support and unwavering love that my young children (Ashar, Eshaal, and Muhammad) gave me during my studies. They taught me that difficult tasks could be accomplished if we do one step at a time. I am entirely grateful to my beloved husband Aqib Yousaf for providing me with all the required resources for my education. Thanks for

taking care of our kids in my absence, taking me home in the harsh Montreal snowstorms, and pushing me to achieve my future goals.

Lastly, I am ever indebted to my parents, Qasra and Khuda Bukhsh, for their constant encouragement and prayers during the whole journey of my education and professional career. Thank you for indoctrinating in me the courage to overcome the obstacles and hardships in life and staying focused towards achieving what I have aimed for since my childhood.

Preface and Contribution of Authors:

The format of this thesis was traditional. Research projects were developed by Asma Aqib. Data for literature review and quantitative analysis were collected and analyzed by Asma Aqib. This thesis was written by Asma Aqib and improvements were suggested by Dr. Tibor Schuster, Dr. Bertrand and Stephanie long. However, my primary supervisor (Tibor Schuster) added few sentences to the proposed dynamic app development scheme in the Introduction chapter and helped me with the data imputations, and analysis of CD4 counts and Viral load data.

1 Introduction

The Human Immunodeficiency Virus (HIV) infection is a worldwide epidemic. According to a recent World Health Organization report [1], about 37.9 million people are living with HIV worldwide and 23.3 million were receiving antiretroviral therapy (ART) by the end of 2018 [2]. Incidence and prevalence of HIV have increased in Canada in the past decade [3]. According to the 'HIV in Canada-Surveillance Report 2017', a total of 2,402 new cases were registered in 2017, which corresponds to an increase of 17.1% since 2014 [3].The national diagnosis rate of HIV increased from 6.4 per 100,000 in 2016 to 6.5 per 100,000 population in 2017) [3].

Adherence to ART plays a cardinal role in reliably preventing HIV transmission and slowing down disease progression [4]. Adherence is referred to "a patient's ability to follow a treatment plan, take medications at prescribed times and frequencies, and follow restrictions regarding food and other medications" (p.835) [5]. Adherence to ART is essential for the achievement and maintenance of an undetectable viral load (defined in Canada as less than 40 -50 copies of the virus per milliliter of blood, depending on the type of test used) [6].

Antiretroviral treatment adherence rates, however, are troublesome to estimate due to the complexity of the available ART regimens and subjectivity in the adherence assessment [7]. Moreover, clinicians are often not able to completely evaluate treatment adherence of their patients and related barriers during their routine clinical consultations [8-11]. A potential solution to this problem is to utilize adherence-promoting tools and instruments that measure individual barriers to ART adherence using a patient-reported outcomes framework that helps to close this impertinent gap in patient-doctor communication [12].

Patient-reported outcome measures (PROMs) are defined as measurements of any facet of a patient's health standing that return directly from the patient (i.e. while not the interpretation

of the patient's responses by a health care provider or anyone else) [13, 14]. PROMs usually embrace general questions on health status, disability, exacerbation of symptoms, quality of life, and satisfaction with care [15]. PROMs have been utilized for monitoring, screening, and facilitating the delivery of patient-centered care to patients by their healthcare providers [16]. In the modern era of technology, electronic versions of PROMs (e-PROMs) can be delivered via mobile health applications (mHealth apps). This transition may play an increasing role in the digitalization of health and medical information [14, 17].

Indeed, the advancement of mobile technology has substantially impacted the health care system [18]. Presently, with the advancement of technology, electronic health interventions like mHealth apps are growingly available [18] and can help address various health concerns by engagement of patients in their disease management [19]. In addition, the adaptability of health apps vis-a-vis software tools' designs and features and a growing number of potential users (76% of the Canadian used smartphones by 2016 [20] and 1/3rd of Canadians use mobile apps to track the health [21]) has attracted health care providers and patients towards using mHealth apps for disease management [22, 23]. Above and beyond, it has been proposed by the United Kingdom Health Department that mHealth apps should be prescribed for chronic disease management such as diabetes, high blood pressure, and post-traumatic stress disorder, etc. [24-26].

In the context of HIV, advancements in ART have transformed HIV from a death sentence to a well manageable chronic condition [19, 27], given that patients are adherent to their treatment. Over the past two decades, there has been a shift in the demographics of HIV patients from intravenous drug users and those who are predominantly illiterate to the patients with higher literacy and greater knowledge of technology [24, 28]. The changing patient profile and

sociodemographic characteristics suggest that the clinical and self-management strategies need to be re-tailored [24].

Rapidly advancing mHealth technology and extended the coverage of cellular networks has swiftly enabled the use of mHealth apps for delivering self-management interventions for HIV care [19, 29]. Mobile health application-based interventions are considered effective for improving adherence-based behavior [29]. For example in a trial conducted on PLHIV showed that ART adherence was improved by a mHealth app that contained imagery illustrating the concentration of ART and its effect on the immune system [30]. Moreover, mHealth-based interventions are cost-effective and scalable for use for people worldwide living with HIV [29, 31, 32]. Notably, barriers to ART adherence may differ between individuals who are initiating therapy compared to patients who have taken it longitudinally [29, 33]. Therefore, in order to be effective, a mHealth app has been recommended to be personalized according to the stage of adherence to the ART [34] and to exclusively target the determinants of non-adherence [29].

To date, most health apps used for HIV care are not optimally tailored to the needs of the individual user, therefore, they have failed to attract the broad attention of users and healthcare providers [24, 35]. Areas for improvement include data safety, confidentiality, and quality of content [24]. Robustillo et al. proposed that "…the application should be designed so that the user can use it efficiently and effectively without adaptation" (p.731) [24]. He also suggested that the mHealth app design should be universal, and it should clearly define who will be the end-users and what will be the purpose and class objectives of the mHealth app. [24]. To achieve these goals, engagement of patients, public health practitioners and software developers for the design and development of the evidence-based intervention is necessary [24, 35].

I-Score (CTN 283 study):

The I-Score study was developed in responding to gaps in both the research and clinical care of HIV. The core objective of the study was to develop and validate a PROM for electronic administration (via a mHealth app), to comprehensively assess perceived barriers to ART adherence in PLHIV being followed at clinical centers in Canada and France. The PROM's conceptual framework was generated based on a synthesis of qualitative studies with PLHIV in developed countries [36]. In addition, several stakeholder engagement initiatives were conducted, including the formation of a patient advisory committee [37] and consultations with HIV clinicians [38] aiming to ensure the instrument's relevance. The ultimate study goal is to deliver the intervention (PROM) via a mHealth app and to continue developing various aspects and design features of the mHealth app to increase usability, acceptability and overall utility of the application.

A central aspiration of the I-score study is to accommodate patient diversity at different stages of ART treatment and associated adherence challenges [34] as well as to contextualize, personalize and generalize the mHealth app to a diverse population of PLHIV in Montreal.

Nevertheless, the development of a user-friendly and useful mHealth app is challenging, particularly, in terms of contextualization and personalization [39]. Given that, end-users may have different sociodemographic and clinical characteristics like the severity of illness, overall health status, and general prognosis. However, some of the difficulties in developing and establishing mHealth apps could be potentially addressed through innovative and cost-effective trial designs that enable rapid evaluation of various aspects of an application (e.g. adaptive trials). Such novel trial designs might be beneficial in the context of mHealth application development, as the inbuilt software is typically dynamic and continually undergoes updates

[40]. Therefore, a dynamic, cost-effective and efficient trial designs such as adaptive trials could be useful in the evaluation of mHealth app development. This is akin to the evaluation of the safety and efficacy of treatments in clinical trials.

Modern adaptive randomized controlled trials allow for a continuous and unbiased comparison of study arms while aiming to minimize the overall study costs and a number of patients being exposed to inferior treatment options [41]. Novel adaptive designs (ADs) have been recommended by Bajpai, R. and J. Car for the evaluation of e-health interventions (e.g. mHealth apps) [42] trials, however, these approaches have not been widely used in the context of measuring and evaluating PROMs.

Platform trials (PTs) are a promising subclass of adaptive RCTs with great utility in the systematic evaluation and development of mHealth apps. The objectives of the I-score study are well aligned with a dynamic approach for mHealth app development using a PT design. The development and evaluation of the app through a PT design will involve patients, clinicians, and software developers throughout all study phases.

Proposed dynamic mHealth app development

A mHealth app for measuring PROMs can be evaluated with regard to various usability and utility aspects. Clinically relevant usefulness for clients (patients) and their interconnected care providers is, arguably, the foremost objective of a mHealth app. The need for improving or adding app features and functionalities may emerge directly through user feedback or might be indicated through general developments in the respective market segment. Feature updates and changes to app functionalities typically yield in a specifically modified (‘updated’) version of the previous software released (Figure 1.1). To enable an unbiased evaluation of a newly released app version compared to one or multiple preceding versions, random assignment of clients (and

their respective health care providers) to either version under comparison is desired. This is in direct analogy to the principle of 'randomized controlled trials' that are established as gold standard in confirmatory medical research for testing new treatments or interventions. In platform trials, randomization depends on the available information regarding the 'performance' of each trial arm overtime i.e. allocation probabilities are dynamic and determined by pre-specified adaptation rules that satisfy desired statistical optimization criteria (Figure 1.2). In my master's thesis research project, I investigate the applicability of the PT design for developing mHealth applications within the context of measuring patient reported outcomes i.e. barriers to ART treatment adherence reported by PLHIV.

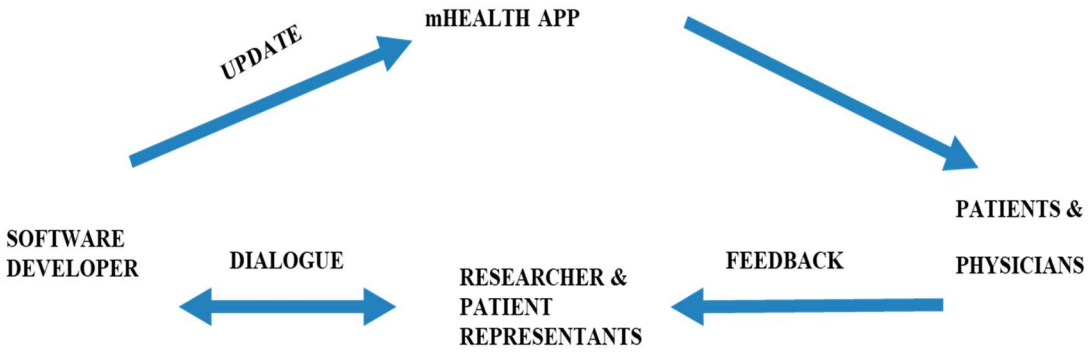

Figure 1.1: Schematic representation of the dynamic app development.

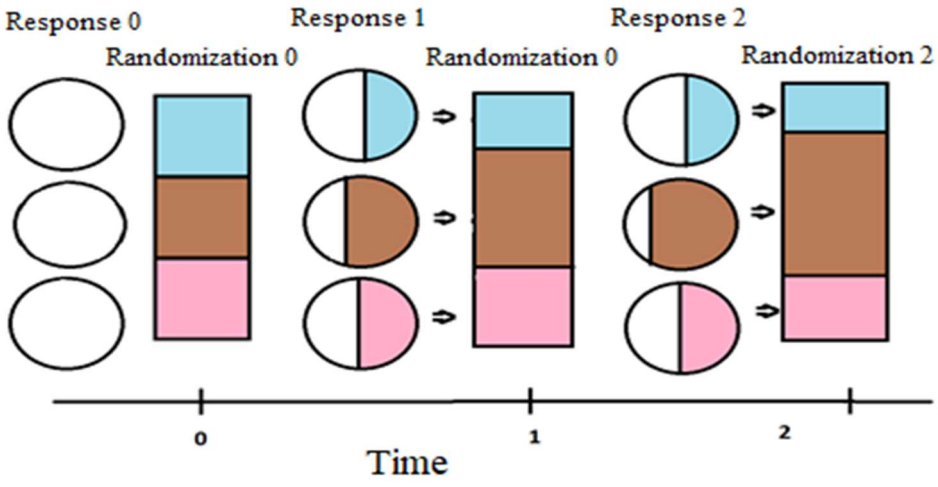

Figure 1.2 : Schematic representation of response-driven adaptive randomization.
Explanation: Initially all app versions have equal randomization probability. As the trial progresses, the randomization ratios will be adapted (using Bayesian or reinforcement learning algorithms) after interim results are available. Ultimately, more participants will be randomized to the best performing trial arm (shown as brown).

Response-driven adaptive randomization is the fundamental principle of the PT that distinguishes it from classical RCTs using fixed randomization schemes. Nevertheless, various ethical, pragmatic or technical issues associated with RAR must be considered before initiating a PT.

Typically, mHealth apps undergo continuous development (i.e. updates) to improve the user-uptake and the perceived value of the app [18]. These updates typically serve to improve revenue or value on the producer side (i.e. the quality, reliability, and usability of the data that is collected by the app), while also increasing the appeal to customers (i.e. usability and perceived value of the mHealth app) [18]. PT, therefore, appears to be a promising approach, not only regarding the identification of better treatments but also to develop a better understanding of barriers to treatment adherence within PLHIV. These patient-reported treatment adherence barriers can be collected with specifically designed electronic health applications.

Newly available tools for measuring barriers (e.g. mHealth apps) can be perceived by health care providers and researchers as treatment-complementary alternatives for obtaining

better health outcomes [24]. All advantages (cost-effective, efficient, requiring small sample size, accommodating the heterogeneous population, etc.) of the PT approach seem to apply also in this context. However, no systematic assessment of the feasibility of PTs in the context of developing mHealth apps has yet been performed. Here, I propose to conduct a descriptive literature review of platform trials conducted in clinical settings (i.e. trials conducted on human subjects).

Likewise, the literature does not inform about the ideal characteristics of the population to be recruited in platform adaptive trials conducted in HIV context. Therefore, before the implementation of the PT, it is critical to characterize the target population i.e. PLHIV taking ART in Montreal, Quebec as such information is pertinent to proper planning and setup of a PT to ensure generalizability of the trial outcomes. Similarly, this objective corresponds to the overarching aims of the I-score study to contextualize and personalize the mHealth app for PLHIV in Montreal; while accommodating diversity in their stages of treatments and sociodemographic characteristics.

The primary goal of my study is to investigate the feasibility, of the 'Platform Trial Design' for the development of a mHealth app incorporating patient-reported outcomes and identifying barriers to treatment adherence in HIV patient populations.

Research objectives:

1) To determine the ethical, pragmatic and technical issues that can affect the implementation or acceptance of platform trials (stated in the clinical trials literature) that are potentially relevant in the context of developing mobile health applications for improving treatment adherence in HIV patients.

2) To characterize a study population, PLHIV under ART included in the 'Cohorte de Montréale' (CM) (MUHC REB #2011-942). CM includes PLHIV in Montreal who receive care at one of four major HIV-specialist care centers. 1) Clinique médicale l'Actuel, 2) Clinique medicale Quartier Latin (CMQL), 3) Unité hospitalière de recherche, d'enseignement et de soins sur le sida (UHRESS) at the Centre hospitalier de l'Université de Montréal (CHUM); and 4) Chronic Viral Illness Service (CVIS) at the McGill University Health Centre (MUHC)/ Centre Universitaire de Santé McGill (CUSM).
3) To characterize a target population, PLHIV under ART, followed at Centre Universitaire de Santé McGill (CUSM*/ Chronic Viral Illness Services (CVIS†) clinic, to be included in a future platform trial evaluating a mHealth app measuring barrier to ART adherence.
4) To identify commonalities and differences in characteristics of the potential target population (PLHIV on ART followed up at CUSM) compared to the study population (PLHIV on ART followed up at clinics included in CM).

Research Questions:

1) What ethical, pragmatic and technical issues may affect the implementation or acceptance of platform trials in clinical trial settings that are potentially relevant to the development of mobile health applications aimed at improving treatment adherence in HIV patients? 2) Are there systematic differences in the sociodemographic and treatment adherence characteristics of PLHIV with ART followed up at CUSM (CVIS clinic) compared to PLHIV in the Cohorte de Montreale (e.g. CHUM, CMQL and L'Actuel) in the year 2016?

---

* Centre universitaire de santé McGill (CUSM) is also known as McGill University Health Centre (MUHC).

† Chronic Viral illness services (CVIS) is located at CUSM (MUHC).

2 Literature Review

Premise

In this section, I will situate the introduction of adaptive designs (ADs) in comparison with traditional randomized control trials (RCTs). I will particularly focus on the PT design including the definition, types, context and various advantages of the trial design. Later, I will explain in detail the steps that I followed for conducting a descriptive literature review and will report the results and highlight the relevant research gap.

Adaptive trials vs. Randomized controlled trials:

Conventional RCTs are different from ADs as these allow fixed randomization only. The participants are equally likely to be randomly assigned to either the "intervention" or the "control" groups (Fig 2.1). Individuals assigned to the "intervention" group receive the new intervention or treatment, while those in the "control" group receive either the gold-standard intervention or a placebo [43]. At the end of the trial, one or multiple pre-specified outcome (s) are then compared between the study groups to determine whether the intervention or treatment is efficacious [43] (Fig 2.1).

In fact, RCTs have been employed in clinical research over decades, and are considered the gold standard of determining evidence [43]. Nevertheless, RCTs are expensive in terms of time and resources and conclusions can only be drawn at the very end of a trial [43].

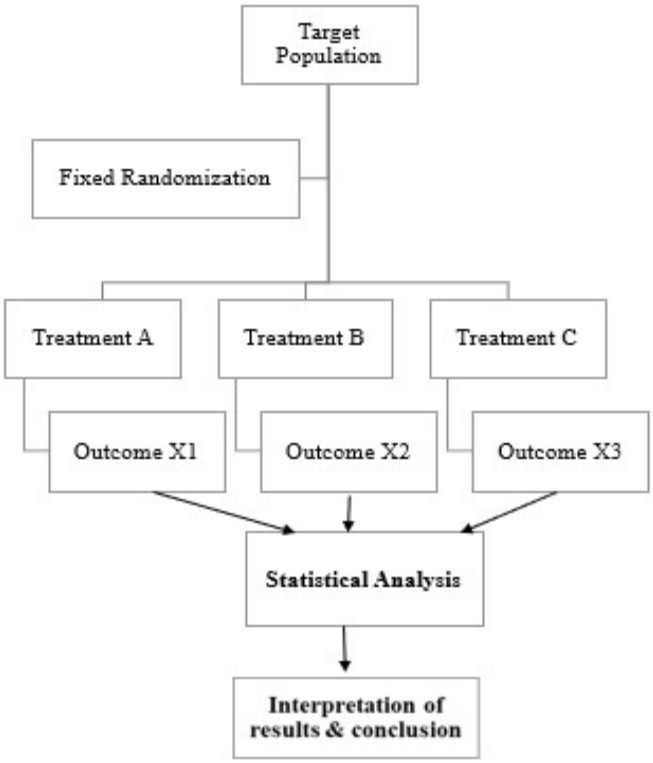

Figure 2.1 : Traditional Randomized Control Trial.

With recent advancements of technology and diagnostic tests, modern RCTs incorporate, more and more, pre-specified adaptation rules to enable better-personalized treatment of trial participants and to minimize the overall risk for the futility of a trial [44].

The Food and Drug Administration (FDA) defines ADs as studies that have a prospectively planned chance for modification of one or many particular aspects of the study design and hypotheses supported analysis of data (typically interim data) from subjects within the study (p.2) [45].

Adaptive designs allow for **interim analyses** in which incoming data is analyzed and used to inform the adaptation of certain protocol aspects. Such aspects may include the choice of primary study endpoints, (dis-)continuation of study arms or the randomization schedule. In

adaptive designs that employ adaptive randomization, the proportions of individuals being allocated to treatment and control arms dynamically change (adapt) to the evidence being generated until the interim analysis; with the goal of allocating a higher proportion of participants to the most effective intervention (Fig 2.2).

The call for more and better-personalized medicine in the modern era demands innovative trial design to efficiently (and more rapidly) evaluate multiple interventions [41] including e-health interventions [42]. Platform trials are a relatively new class of ADs, which allow for confirmatory evaluation of multiple interventions while meeting some of the challenges of personalized medicine [46].

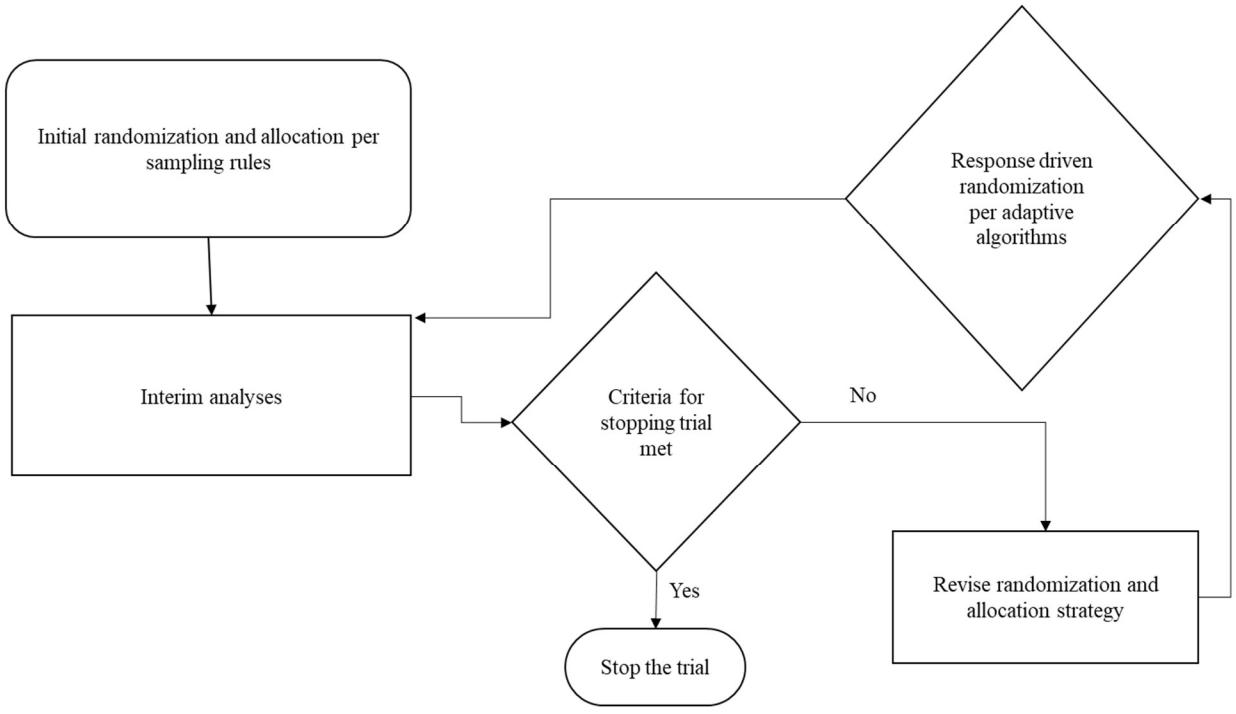

Figure 2.2 :Schematic representation of Adaptive trial designs [47]

2.1 Platform trials

2.1.1 Definition

According to Scott M. Berry, "A platform trial is defined by the broad goal of finding the best treatment for a disease by simultaneously investigating multiple treatments, using specialized statistical tools for allocating patients and analyzing results" (p.1619) [48]. Response Adaptive randomization (RAR) and Multi-arm Multistage design (MAMS) are two subgroups of PTs [46] (Fig 2.3).

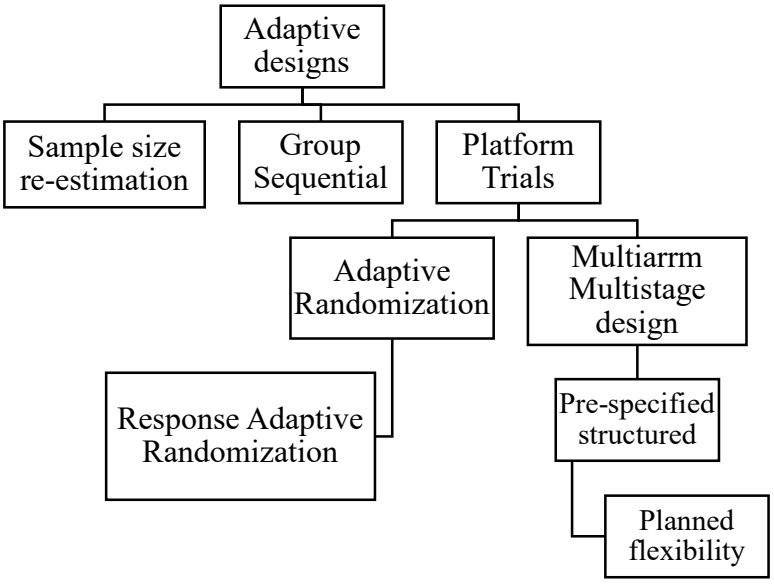

Figure 2.3: Terminology and classification of Adaptive designs [46]. Reproduced with permission of authors (Appendix B)

2.1.2 Response-adaptive Randomization (RAR) / Outcome-based randomization:

In RAR, it is anticipated through outcome-driven adjustment of the randomization scheme that a higher number of patients receive the most effective treatment under study [49] (Fig 1.2). Moreover, when multiple treatments are evaluated, within each interim analysis, ample information regarding the desired treatment effect and possible unintended effects (i.e. adverse

effects of a drug) becomes available. Concurrent planned modifications of the trial shorten the evaluation time for the best intervention.

In an effectively planned, designed and conducted PT, the majority of trial participants will receive the best intervention [50]. If differences in outcomes emerge across treatment arms, the RAR maintains a high power even with a relatively low sample size [51]. Nonetheless, to achieve effective adaptive randomization, primary endpoints need to be accurately determined [52].

2.1.3 Multi-arm Multi-stage design (MAMS)

MAMS designs allow comparison of multiple treatment groups against a single control group, which has an ethical advantage of depriving a smaller number of participants from a potentially useful treatment. Like platform trials and other adaptive trial designs, adaptive features of the trial are bound to be pre-specified criteria and pre-trial selection rules, which allow planned flexibility [53]. Expressed in lay terms, the MAMS design follows a strategy of "drop all losers"‡ and "keep all promising" (keep a pre-specified best treatments) [53]. Pre-planned interim analyses are a core design aspect that predefines and controls the statistical power and type 1 error of the study and consequently, preserves the operational feasibility of trial [53].

2.1.4 Similarities and differences between RAR and MAMS designs

Both RAR and MAMS designs allow planned modifications of the study conduct according to the results of interim analyses. In RAR designs, participants in the superior treatment arm (according to interim analysis results) are increased whereas MAMS design

‡ "Drop-the-losers designs are statistical designs which have two stages of a trial separated by a data-based decision. In the first stage k experimental treatments and a control are administered. During a transition period, the empirically best experimental treatment is selected for continuation into the second phase, along with the control"54. Sampson, A.R. and M.W. Sill, *Drop-the-losers design: normal case.* Biometrical Journal: Journal of Mathematical Methods in Biosciences, 2005. **47**(3): p. 257-268.

focuses on determining the stopping boundaries for the least efficacious arm [46]. According to Wason and Trippa [55], if multiple interventions being compared are efficacious, MAMS design performs superior to RAR designs. Oppositely, if only one treatment is effective, RAR performs superior to MAMs.

2.1.5 Characteristics of Platform Trials

Platform Trials embrace several different sub-trials or cohorts under one principal overarching clinical adaptive trial protocol called **master protocol** [49]. The master protocol is defined as: "any top-level or overarching clinical trial protocol that comprises several parallel biomarker-based, or genomically based sub-trials or cohorts" (p.218) [49]. Hence, PT designs accommodate cohorts or sub-trials that can be different from each other with respect to genome, biomarkers, or other characteristics of interest [49]. PTs encompass a **dynamic strategy** that allows sub-trials (i.e. treatment or intervention arms) to be added or eliminated from the trial based on either graduation (for being successful) or failure (lack of demonstrating efficacy). The design typically follows **a Bayesian strategy** for randomization that summarizes all available information regarding the efficacy of the trial arms understudy, with increasing probability of randomizing participants into more efficacious treatment groups [41, 49](Fig 2.4).

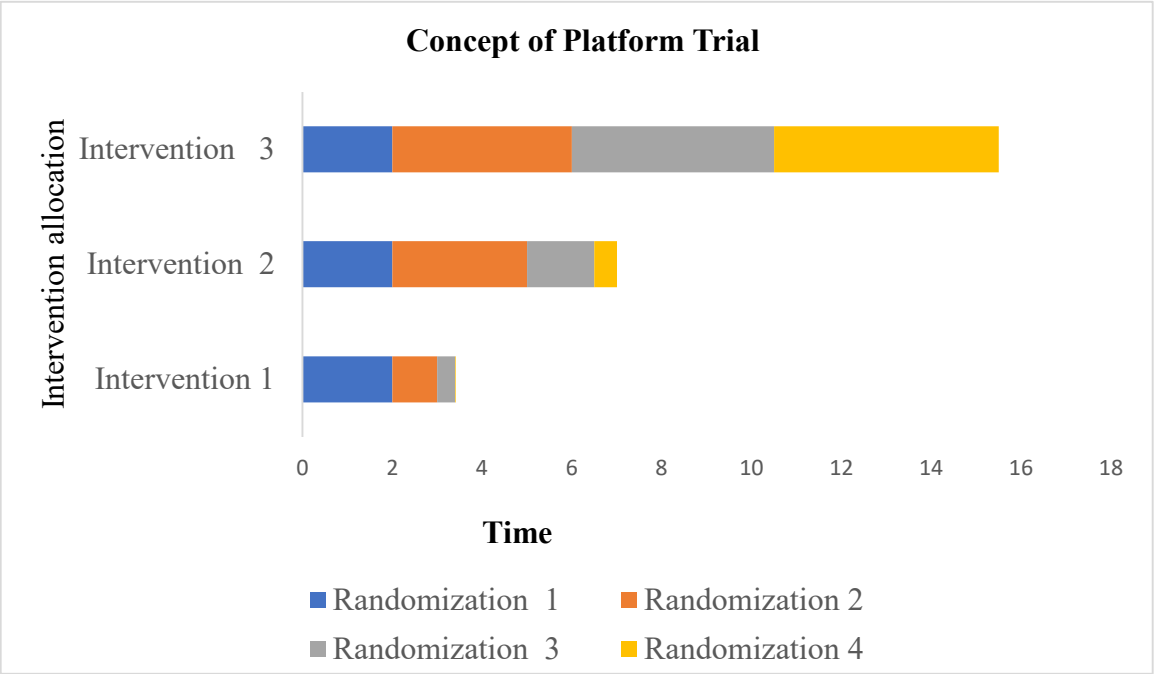

Figure 2.4: Platform Trials

Explanation. Initially all the participants have equal randomization probability for all the study arms (intervention 1,2 & 3). However, after interim analysis the randomization probability changes in favor of successful arm.

2.1.6 Types of platform trials

Open (perpetual) platform trials

The open PT design is flexible and allows new treatment arms to be added or dropped during the trial. Multiple treatment arms are compared against a single control arm [41] (Fig 2.5).

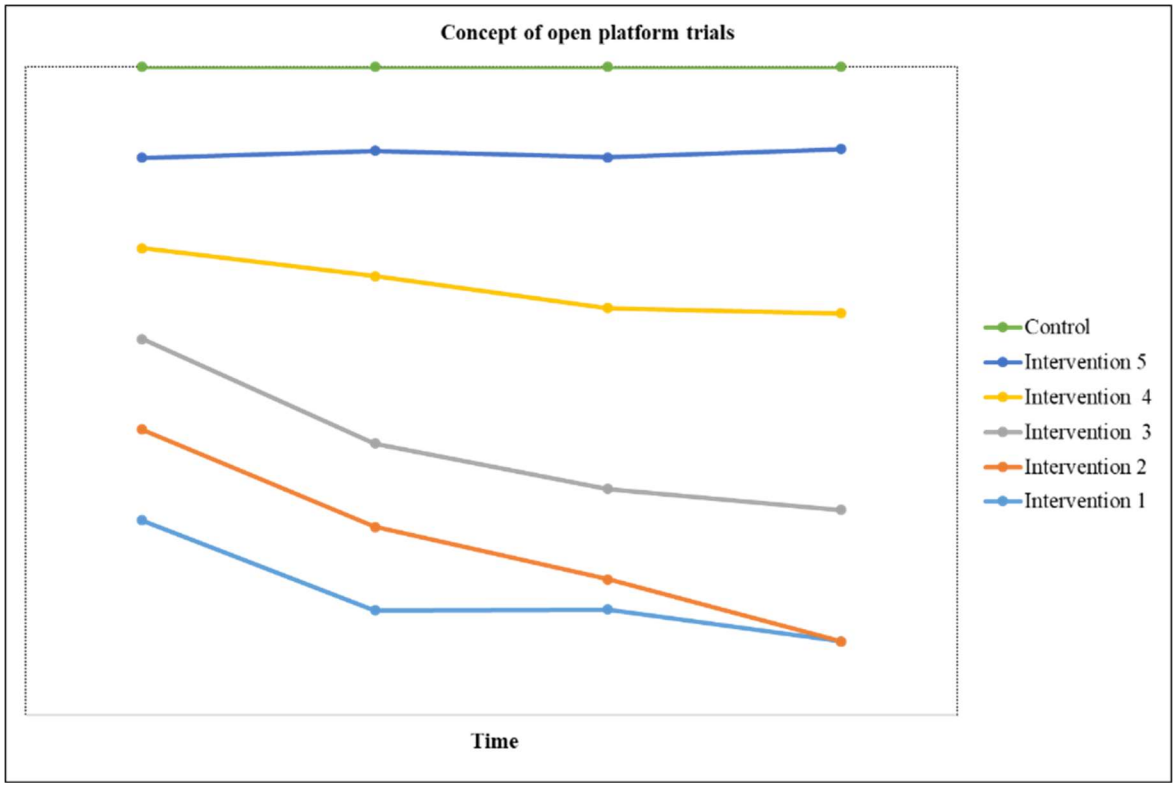

Figure 2.5: Open Platform Trials (reproduced from [41])

Expalanation: Five interventions are compared against one control arm. Ineffective interventions are replaced with the new interventions untill a significant difference is found.

Closed platform trials

Unlike perpetual PTs, in closed PTs, additional treatment arms may not be added to the trial. Each treatment arm has a fixed sample size, randomization scheme, and its own control [41] (Fig. 2.6).

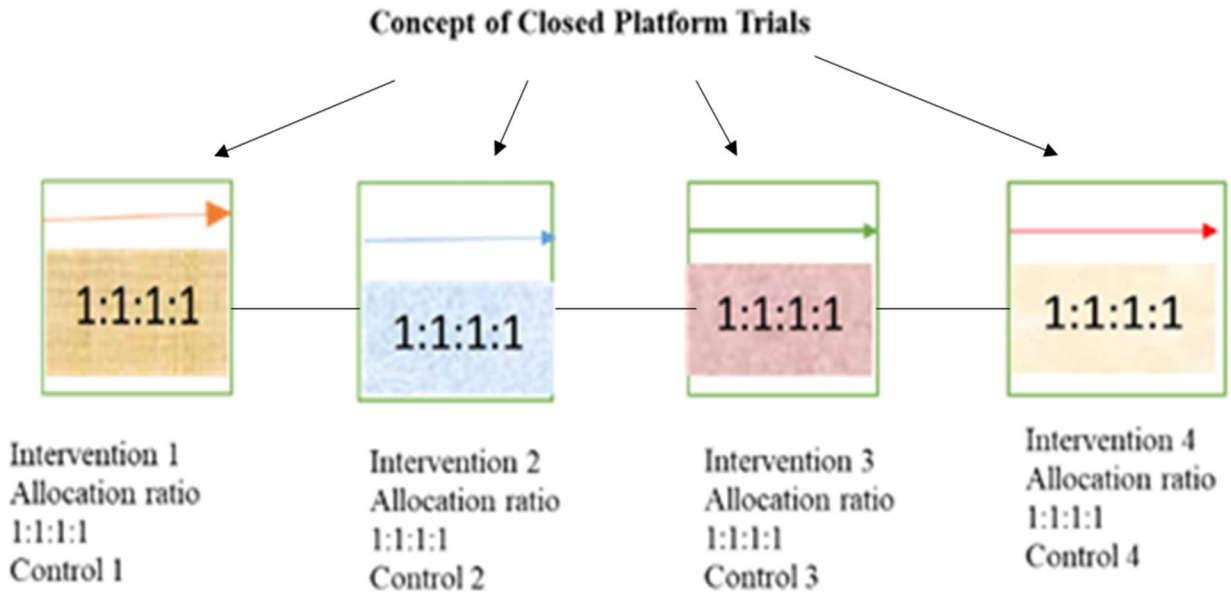

Figure 2.6: Closed Platform Trials [41]

Explanation: Each box indicates a treatment group, and different colors of arrows indicate different control within each treatment group.

2.2 Advantages of Platform trials

2.2.1 Efficiency

Platform trials have recently caught the interest of pharmaceutical companies for being efficient and pragmatic in multiple treatment comparisons [46]. Many multi-centered trials like FOCUS4 [56], STAMPEDE [57], SPY2 [58], and BATTLE [59] have successfully been conducted [46]. PT designs have the potential to efficiently reduce required sample sizes, save time and costs of a drug development process without undermining the validity and integrity [46].

Ariel D. Stern and Sarah Mehta explained well the efficiency of adaptive platform trials using a case study [60]. According to them: “Each time a new trial began, regardless of its design, investigators invested time and money into developing a study protocol, hiring a CRO, recruiting a team of clinicians and statisticians, securing funding, obtaining ethical approval,

identifying clinical trial sites, working with regulatory agencies, instituting standard operating procedures, establishing contracts, and enrolling participants. Because adaptive platform trials were ongoing, many of these "fixed costs" could potentially be shared across multiple treatment arms in a way that was not possible in traditional RCTs. Some investments, such as CRO selection, ethical approval, and statistical design needed to occur only once, at the time the platform was launched. Thereafter, new drugs could be added to the trial continuously with lower start-up costs relative to what would be required to launch a new traditional RCT" [60](page 7-8)(Note: I am quoting this with the permission of Harvard University) (Appendix B).

2.2.2 Cost Effectiveness

Traditional RCTs use the frequentist approach for statistical analysis which is scientifically rigorous; however, may result in an overly large trial setup requiring high costs [61]. In contrast**,** Bayesian adaptive trials (e.g. Platform trials) can effectively reduce the costs of drug (or intervention) development without compromising the scientific rigor of the study [61].

In addition, PTs establish a long term resource that may combine various funding resources like federal, pharmaceutical and for-profit organizations for drug development [48]. Since, PT designs use the same trial infrastructure, a master protocol, underneath which multiple interventions can be evaluated, this single infrastructure saves associated infrastructure costs required for the evaluation of multiple treatments and interventions. Furthermore, the ability of PTs to drop ineffective or futile treatment arms potentially saves substantial costs in carrying on investigating these trial arms[50].

2.2.3 Operationally feasible

Platform trials embrace operational seamlessness and efficiency [49]. A predefined master protocol enables researchers to determine various operational characteristics of a trial in

advance e.g. type 1 error, sample size, and power [62] which predetermines the operational feasibility and success of the trial. In addition, PT designs overcome the barriers for trial initiation and entry of new trial arm.

A few concrete examples support this desirable attribute of platform trials as the parallel entry of new sub-trials (e.g. based on new markers) does not require the larger trial to be stopped because of modification of the trial's protocol [49]. Likewise, if drug graduates from phase II, it can immediately enter a preplanned phase III confirmatory trial [49] without unnecessary delay. These examples indicate the usability of this novel adaptive design in comparative effectiveness clinical trials. In addition, Bayesian adaptive trials such as PTs conducted for comparative effectiveness research enable continuous quality improvement as compared to conventional non-adaptive trial designs [62].

2.2.4 Heterogeneity

Traditional RCTs are typically conducted in homogenous populations and large variation in terms of biomarkers, genetics, the severity of illness, etc. is not considered which may substantially affect the trial's generalizability [48]. In contrast, PT design explicitly considers and adjust the heterogeneity in the study populations, with a goal to find the best treatment match for the patients defined by these subgroups [48].

2.2.5 Preplanning

Important aspects of study design can be pre-planned such as intervention allocation strategy, conduct and methods for data analysis [63]. If required planned changes can be done even while the trial is in progress, for example, re-estimation of sample size, re-definition of study end-points and stopping rules [63]. However, such potential changes must be considered

and incorporated at the design stage of the trial and extra costs or resources required to implement these modifications can be calculated in advance [63].

2.2.6 Acceptability among stakeholders

According to the US FDA; flexible trial designs (e.g. PT) are more acceptable to stakeholders than non-flexible trial designs[45]. Patients may be more willing to participate in a trial in which there is a greater likelihood of receiving the most effective intervention [45]. Likewise, the funding sponsors may feel encouraged to participate in a flexible trial, allowing planned modification and early decisions [45]. The platform trial design is particularly acceptable to pharmaceutical companies as it opens the door for testing and developing drugs for rare diseases and diseases with high and fast mortality e.g. Glioblastoma and Ebola virus [60].

Objectives and research question

These enormous advantages of PT design have drawn the attention of the I-score study to implement this design for the evaluation of multiple versions of a mHealth app for measuring patient-reported outcomes. However, no systemic inquiry about the feasibility of platform trials in the context of developing a patient-reported outcome-based mHealth app has been done so far.

In addition, the concept of PT is yet emerging and varying terminologies are alternatively used in the literature (e.g. platform trials, adaptive platform trials, Bayesian adaptive randomization, outcome-based randomization, response adaptive randomization, multi-arm multi-stage design, etc.) and diversity in the context of trials' conduct is alike.

A few Platform trial/RAR/ MAMS were conducted on animals (rats), some focused-on mortality and survival outcomes. Many of the articles have a pure methodological and statistical calculation of adaptive algorithms and a few articles are descriptive that define PT, explain the

advantages or disadvantages and ethical issues associated with PT. However, information regarding ethical pragmatic and technical issues is dispersed in the medical literature.

Reasons for conducting a descriptive review in a comprehensive way were to 1) conceptualize the ethical, pragmatic and technical issues related to clinical PT. 2) extract and summarize the disperse information regarding these challenges from the published literature. 3) to identify the gaps in the literature. 4) summarize the literature to generate new perspectives on the topic (e.g. how can we use platform trials in the context of developing a mobile health application employing patient-reported outcome measures). 5) inform research (specifically the I-Score study).

I addressed the following question:

What ethical, pragmatic and technical issues may affect the implementation or acceptance of platform trials in clinical trial settings that are potentially relevant to the development of mobile health applications aimed at improving treatment adherence in HIV patients?

Literature review methodology (Phase 1)

The literature review was developed in the light of Sharon D. Kruse's guide for the master's student for " Developing a comprehensive review: an inquiry into method" [64]. This article provides guidance for students on conducting a comprehensive review of the literature for the thesis or dissertation required for the synthesis of results and discussion.

Following Dr. Kruse's instructions, I conducted this literature in a systematic way using an array of strategies for identifying and selecting the relevant literature, recording, understanding and presenting information pertinent to the topic of interest [64].

Working with an academic librarian (GG), a search strategy was developed and implemented using two research databases: EMBASE (Ovid) and Medline (PubMed) on Feb 05,

2018 (Appendix A, Tables 1 and 2). The search was limited to the English language because of my linguistic competence, limited time and resources for translation. Nonetheless, I anticipate that this limitation may not bias the findings of review as some evidence from the literature suggests [65]. In addition, to portray a full picture of the literature, I decided to include all the articles since the establishment of the databases (EMBASE 1947 & PubMed 1966) to 2018.

In the pilot search, I included the following concepts: 1) Platform trails (PT) /Multiarm-Multistage (MAMS)/Adaptive randomization (AR) 2) HIV/Chronic disease (Appendix A Table 2). The PubMed search yielded 419 studies out of which 320 were related to the concept of PT/MAMS/AR. However, the broader concept of PT/MAMS/AR was the primary focus of my research. In addition, published studies were not too many; therefore, as advised by the librarian, I restricted the search strategy to the key concept of platform trials and its synonyms: Multiarm-multistage design and response adaptive randomization. All the keywords related to the concept of platform trial, MAMS and AR were combined using the Boolean operator "OR" which yielded the relevant research.

Table 2.1: Eligibility criteria

Inclusion Criteria	Exclusion Criteria
All empirical research (quantitative randomized/ quantitative non-randomized, quantitative descriptive) conducted on human beings only.	Exclude non-medical/patient settings or the trials conducted on non-human subjects or based on mortality or survival outcomes (not compatible with the context of I-score study).
All descriptive summaries, reviews, letters, etc.	Technical/methodological papers without real-world applications in a medical/patient setting (e.g. development of a specific Bayesian algorithm).
Only platform trials, adaptive randomization, and Multiarmed multistage design in a medical/patient setting	Not explicitly mentioning adaptive randomization, platform trials or MAMS design.
All countries included	
Limits: 1) Published in English (because of my linguistic competence, time and resource limitation 2) Limited to 1947-2018.	

Later on, another search was conducted on Nov 2, 2018, in PubMed using two key concepts 1) platform trials and 2) patient-reported outcomes (PROs). Combinations of relevant Mesh/index terms, subject headings, and keywords (Appendix A Table 5) were used, including all other synonyms or relevant words or spelling variations for each key concept. The search for the concept of PT when combined with PROs yielded zero results, indicating that this trial design has never been used in the context of developing patient-reported outcomes. I reconfirmed these results on April 24, 2019 (Appendix A Table 3 and 4). Nonetheless, I restricted the search strategy to the concept of PT/MAMS/AR only.

All the retrieved articles (titles and abstracts) were exported to Endnote X8.0.2 a citation management software. All duplicated records were deleted. To facilitate the title and abstracts screening, subsequently, endnote deduplicated references were exported to Rayyan software (https://rayyan.qcri.org). Articles that passed the initial title and abstract screening were imported

to EndNote x 8.2. The full-text articles were searched (in pdf format) using the McGill library search engine were downloaded and stored (as pdf) in a desktop folder.

Due to limited time and resources for hiring another reviewer, I independently screened titles and abstracts and full text-articles by applying predefined eligibility criteria (Table 2.1). The eligibility criteria were iteratively developed after frequent discussions with my thesis supervisor (TS). Similarly, we developed data extraction charts in Microsoft Excel to facilitate the storing and organization of data (Appendix A>data extraction charts).

Henceforth, after the full-text review, the following relevant information about PT/MAMS/AR design was extracted about the disadvantages, ethical, pragmatic and technical issues associated with PT in clinical settings. Given the review investigates how were the attributes of PT/MAMS/AR described in the literature, relevant information was extracted from the entire article i.e. from the abstract to the conclusion.

Following Dr. Kruse's instructions about the data charting and organization of the information, I charted relevant and important points, topics, ideas, summary, abstracts and conclusions with specific attention to the citations, methods, major findings, conclusions, etc. (Appendix A> Data extraction charts> chart1). As recommended, I built the charts in the order I read the articles. Afterward, I aligned the literature in a historical manner (from most recent to the least recent) by using the word processing sort tool which helped me to review the evolution of my topic. Then I reviewed the citations and identified or traced the researcher or the study whose name was repeated more than three times. This assured me that I did not miss any of the important aspects of the literature. After the first reading, I specifically paid attention to findings (important points) and tried to identify the most common themes. I could identify the theoretical concepts, research gaps, identified in the previous literature and got an overview of important

aspects of the literature. Then I organized the themes identified in the literature into a second chart (Appendix A >Data extraction charts>chart 2a &2b). This shorter and comprehensive chart helped me to organize my thoughts and the subsequent work into concise categories of themes and sub-themes. Then I wrote down the information of the columns, completed each section stepwise, and then added a synthetic conclusion to the final document as proposed by Dr. Kruse.

The search was conducted in PubMed (Medline), EMBASE (OVID). The search strategy yielded a total of 766 articles. Upon removal of duplicates, 459 articles were screened for titles and abstracts. After applying eligibility criteria 295 records were excluded as they were not explicitly MAMS/PT/AR trials and were conducted on animals and were accounting survival and mortality outcomes. The remaining 164 articles were eligible for full-text review. Out of these 137 were excluded as these were in non-medical context, too methodological papers without real-world data application. Only 27 articles were included in the final synthesis (Fig 2.7)

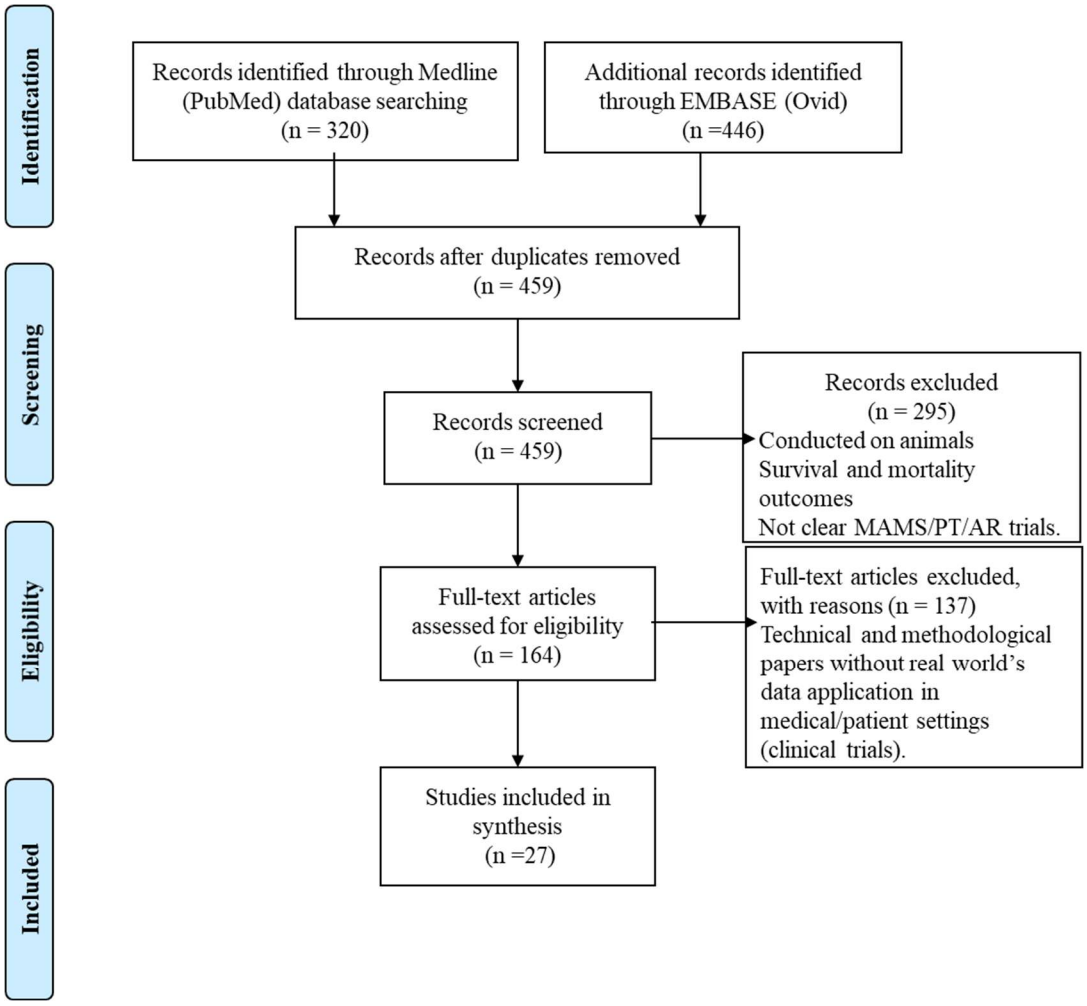

Figure 2.7: PRISMA Flow diagram

2.3 Summary of the relevant literature

The following subsections summarize pertinent information on platform trials that were extracted from the relevant literature.

2.3.1 Ethical issues

Response-adaptive randomization (RAR) is widely considered ethical as the randomization allocation is updated according to the interim results, and a great proportion of participants are allocated to the best performing treatment arm [66]. Nevertheless, several authors have argued the ethical implementation of PT in terms of loss of equipoise, injustice,

complexities of informed consent, and some potential threats to the integrity and validity of the studies [67]. Details of which are elaborated on in the following section.

Equipoise

Equipoise refers to "a state of genuine uncertainty on the part of the clinical investigator regarding the comparative therapeutic merits of each arm in a trial" (Section14, abstract) [68]. Equipoise is considered to be an ethically indispensable condition in clinical research and clinical trials [68]. According to Freedman, research is ethical, only if all the norms of beneficence (maximizing benefit) and non-maleficence (minimizing harm) are applicable [68]. Therapeutic obligations of the clinicians strictly enforce them to enroll their patients into trials where there are increased chances of receiving a maximum effective intervention/treatment [68]. If a clinician is involved in a trial, equipoise disappears as soon as he observes one treatment performing better than another [69].

In ADs, stringent adherence to equipoise is possible only at the beginning of the trial when the performance of the competing trial arms is unknown. However, in RAR, once the interim data (results) are available, the equipoise potentially disappears [69]. Nevertheless, it may return if the success rates of the two therapies are equal in the long run [69].

Proponents of RAR like Scott Brain [70] contended that absolute equipoise or uncertainty is not mandatory for the ethical conduct of a trial. Although preliminary results may give some clues towards the best effective therapy; they are typically insufficient to be considered as evidence [70]. He further argued that research is ethical if therapeutic obligations based on 'prima facie'[68] are not forgone. In some scenarios, equipoise is not compromised as if intervention under test has minimal or no side effects or if the patient's ultimate health outcomes are not affected by withholding standard treatment [70].

Thall [69] mentioned that the development of computer software like MDACC enabled ethical the implementation of adaptive randomization. Such tools preserve equipoise in clinics and have brought up substantial change in oncology [69]. Hence, it can be assumed that the development of software tools that can determine the randomization allocation scheme can be helpful in preserving the norms of equipoise.

In short, equipoise is considered an important requirement for the ethical conduct of a trail. [71]. Beyond the notion of equipoise, the Declaration of Helsinki states that "In medical research involving human subjects, the well-being of the individual research subject must take precedence over all other interests" (p.1532-33) [72].

Informed consent

A valid informed consent obtained from trial participants has four basic components: capacity, disclosure, comprehension and voluntary agreement [73]: I will focus on *disclosure* and *comprehension*. *Disclosure* is a process of revealing and discussing information that patients must consider when deciding to participate in a research study. This information typically includes details about the treatments under study and the chance of receiving one therapy or another (the randomization scheme) [70]. *Comprehension*/understanding of risks of participating is only feasible when patients are competent and autonomous and how transparently a physician/researchers inform the participant about the research study for which they are being recruited [69].

The adaptive randomization scheme is complex and may not be understandable to participants, which will ultimately compromise autonomous decision-making [70, 74]. Buyse, Saad, and Burzykowski commented: "…adaptation mechanism will remain a black box to many patients, amplifying the challenges of effective communication" (p. 1593) [75].

Saxman proposed "By necessity patients would have to be informed generically that they will be assigned to one arm or the other and that the chances will vary over the course of the trial. They would also need to be informed of the reason for this imbalance, the fact that it will change irregularly based on accumulating data throughout the duration of the trial, and that even though sufficient information had become available to alter the randomization scheme they could be randomized to what the computer now has calculated as more likely to be the inferior therapy" (p.64) [70].

To conclude, despite being challenging, the researchers and physicians must show transparency and make maximum efforts to take an ethically sound informed consent.

Justice

Justice is an important ethical implication of the PT design because with time, the randomization allocation is changed and more patients are recruited in the treatment arm that is performing better [70]. Therefore, the patients enrolled later in the study will receive superior treatment and improved outcomes compared to those enrolled earlier. This inequality is unique in outcome-based adaptive designs and this may provoke a feeling of injustice in the participants recruited at the beginning of trial [70].

Nevertheless, ethical norms of justice can be preserved if consequences of RAR on treatment allocation are explicitly explained to the participants in advance and if they are given the right to withdraw any time from the trial if they feel injustice [70].

Confidentiality/ Integrity

Platform Trials with frequent interim analyses will obtain information about the effectiveness of an intervention but confidentiality (blinding of stakeholders) to the interim analysis results is very important to maintain the integrity of the study [76]. Blinding of

stakeholders prevents ad hoc decisions being made regarding eliminating a trial arm [76]. However, when recruitment to a trial arm must be stopped, it becomes hard to hide information from stakeholders because this information has to be updated in the patient information form [77].

Therefore, to maintain the confidentiality of results and integrity of the trial, the selection of a highly-trained independent Data Management Committee (DMC) and the Trial Steering Committee (TSC) is recommended [78]. DMC's role has been defined in the journal of clinical trials "…as an advisor to the research sponsor on whether to continue, modify or terminate a trial based on periodic assessment of trial data" (p. 342) [78]. Whereas, TSC, oversee the trial and makes ad hoc decisions suggested by the DMC [76, 78, 79]. DMC and TSC are important for clinical trials. It is important to have a clear agreement between the two committees throughout the trail.

To maintain integrity and confidentiality, the TSC should be kept blind to the interim results [78, 79]. A clear agreement regarding the trial design and adaptations should be signed between DMC and TSC at the beginning of the trial [77]. Planned modifications should be done only by DMC without further involvement of TSC [76]. Note that the permission of TSC is not necessary for doing planned modifications because these are considered the part of the initial agreement [76]. However, if unforeseen modifications are required for safety and futility reasons, then per adaptation rules, the TSC should be involved to make further decisions about the continuity of the trial conduct [45].

In short, maintaining confidentiality to interim results (by blinding the investigators and all other stakeholders to the interim results) is of paramount importance. For the sake of

integrity, it is advisable to make maximum efforts to avoid disclosure of sensitive information to the research team [77].

Validity

"Validity" refers to the selection of an applicable methodology to properly answer a search question and suitably estimate the effect parameter of interest and also the associated confidence interval and p-value [76]. According to Laage et al.: "Preserving trial validity is an ethical consideration equally important as the health and survival of the participants and one requiring investigator education" (p.196) [80].

Although, adaptive trials swiftly meet the regulatory criteria for safety and efficacy [80]. Nonetheless, both opponents and proponents of RAR acknowledge that outcome-based allocation strategy may induce confounding and selection biases that threaten the internal validity of a study [81]. To elaborate, the trial design is dynamic, both the population under study and the care/treatment standards can change over time [81]. The addition of new trial arms results in modification of trial protocol and changes in the characteristics of the population over time [81]. Henceforth, these factors can potentially introduce a bias in the treatment comparison with changes in the allocation ratio [81].

Treatments under evaluation and study protocols are subjected to modifications during the conduct of the trial which may introduce confounding and other biases[§](selection bias, investigators bias/ operational bias), etc. [82]. Patients who enter the trial earlier are at more risk of being exposed to a less efficacious intervention as compared to those who enrolled at the end [81]. Besides, patients who enter later in the trial might be different from those enrolled earlier [82]**.** Furthermore, outcome-based allocation leads to imbalanced randomization of the patients

[§] Biases are explained in the following section.

to a superior treatment arm which amplifies the imbalanced randomization of important factors (variability); resulting in confounding and erroneous conclusions [81]. However, trials where the outcome is observed at a short period, stability and fewer threats to validity are anticipated [81].

The ethical balance between sample size and resource allocation

A fundamental ethical principle in conducting trials is to minimize the overall costs and reduce the research burden [81]. Every health care system demands efficient strategies to answer a research question while preserving the resources to be directed to resolve other uncertainties [81].

Recruiting a large sample is advantageous in outcome-based trials given that a large number of patients receive superior treatment [83]. However, this is only true if the alternate treatment proves superior to standard treatment (large effect size). Otherwise, the economic burden for bearing more patients in a trial cannot be neglected [81]. In such a situation, more resources are required to accommodate a large sample size and to answer a single research question in an adaptive trial as compared to 1:1 randomization [81]. Hence increasing the research burden may bring up ethical challenges [81].

In short, important research enterprises, including FDA, consider adaptive trials as complex and expansive and they are ambivalent towards the efficiency of PT designs [45].

Conflict in design and conduct of the trial

According to Thall "The main difficulty in the design and conduct of a clinical trial is that, ideally, it must provide the patients in the trial with the best available treatments while also generating data that will provide a valid basis for making inferences aimed at developing improved therapies. These two goals, each ethically motivated, are often in conflict" (p.432). [69].

The main idea of treating the patient is a clinical trial is to generate the data which contradicts with the clinical practice, where first and foremost thing is to treat the patients. However, certain limitations imposed by the clinical trial protocol/ mandate may not be ethically acceptable to participating physician which may limit them from enrolling their patients in the trial [69]. This withholding of eligible patients is considered unethical by some clinicians and statisticians because it compromises the data collection (as fewer patients are enrolled) and deprives the patients of receiving a promising investigation [69].

Ideally, patients enrolled in PTs are likely to receive the best available treatment if they remain in the trial long enough. [69]. However, practically, these ethical goals are in conflict [69]. Also, the practical utility of the statistical data obtained during the trial, for clinical treatment of the future patients, is limited [69]. Because there are fewer data points at the start of a study so any measure of the outcome is subject to uncertainty. Hence, the ethical conflict between the design and conduct of the trial is not negligible.

To summarize, the medical trial literature debates various ethical issues associated with ADs. Certain parameters like loss of equipoise, injustice, complexities with informed consent are inherent to the adaptive designs (including platform trial) [70]. However, it can be concluded that "Equipoise", although is important for the ethical conduct of the trial, is not considered mandatory [68]. As long as the research is conducted in the light of the declaration on Helsinki, and rules of beneficence and prima facia are not ignored we may compromise the equipoise.

The complexities associated with informed consent threatens achieving the principle of autonomy (to be able to make uncoerced and informed decisions) [70]. However, physicians and researchers should make efforts to keep the consent process very transparent so that the participants can understand the risks and benefits of participating in the study.

Similarly, justice and personal perceptions may influence the operational feasibility of the research and delay the completion of a trial [70]. This inherent injustice or unfairness may negatively influence the professional relationship between stakeholders and the research enterprise [70].

2.3.2 Biases

The following sections aim at elaborating important types of biases that can occur in the context of trials:

Operational Bias

Inconsistencies across various stages of a trial may result in so-called operational bias [49]. For instance, adding of a new intervention arm, sometimes require modification of the inclusion criteria, likewise, the addition of additional test/biomarker changes accrual rate and the study population over time, which may provoke bias alterations in how the study is operationalized.

Investigators Bias

Knowledge or just speculations of interim results are likely to alter the behavior of the stakeholders in the trial, including the patients, researchers, and investigators [84]. Therefore, protecting the confidentiality of data and interim results is of paramount importance [84]. Likewise, rigorous planning and transparency are necessary for the conduct of a trial [76, 78]. Proposed strategies are 1) funding sponsor's role in decisions making should be predefined. 2) independent and competent Data Management Committee (DMC) should be established, 3) trialists should be kept blind of interim results, and 4) results should be appropriately reported [78].

Biased effect size estimation

Planned modifications and adaptations during the conduct of the trials may lead to statistical biases, biased estimation of treatment effect and erroneous conclusions [45]. Likewise, confidence intervals may not correlate with the true treatment effect [45]. Additionally, drastic changes in inclusion and exclusion criteria or the addition of a high volume of new study sites may affect the response rate and may lead to erroneous conclusions [62]. Critics of AD argue that the dynamic patient populations may cause a bias in parameter estimates [62]. Proponents of AR [85], conducted simulation studies and found that the effect of the bias is minimal.

Sample size and power conflict

Although PTs have received attention by the clinical trialist, and pharmaceutical companies due to smaller sample size requirements; sometimes an increased sample size is required which may affect trial efficiency [45]. PT / MAMS drop the least efficacious arm during the conduct of study which can lead to uncertainty about total sample size requirements [46]. Likewise, adding trial arms raises funding and logistic issues [45, 46] which could be eliminated by keeping a fixed number of trial arms at each stage of the study [46].

In PT/MAMS/RAR designs, there can be a higher risk of type 1 error and power loss as compared to conventional RCTs [49]. Likewise, with increasing sample size, a high proportion of participants are allocated into the superior arm, and an inferior arm [86, 87]. Therefore, in trials where multiple superior treatments exist, power estimation becomes an additional challenge [88]. Recruiting a diverse and adequate sample of participants has been proposed to optimize the power of a study[88].

Allocation Concealment / Selection Bias

Allocation concealment prevents selection and confounding biases [89]. In AD (e.g. PT), imbalanced randomization compromises the allocation concealment potentially leading to a

selection bias. Investigators can minimize the potential selection bias by concealing the allocation strategy from the clinicians and researcher. Still, sometimes the nature of the intervention might make the concealment impossible [82].

Time trends

In PT, balanced random allocation at the start of the trial is gradually altered to an unequal randomization frequency, favoring greater allocation into the best performing trial arm [90]. According to L. Korn & B. Freidlin, "Any time trends in the prognostic characteristics of the patient population enrolling in the trial will bias the results of the trial. For instance, if in the earlier part of the trial, patients are randomly assigned equally to the experimental and control treatment arms, but are randomly assigned 9:1 in favor of the experimental arm later in the trial, then an improving prognostic pool of patients being randomly assigned in the trial will translate into a bias in favor of the experimental arm" (p.3).[90].

In short, time trends result in erroneous conclusions such as an inflated effect estimate [90]. This concealed bias (time trends) makes RAR less suitable for chronic long-term conditions [91]. Although Simon has proposed special statistical methods, which can amend this issue at stake of compromising the power of the study [86, 92] however; these details are beyond the scope of my thesis.

2.3.3 Temporal drift

Temporal drift defined as "changes in patient or treatment characteristics over time" (p.1090). [82] This phenomenon is frequently encountered in PT due to the flexibility of adding and eliminating different treatments and participant recruitment over the conduct of the trial [82]. This issue can be addressed by selecting similar patients e.g. in terms of severity of illness across various phases of the study [82]. Practically, it is important to collect baseline characteristics of

participants recruited at different sites and stages of a study. Information about the baseline characteristics can explore the heterogeneity of the recruited population across different stages/phases of the trial. Nevertheless, sometimes it is difficult to differentiate between heterogeneity from population drift [76].

2.3.4 Generalizability

In the PT design, populations recruited at the end of the trial may be substantially different from those recruited at the beginning of the trial [45]. This may challenge the interpretation of results and consequently limit the generalizability [45]. However, proponents of AD support the notion that if comparative effectiveness studies are conducted in a realistic setting using real data, the results can be generalizable because more variability is expected [62]. However, this may require a large sample size and increased costs [62].

2.3.5 Logistical issues

Adaptive randomization may have a restricted utility if the primary endpoint/outcome is evident over an extended period, for example, if the effect of the drug/ intervention under study is observable only after a long period [45]. PT may take more than expected time to complete because sometimes there is no "well-defined end of the trial" (p. 220) [49]. There can be extended delays between start (patient enrollment) and end (outcome observation), particularly, if the outcome is the manifestation of a disease or death [49]. PT, like other ADs, may require added time for making decisions based on interim results. Statisticians may ignore this important logistical issue; instead, based on mathematical convenience they may assume that a trial is efficient as the interim results can indicate the outcome instantly [49].

2.3.6 Statistical issues

Korn &Edward argued that adaptive randomization may be statistical inefficient because of having unequal randomization in the treatment arms [53]. They further elaborated that a trial with an equal number of patients in the treatment and control arm has a more precise estimate than the trial with 90 patients in the treatment arm and 10 patients in the control arm [53]. Hence, to get the adequate and same information effect about the effect of intervention the RAR may require a longer time, and therefore, delay the development of effective intervention and may also expose more patients to an ineffective treatment [53].

Assessment of a trial requires treatment effect estimates, computation of confidence intervals(CIs) or estimate accuracy (mean standard error) and p-values, [76]. The US Food and Drug Authority elaborated on the reasons why the sample mean, traditionally used for estimating treatment effects, is affected by the adaptations. Hence, it overestimates the true treatment effect on the population under study, which is true for both the primary and secondary study endpoints [93]. Likewise, the CI may not accurately cover the true treatment effect with the nominal probability assigned [93].

2.3.7 Risk of type 1 error inflation

Platform trials, like other adaptive designs that employ interim analyses of the primary outcome, are at risk for type-I error [45]. The planned trial modifications based on interim results and population drift can markedly inflate type-I error [49, 94]. For instance, as per adaptations, the inclusion and exclusion criteria are changed and the population recruited at the end of trial may be different from the source population recruited at the beginning of trial [95, 96]. This population drift may not only affect parameter estimates but also inflate the type-I / type-II error rates. Suggestions are made to adjust the estimates of parameters with time [95] using linear

modeling [96]. Likewise, Villar, Bowden,and Wason [97] accessed the risk of type 1 error inflation in RAR by generating various scenarios They conclude that various adjustments can significantly alleviate the type 1 error (secondary to population) in RAR, but it is important to correctly specify the trend and accurately measure the responsible covariates [97]. However, for MAMS designs such adjustments are less effective in preserving the power of study; nonetheless, simple protection of the "allocation to the control arm" is more useful [97].

In the PT design, population drift, long term or delayed outcomes, etc. may complicate the prediction of the actual type 1 error rate [98]. Likewise, exactly defining the type 1 error rate in a multi-arm trial is challenging [98]. Lipsky elaborated on this issue with an example: if two tests (of two drugs) in two independent trials had a 2.5 % error rate. Now, when these drugs are combined in a single trial, should the new trial has combined a 2.5% error rate or should it be 2.5 % for each of the arm [98]. Additionally, the required simulations in PT designs may complicate the control of Type I error rate [98]. Therefore, FDA recommends that adaptive design proposals should thoroughly address the omnipresent type 1 error issue and how it will be addressed [99].

2.4 Knowledge gap

Advanced adaptive trial designs including PT/MAMS/AR are, despite being known for more than 25 years, still underutilized in research [100]. To my knowledge, the PT/MAMS/RAR designs have been used for treatment evaluation in oncology (e.g. breast cancer [101], lung cancer [102], brain cancer [103], melanoma [104], glioblastoma [105]), infectious diseases (pandemic influenza and community-acquired pneumonia [103], Ebola [51], tuberculosis [88] etc.), neurology (status epileptics [106]), dermatology [107], intensive care [108], and cardiovascular diseases [109] etc. In addition, some of the trials were multi-center trials like FOCUS4 [56], STAMPEDE [57], SPY2 [58], and BATTLE [59].

PT/MAMS/RAR designs have barely been used in the context of HIV cure research [100]. Probably, the complexity of the trial design and practical challenges for implementing such trials demand more careful planning in the context of HIV research [100].

Thus far, these designs have been used to test the efficacy of antiretroviral therapy ART in neurological disorder [110], the EHVA T01 trial for HIV vaccine testing [100], evaluating rosuvastatin in pulmonary disease in PLHIV [111] and HIV self-testing interventions [112]; however, *have not been used in the context of developing PROM based mHealth interventions* (Appendix A, tables 3&4).

The I-Score (CTN 283) study aims to develop a PROM based mHealth intervention and evaluate different evolving features of the health app by using a PT design. For the proper planning of a PT and to ensure external validity (generalizability) of the trial outcomes, it is important to characterize the study and target population.

To understand the key characteristics of PLHIV that had already been included in HIV cure related platform clinical trials I explored and compared the demographic and clinical characteristics of the PLHIV in previous HIV- PTs [111, 113]. Both trials identified from the literature [111, 113] included young (44-50) year old, males (70-85%) > females (15-30%), Black (45-55%) and Hispanics (4.5-15 %) participants. However, the eligibility criteria and context for both trials were different and no information could be retrieved regarding the ideal characteristics of PLHIV that are suitable for participation in a PT (Appendix A, Fig 1 & 2). Similarly, neither of the trials highlighted the practical challenges that may be particular in the context of an HIV care-related trial.

Since the I-Score investigators aim to conduct the trial on PLHIV in Montreal, my research revealed that previously a cross-sectional assessment of Cohort de Montreal data (CM)

for the year 2015 was done to determine the prevalence of PLHIV on ART and PLHIV with suppressed HIV viral loads in this urban context [114]. Some of the demographic and clinical characteristics of PLHIV determined in 2015 from the CM data were: total N= 6, 364, (85%) male, (15%) female, (51 years) average age, (67%) MSM, (11%) IDU, (11%) from endemic countries, (9 %) heterosexual [114]. Nonetheless, this information was not updated and was not sufficient to address my research question that is to identify the characteristics of the potential target population[**] and compare it with the characteristics with the I-Score study population[††].

Here, I proposed to conduct a descriptive cross-sectional analysis of updated (2016) CM data to characterize the study population and the target population for a future PT in the context of the I- score study. *I aimed to identify commonalities and differences in PLHIV followed up at the CUSM with other HIV clinics included in CM*. This information would later be used to define the eligibility criteria along with the proper set-up and planning for the future platform trial in Montreal. Ultimately, these findings would contribute to understanding the (limits to) generalizability of the PT's findings. Specifically, I addressed the following research question:

Are there systematic differences in the sociodemographic and clinical characteristics of PLHIV with ART followed up at CUSM compared to PLHIV in the Cohorte de Montréal (e.g. CHUM, CMQL and L'Actuel) in the year 2016?

[**] PLHIV on ART are followed up at the CUSM
[††] (PLHIV on ART are followed up at other clinics included in CM)

3 Methodology for quantitative analysis (phase 2)

I conducted a secondary analysis of Cohort de Montréal (CM)[‡‡] data. The data was collected after the city of Montreal signed the Paris Declaration on "Fast-Track Cities" on Dec 1, 2017, and became one of the Fast Track Cities fighting against HIV [114, 115]. Fast-Track Cities commit "to build upon, strengthen, and leverage existing HIV-specific and-related programs and resources to attain the 90-90-90 targets [116] (90% of people living with HIV (PLHIV) diagnosed, 90% of diagnosed PLHIV on antiretroviral (ARVs), and 90% of PLHIV on antiretroviral and virally suppressed); increase utilization of combination HIV prevention services; try to reduce to zero the negative impact of stigma and discrimination; and establish a common, web-based platform to allow for real-time monitoring of progress." (p.1) [114, 115] CM data was collected to access the quality of care provided to PLHIV that were already registered with HIV care providers in Montreal, Quebec, Canada [114].

Study Design

I conducted **a descriptive cross-sectional study of Cohort de Montreal data** collected from Jan 1, 2016, to Dec 31, 2016. The descriptive cross-sectional design is inexpensive, fast, and can be useful for the planning of future studies [117]. This type of analysis is useful for assessing the prevalence of diseases and risk factors from the clinic and hospital data [118]. It allows obtaining an overview of what is happening in the population without manipulating variables of interest [119].

[‡‡] CM includes all PLHIV in the Montreal who receive care at one of four major HIV-specialist care centers.
These clinics are 1) Clinique médicale l'Actuel, 2) Clinique medicale Quartier Latin (CMQL), 3) Unité hospitalière de recherche, d'enseignement et de soins sur le sida (UHRESS) at the Centre hospitalier de l'Université de Montréal (CHUM); and 4) Chronic Viral Illness Service (CVIS) at the McGill University Health Centre (MUHC)) [2]. I analyzed the latest available (2016) one-year EMR data from four major clinics included in CM.

According to Gaille, the cross-sectional design "…provides better descriptions for the data points which occur, making it possible for the information to lead toward possible solutions that may not have been previously considered" [120]. This design may bring up an important hypothesis for future research and collected data may be useful for multiple secondary analysis [120].

I chose the descriptive cross-sectional design, as I aimed to determine the outcomes (prevalence and the differences in the characteristics of the study populations vs target population), without changing the exposures or outcome. Besides, I was able to assess relevant associations among different variables before the purpose of hypothesis generation and to be confirmed in future research. In addition, the descriptive study enabled me to identify the deficiencies in the data collection methods used in HIV clinics in Montreal, which potentially contributes to quality improvement in the CM data and ensure the validity of future research, conducted using CM data.

Despite the various advantages, cross-sectional studies are not without limitations. Due to the collection of data only at a single time point, the results may not be generalizable to the general population unless data is being collected from the entire population studied, including vulnerable groups [120]. In other words, the results of a cross-sectional study may not apply to the source population; therefore, before generalization of the results, the size and composition of the study population must be accessed [121].

The cross-sectional design does also not allow for measurement of incidence rates. Researchers can explore the association between variables; however, they cannot establish a causal relationship between variables as the temporal order of measurements is not always clear

and confounding variables (necessarily preceding the measured exposure and outcome variables), are typically not available.

Generally, when using data for secondary analysis the bias of the investigators and researchers may influence the quality of the collected data and the secondary studies may not be aware of the bias. Also, the secondary analysis does not allow control over the data collection strategy [120]. Therefore, details about the data collection method, purpose, and choices made needs to be hand over during the transfer of data for the secondary analysis [120]. The population included in the data must be large and appropriately defined particularly when exposure and outcome of interest are rare; otherwise, it will affect the accuracy of the results [120]. Similarly, there might be minor differences among the participants included in the study; however, these minor differences may significantly influence the results of the study [119]. Moreover, the cross-sectional design may insufficiently capture health events of short duration (e.g. rapid recovery). Hence, only the association between the exposure and long-term event is observable. However, such an association is not necessarily representative of the entire population and individuals experiencing the same exposures and outcomes [121].

Since I am conducting a secondary analysis on an established dataset, I anticipate that some of the inherent biases (for instance during the data collection or transfer) could have affected my study's results. The potential biases will be discussed in detail in the limitation section of this master's thesis.

Study population

Inclusion/exclusion criteria

I included adult PLHIV from the CM who were on ART and had at least one follow-up visit in 2016.

Period studied

I included EMR data from PLHIV collected between January 1, 2016, and December 31, 2016.

Description of the Cohorte de Montreal data

The CM data was kept at the lead site: Centre De Recherche Du CHUM. Dr. Marina Klein is the Principal Investigator of Cohorte de Montréal at MUHC. The sponsor principal investigator: Dr. Jean-Guy Baril permitted me to access the CM data. Additionally, a data transfer agreement signed between Dr. Bertrand Lebouché (RI-MUHC) and Dr. Jean-Guy Baril at the CHUM. This study obtained ethics approval from the MUHC Research Ethics Board (REB).

Ethical Considerations

This study was conducted in accordance with the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (2014), as well as in respect of the requirements set out in the applicable standard operation procedures of the Research Institute of the McGill University Health Centre Research Institute and of the McGill University Health Centre (MUHC) Research Ethics Board (REB # 2020-5716) (Appendix B). The board reviewed and approved this study and was responsible for monitoring all the participating institutions in the health and social services network in Québec.

Confidentiality:

I analyzed a subset of prespecified variables that were relevant to my thesis. All the collected data was secured and pseudonymized in order to protect patient confidentiality. All personally identifiable information removed to ensure that it was not possible to identify participants.

Informed consent:

In lieu of individual informed consent of participants, authorization to access patient charts was obtained from the Director of Professional Services (DPS) of FRSQ Registre main study La cohorte de Montréal – register de données communes VIH (MUHC REB #2011-942).

Operational Definitions of Variables

The CM dataset included the following variables

3.1.1 Pseudonymous patient identifier

In the CM dataset, there were two identifiers:

BdCid_2018

This number is a unique identifier for each of the patients included in the CM.

Site_Pid

This identifier generated at the time of encryption. This ID starts with the name of the clinic visited by the patient (e.g. CUSMxxx/ CHUMxxx / CMQLxxx / ACTxxxx for CUSM, CHUM, CMQL and L'Actuel respectively), since, some of the patients visited more than one clinic included in the CM, they may have multiple Site-Pids; thus, I only considered BdCid_2018 as a unique identifier of the participant.

3.1.2 Gender

I only considered male and female patient records in the analysis. There was only one transgender in the cohort that I intentionally removed from the dataset. Being the only transgender patient followed up at a specific clinic could easily reveal who the patient was. Therefore, I was ethically obliged to ensure the anonymity of the participant and remove the entry from the study database.

3.1.3 Ethnicity

This categorical variable in the dataset describes the ethnicity of the patients at ten levels: White, Black/ Black (Africa)/ Black (Caribbean), Asian, Hispanic, Latin, Aboriginal, others, and unknown. I reported all the Black population as one entity, irrespective of their origin from Africa or the Caribbean.

3.1.4 Age

This variable was not in the original dataset; however, using the month and years included in the CM data, I calculated the age of the participants.

3.1.5 Type of infection (HIV mono, HIV-HCV co-infected)

This categorical variable provides information on the patient's infection status:

- HIV-mono infection: patients only diagnosed with HIV
- HIV-HCV co-infection: patients that have an HIV infection and presenting with the antibody for HCV (anti-HCV), a marker for HCV infection.

In the original dataset, there were three levels for HCV antibody status: positive, negative or not done yet.

3.1.6 Visits data

Visits data was a time-series data that includes patient identifiers (BdCid_2018 & Site-Pid), date (d/m/y), and the clinical site of follow-up visits.

3.1.7 CD4 counts (numeric)

CD4 counts is a blood test that quantifies CD4 T lymphocytes (CD4 cells) in the blood sample. This test is a significantly important indicator of the immune system performance and therefore, the strongest predictor of HIV disease progression. The CD4 count is additionally accustomed to monitor a person's response to ART [122].

In healthy adults not infected with HIV, the normal range of CD4 in a human blood test is about 500-1,500 cells/microliters [123]. CD4 cell count begins rising once the HIV infection becomes under control with effective ART. In contrast, CD4 cell counts start declining with the progression of HIV infection. CD4 counts <200 cells/microliter indicate a severe deficiency of the immune system which ultimately results in series of opportunistic infections that all together are known as Acquired Immune Deficiency Syndrome (AIDS) [123].

In the CM dataset, this variable had related information documented including 1) date of tests on which the blood sample was taken 2) the percentage of CD4, and 3) the ratio of CD4/CD8 in percentage. The data was time series.

3.1.8 Viral loads (numeric)

Viral load test

A blood test that quantifies the amount of HIV as the number of copies of HIV RNA per milliliter of blood. The viral load test is an important diagnostic test for acute HIV infection, it guides treatment decisions, and monitors response to ART [124].

Undetectable Viral Load/ Viral suppression:

HIV viral load is considered undetectable when HIV viral load (HIV RNA) in the patient's blood is below the detection limit of the viral load test. If the viral load remains undetectable for six months after the first undetectable test result it is considered "durably undetectable" [125]. Effective ART reduces the viral load to undetectable that means the virus has become dormant in the immune system (suppression) [125, 126]. However, if ART is stopped the virus will likely become active and viral loads be detectable [126].

The virus load is considered undetectable if there are <50 copies of the retrovirus per ml of plasma. The undetectable limits also depend on the sensitivity of the test used [127]. Some

tests can only detect levels up to <50 copies/ml and others can go as low as 20 copies of HIV per ml of blood [127].

In the CM dataset, this numeric variable provides the following information: 1) date of blood test (d/m/y), 2) site (Site-Pid) and 3) type of blood test (BDNA-US, BDNA, PCR, PCR-US, NASBA, Abott RealTime HIV-1 PCR, and others) & 4) Viral loads.

3.1.9 Injection drug use status (dichotomous)

Injection drug use is defined as the drugs that are injected with any needle or syringe into the body (muscle, vein or beneath the skin). [128, 129]. The dataset being used has many levels (e.g. do not use drugs (who never used drugs), ex-consumer (who consumed drugs in the past but do not use now), current consumers (who are currently consuming drugs), unknown (their IDU status is unknown).

3.1.10 Risk factors for HIV acquisition

In the CM dataset the risk factors for HIV acquisition were:

Natives of countries

Patients born in countries where HIV is endemic (e.g. black people of African or Caribbean descent) [130].

Transfusion of blood

Transfusion of blood is referred to the transfer of blood or any component of blood products (e.g., plasma, platelets, etc.) from the donor to the recipient [131].

Vertical transmission:

Vertical transmission is the passage of disease-causing organisms from mother to baby at a time either before or immediately after the birth of the baby. Transmission can occur through the placenta, breast milk, or direct contact, during or after birth [132].

Homosexuality

Homosexuality is referred to as sexual attraction, desire or orientation towards a person of the alike sex [133].

Heterosexuality

Heterosexuality is referred to sexual attraction, desire or orientation towards a person of the opposite sex [134]

Bi-sexuality

A person who engages in both heterosexual and homosexual sexual relations [135].

3.1.11 Social-economic status (categorical)

The socioeconomic status of study individuals was measured by social assistance benefits (yes/no/unknown). Educational status (categorical): This is a categorical variable encompassing different levels of education: 1) who finished university, 2) Cégep, 3) completed secondary and 4) unfinished secondary (the codebook does not provide any information either it was secondary in progress or drop out).

Data Preparation

The pseudonymous data (from all four clinics included in CM) was received as one Microsoft Excel file including separate sheets for demographics, number of visits, risk factors, anti-HCV antibody, CD4 counts (absolute and percentage) and viral loads. The original data and codebook were in French but due to my linguistic competence in the French language, I decoded and translated the data into English. I added additional variables of "age" (which was calculated from the date of birth) and "site" (by using the variable site-pid).

In addition, there were inconsistencies in the units of measurement for the CD4 counts and viral loads data across sites (particularly, L'Actuel) which required a lot of work for data

cleaning. There were 6,204 entries for the demographic data. I deleted transgender data (n=1) to avoid the risk of identification. Later on, I found that 1,555 of the unique IDs (Bd-Cid2018 had repeated measurements across the sites. I grouped the data by unique IDs (Bd-Cid2018) and selected the first value from the remaining columns in Microsoft Access. Then I created new data sets for the demographic, socio-economic and HCV antibody data that were without repeated measurements. I used these new datasets for the analysis (Appendix B datasets used in analysis). On the other hand, visits, CD4 counts and viral loads data were time series data with repeated measurement, so, to preserve the richness of the data; I retained the repeated measurements. The statistical analyses were performed using the software packages R [136] and R-Studio (Version 1.0.153 – © 2009-2017 RStudio, Inc.).

I visualized the missing data and found a total of 26.4% of the demographic data was completely missing. Individually, (29%) of country of birth, (25.71%) of IDU (5%) of risk factor 1, (63.4%) of the risk factor 2, (93 %) for the risk factor 3, and (99.95%) of the data for risk factor 4 were missing. Likewise, in total (12.75%) of the socioeconomic data and (31%) of education data and (0.1%) of the HCV data was missing (Appendix A, Figs 3-5). The and socio-demographic data for the CUSM were completely missing which I handled by using multiple imputations, the "mice" package in R [137]. (Cleaned and sorted datasets and R scripts attached in Appendix B).

Data analysis

Descriptive statistics including frequency distributions, means, standard deviations, and quartiles were computed to describe the sociodemographic and clinical characteristics of the study population (CM). The variables of interest across various sites were plotted as two-dimensional bar plots (gender, country of birth,) two-dimensional stacked percent bar plots

(ethnicities, injection drug use, HCV antibody status) and box plots (age and number of visit) and dot plot (for prevalence of risk factors combination for HIV acquisition). Laboratory variables including CD4 counts and viral loads were analyzed using bivariate density plots. (R script in Appendix B).

4 Results

Descriptive analysis of Cohorte de Montreal data.

Sociodemographic Characteristics of PLHIV in Cohorte de Montreal.

Between January and December 2016, there were 4,881 individuals in the CM dataset. This population was organized into four sub cohorts: CUSM (n=1,261), CHUM (n=464), CMQL (n=952) and L'Actuel (N=2,204).

Overall, there was a greater proportion of males (85%) and the mean age in the study population was 55 years±11. Among the sub cohorts, the CUSM (67%) had the smallest proportion of males, and CMQL (96%) had the largest. The mean age of the PLHIV at the CUSM (54±12 years) was comparable to the **age** distributions of the three other sub cohorts, differences in mean age were ≤5 years (Figure 4.1, Table 4.1).

Data on **socioeconomic status** was missing for CUSM. To reduce the likelihood of selection bias induced by missing data, multiple imputation methods were applied to impute missing values in demographic variables based on observed covariate realizations. Most of the PLHIV (49%) in the CM were not receiving any social welfare aid. Among the sub cohorts' most of the PLHIV were on social aid at CUSM (100%) which was comparable to CMQL (100%). However, the majority of the PLHIV followed up at L'Actuel (82%) and CHUM (45%) was not getting social welfare aid (Table 4.1).

Table 4.1: **Socio-demographic characteristics of PLHIV included in CM (2016).**

Variables:	CUSM n=1,261	CHUM n=464	CMQL n = 952	L'Actuel n =2,204	Total N= 4,881
Age (years)					
Mean (SD) [Min-Max]	**54 (12) [23-89]**	59 (10) [28-89]	56 (11) [2-88]	54 (11) [22-87]	55 (11) [22-89]
Median [IQR]	**55 [46-62]**	58 [53-65]	57 [50-63]	55 [47-61]	56 [48-62]
Gender (%)					
Males	**845 (67)**	373 (80)	911 (96)	845 (92)	4,147 (85)
Females	**416 (33)**	91 (20)	41 (4)	416 (8)	734 (15)
Socioeconomic Status of the Patient Measured by having Social Assistance Benefit ∗(%)§§					
Yes	**332 (100)**	59 (7)	724 (100)	271 (16)	1,386 (39)
No	**0 (0)**	373 (45)	0 (0)	1,363 (82)	1,736 (49)
Unknown	**0 (0)**	392 (47)	0 (0)	30 (2)	422 (12)
Injection drug use (%)					
Current user	**1 (0)**	7 (1)	261 (27)	24 (1)	394 (6)
Ex-user	**1 (0)**	0 (0) Φ	19 (2)	266 (12)	297 (5)
Do not use drugs	**NA Φ**	51 (11)	521 (55)	305 (14)	1,096 (18)
Unknown status	**1, 259 (100)**	406 (88)	151 (16)	1,609 (73)	4,417 (71)
Prevalence of primary risk factors for HIV acquisition (%).					
Bisexual	**6 (0)**	0 (0)	4 (0)	73 (3)	83 (2)
Blood /blood products transfusion	**12 (1)**	6 (1)	1 (0)	2 (0)	21 (0)
Heterosexual	**288 (23)**	52 (11)	31 (3)	284 (13)	655 (13)
Homosexual	**341 (27)**	250 (54)	787 (83)	1,451 (66)	**2,829 (59)**
Injection Drug Use	**45 (4)**	60 (13)	31 (3)	99 (4)	235 (5)
Native of endemic country	**244 (19)**	64 (14)	16 (2)	54 (3)	378 (8)
Vertical transmission	**23 (2)**	1 (0)	1 (0)	0 (0)	25 (0)

∗ Predicted after data imputation.
§§ After the data imputation Total N= 3,544
Φ Data not available
Φ Data not available

Unknown	**302 (24)**	31 (6)	81 (9)	241 (11)	655 (13)
HCV antibody (%)Ψ					
Positive	**2 (12)**	77 (11)	76 (7)	189 (10)	344 (10)
Negative	**10 (59)**	521 (72)	925 (91)	1501 (86)	2957 (84)
Not done	**5 (29)**	122 (17)	20 (2)	63 1 (4)	210 (6)
Number of visits during 2016					
Mean [Min-Max]	**3 [1-15]**	8 [1-169]	5 [1-37]	2 [1- 15]	
Median [1QR]	**3 [2-4]**	5 [2-8]	4 [3-7]	1 [1-2]	
No visits/ missing = NA	**3416**	4394	3555	3100	

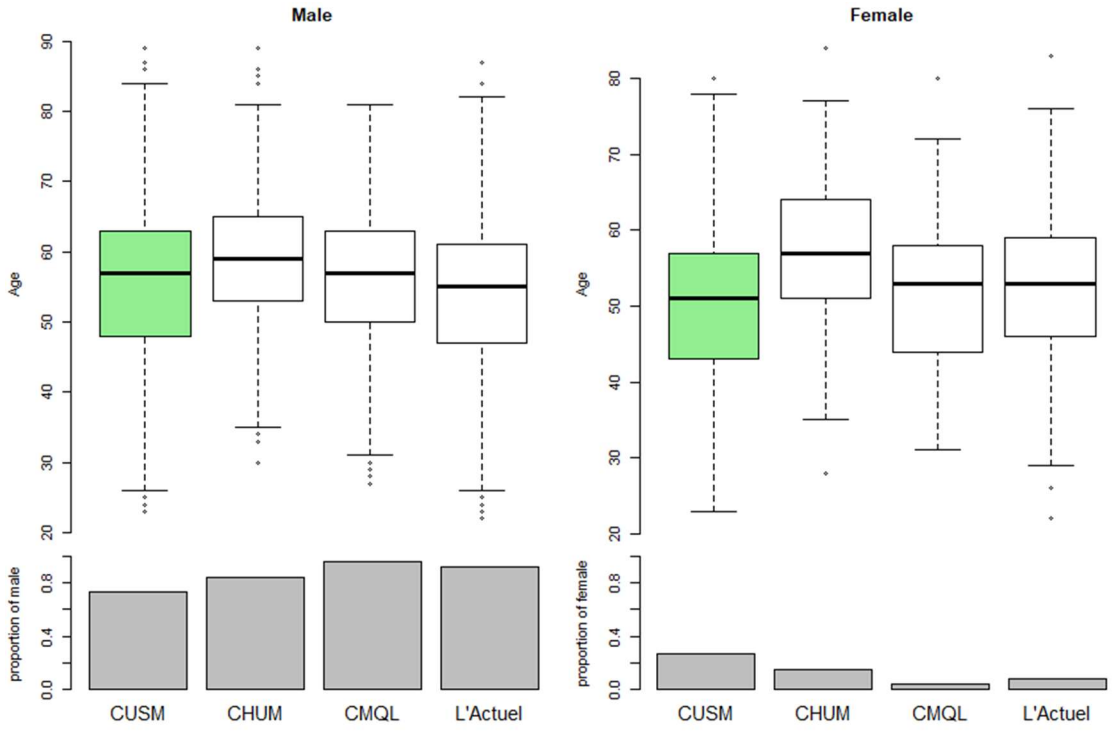

Figure 4.1: Age and Gender distribution in the CM (2016).

Ψ Total N=3,511 for the HCV data set. Sub cohort sizes were n=17 at CUSM, n= 720 at CHUM, n= 1,021 at CMQL, n= 1,753 at L'Actuel,

Ethnicities

In general, the prevalence of white ethnicity is high (39%) among the PLHIV included in CM, other ethnicities in order of frequency include; Unknown (43%), Black (13%), Hispanic (3%), Asian (1%), Aboriginal (0%). Among the sub cohorts, the prevalence of white ethnicity at CUSM (30%) was lesser than CHUM (71%), CMQL (84%) and more than L'Actuel (16%) (Table 4.2 & Figure 4.2).

Table 4.2: Ethnic characteristics of PLHIV in CM (2016).

Ethnicities N (%):	**CUSM n=1,261**	**CHUM n=464**	**CMQL n = 952**	**L'Actuel n =2,204**	**Total N= 4,881**
Aboriginal	**17 (1)**	2 (0)	1 (0)	4 (0)	24 (0)
Asian	**25 (2)**	10 (2)	11 (1)	12 (1)	58 (1)
Black	**489 (39)**	97 (21)	28 (3)	40 (2)	654 (13)
Hispanic	**89 (7)**	0 (0)	54 (5)	23 (1)	166 (3)
Latin	**0 (0)**	20 (4)	0 (0)	7 (0)	27 (1)
White	**374 (30)**	324 (71)	790 (84)	357 (16)	1,845 (39)
Unknown	**267 (21)**	11 (2)	68 (7)	1,761(80)	2,107 (43)

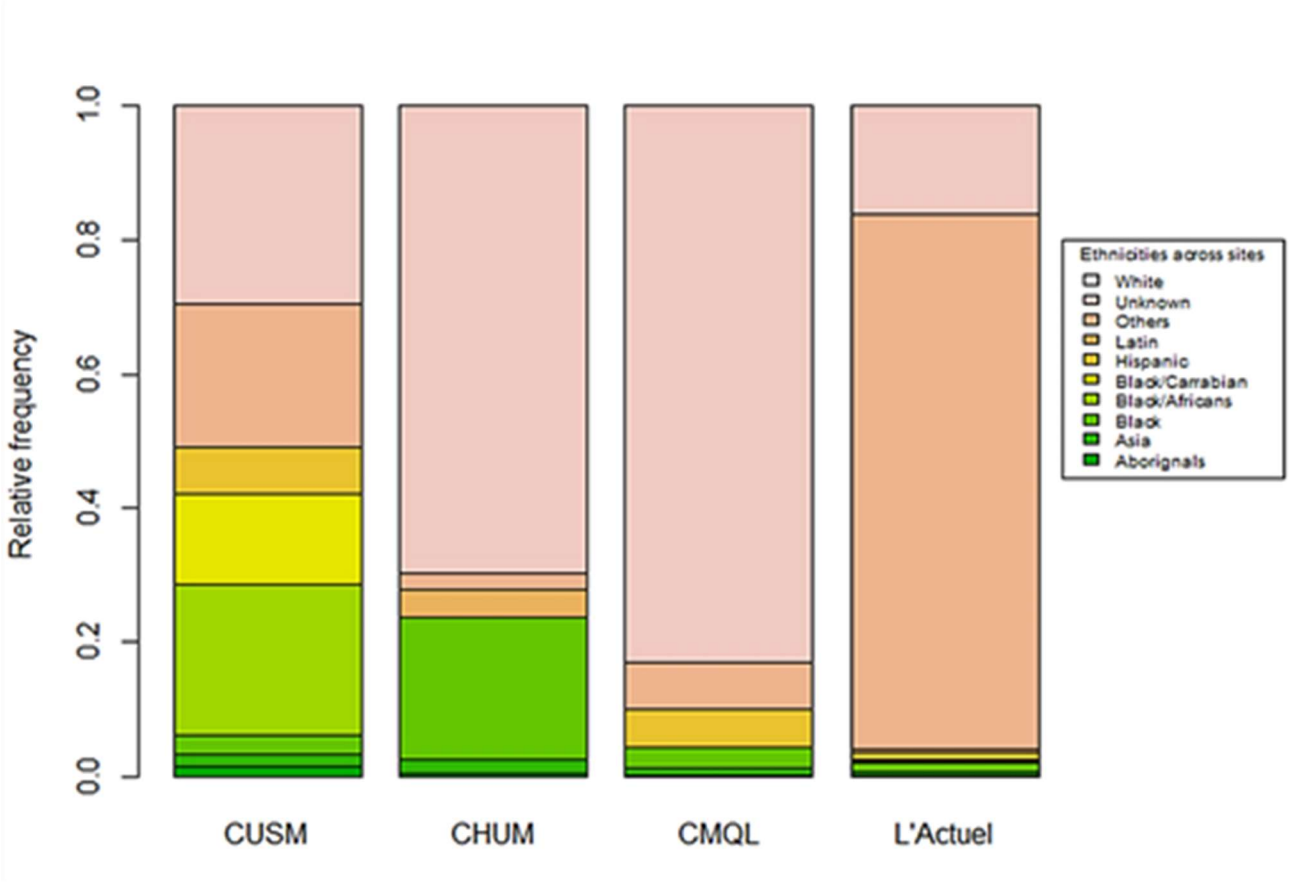

Figure 4.2: Prevalence of ethnicities among PLHIV in CM.

Injection Drug use (IDU) status

Among PLHIV in CM, IDU status was unknown for the majority of the participants (71%). CUSM have not had data for any of the group in IDU data in 2016. For the remaining three sub cohorts, IDU status for the CHUM (88%) and L'Actuel (73%) is unknown. Although CMQL has apparently a higher proportion of non-IDU (55%) and current IDU consumers (27%). However, due to insufficient data, no meaningful difference in IDU status across the sites could be detected (Table 4.1, Fig 4.3).

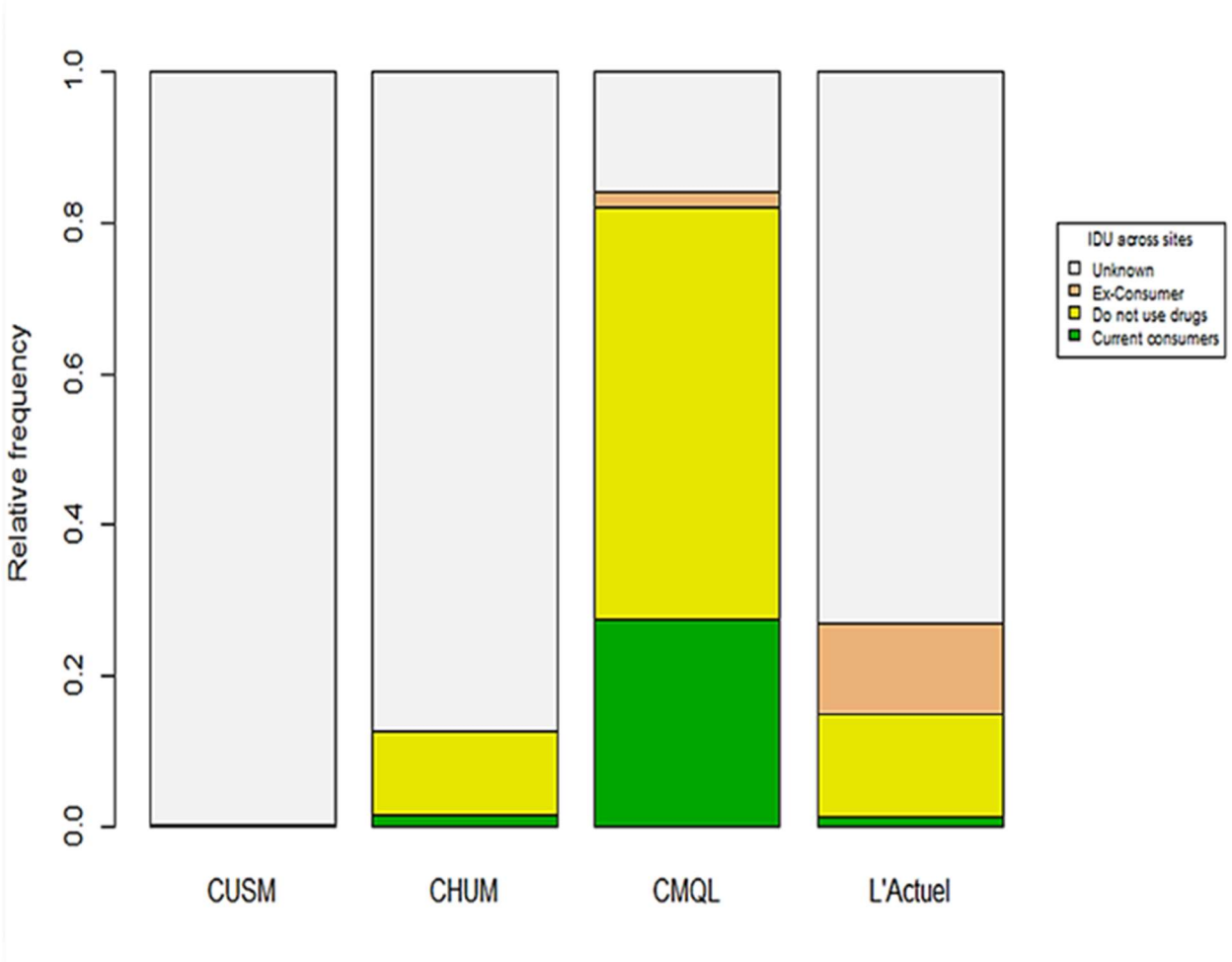

Figure 4.3: IDU among PLHIV in CM (2016).

Infection Type

Across the four-sub cohorts, the prevalence of HIV mono-infection [***](77%) was greater than HIV-HCV co-infection[≠] (10%) (Table 4.1). The prevalence of HIV-HCV co-infection was comparable between CUSM (12%) and the CMQL (10%), and CHUM (11%), but a little greater than L'Actuel (10%) (Figure 4.4). However, 27% of PLHIV at CUSM have never had test for HCV antibody status. Therefore, one cannot affirm the conclusion based on the incomplete data.

*** PLHIV whose test results were negative for anti HCV antibody.
≠ PLHIV whose test results were positive for anti HCV antibody.

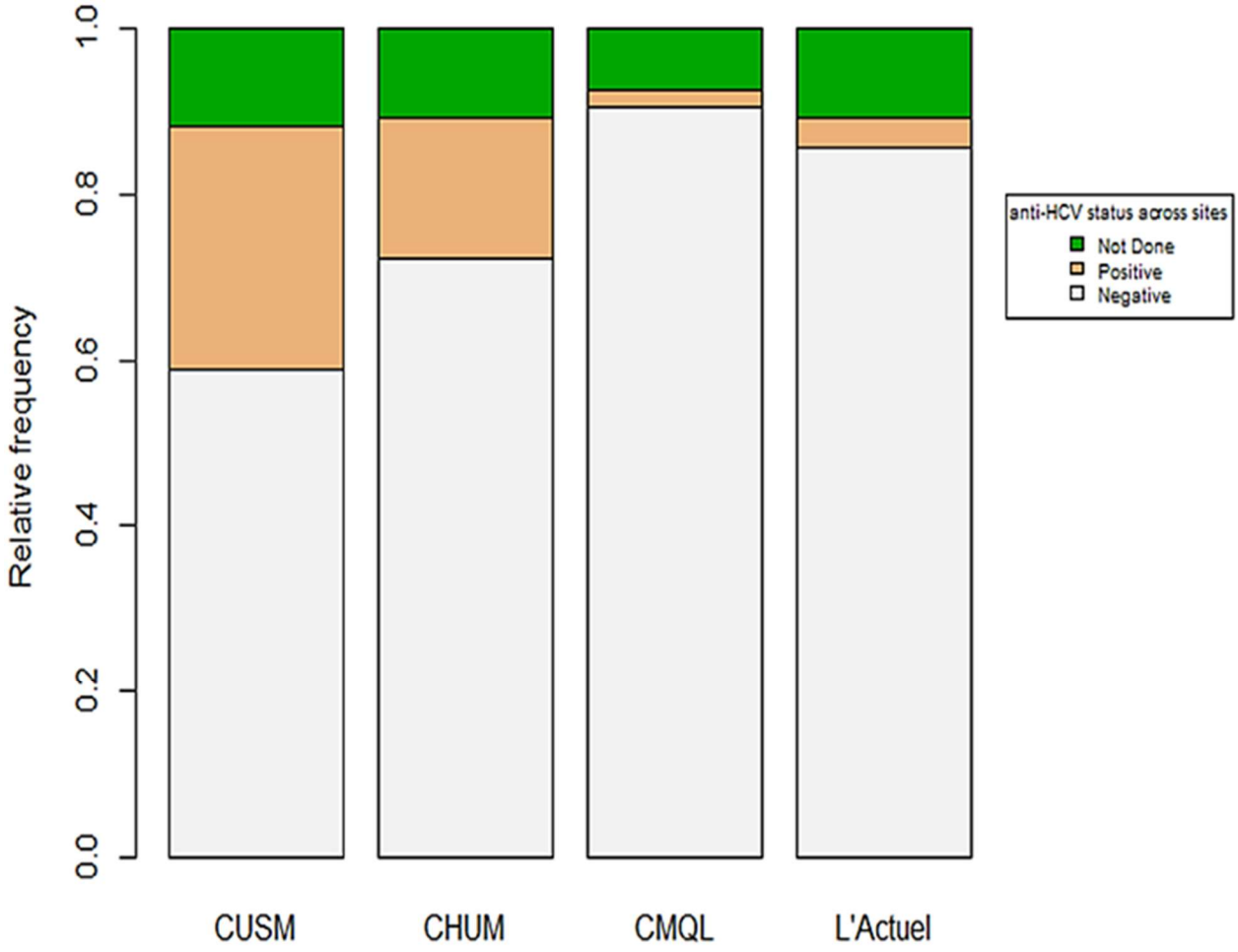

Figure 4.4: Prevalence of HIV-HCV co-infection among the PLHIV in CM.

Distribution of visits frequency across the sites

The average number of follow-up visits by the PLHIV at CUSM [three visits, IQR: 2-4] was lesser than at the CMQL [four visits; IQR: 3-7] and the CHUM [five visits; IQR: 2-8] but was greater than the L'Actuel [one visit; IQR: 1- 2]. In fact, there was an overlap in the clinical

visits across sites. Only 4,881 PLHIV in CM followed up at one clinic and 1,555 PLHIV had follow up visits at two or more sites (Table 4.1, Figure 4.5).

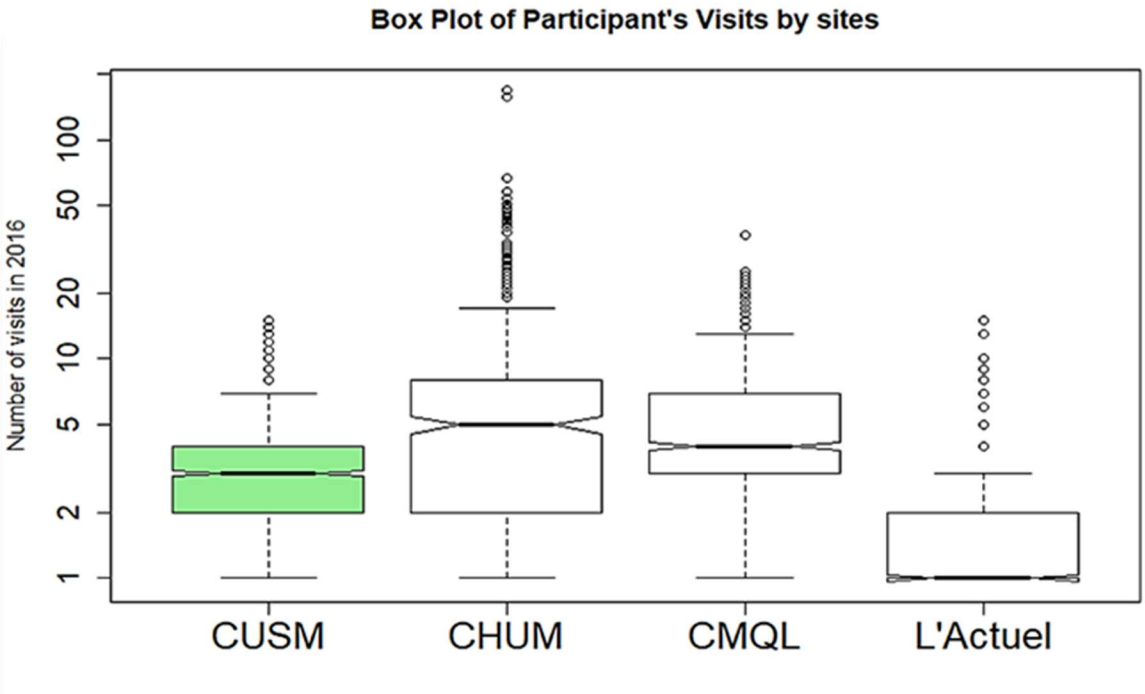

Figure 4.5: Participant's visits at four clinics included in CM**.**
Note: *the y-axis is log-transformed to improve readability.*

Prevalence of risk factors for HIV acquisition

In the CM dataset, there were four columns indicating the presence of risk factors. The first column has primary risk factors for HIV acquisition and the remaining data in columns two to four includes any additional risk factors that were prevalent. I visualized and identified the missing data for the risk factor one (5%), risk factor two (63.4%), risk factor three (93 %), and risk factor four (99.95%) (Appendix B, Figure 3). Since a substantial amount of the data were missing for risk factors 2, 3 and 4, only the primary risk factors were described (Table 4.1). Given that 99.5% of data in column 4 was not completed, this data was not included in the analysis (Appendix B, Figure 3).

Homosexuality (59%) was the most prevalent risk factors for HIV acquisition in the CM followed by heterosexuality (13%), native of endemic countries (8%), injection drug use (5%), and bisexuality (2%), blood transfusion (0%) and vertical transmission (0%) (Table 4.1).

Homosexuality was less prevalent at CUSM (27%) than the other clinics: CMQL (83%), L'Actuel (66%) and CHUM (54%). However, native of endemic country (19%) and heterosexuality (23%) were comparatively common at CUSM (Table 4.1).

CD4 Counts Analysis

Average CD4 count for PLHIV at the CUSM were similar to those followed up at the CHUM but were lesser than the CD4 counts of PLHIV at CMQL and L'Actuel (Table 4.3). Since the values for L'Actuel clinic had outliers, therefore I rely on median values for comparing the results.

Table 4.3: Frequency of mean, median and standard deviations of CD4 count across sites.

	CUSM	CHUM	CMQL	L'Actuel
Mean (SD)	487 (159)	481 (150)	546 (174)	720 (213)
Median [IQR]	459 [321-625]	459 [332-608]	526 [401-667]	545 [393-741]

Median CD4 counts of PLHIV followed up at CUSM (459) were comparable to CHUM (459) but were lesser than CMQL (526) and L'Actuel (545). In addition, the proportion of PLHIV in different subgroups based on the CD4 count shows that the proportions of PLHIV in each of the subcategories (CD4 count <200 or 200 -500 or >500) were comparable across all four sites. However, clinic L'Actuel had a relatively less (36%) proportion of PLHIV with CD4 counts of more than 200 but less than 500 cells/mm3.

Despite being different in socio-demographic characteristics, the quality of care provided and treatment adherence characteristics were similar for all the sub cohorts included in CM 2016 (Figure 4.8, Table 4.4).

Figure 4.2: Prevalence of average CD4 counts among PLHIV in CM (2016).

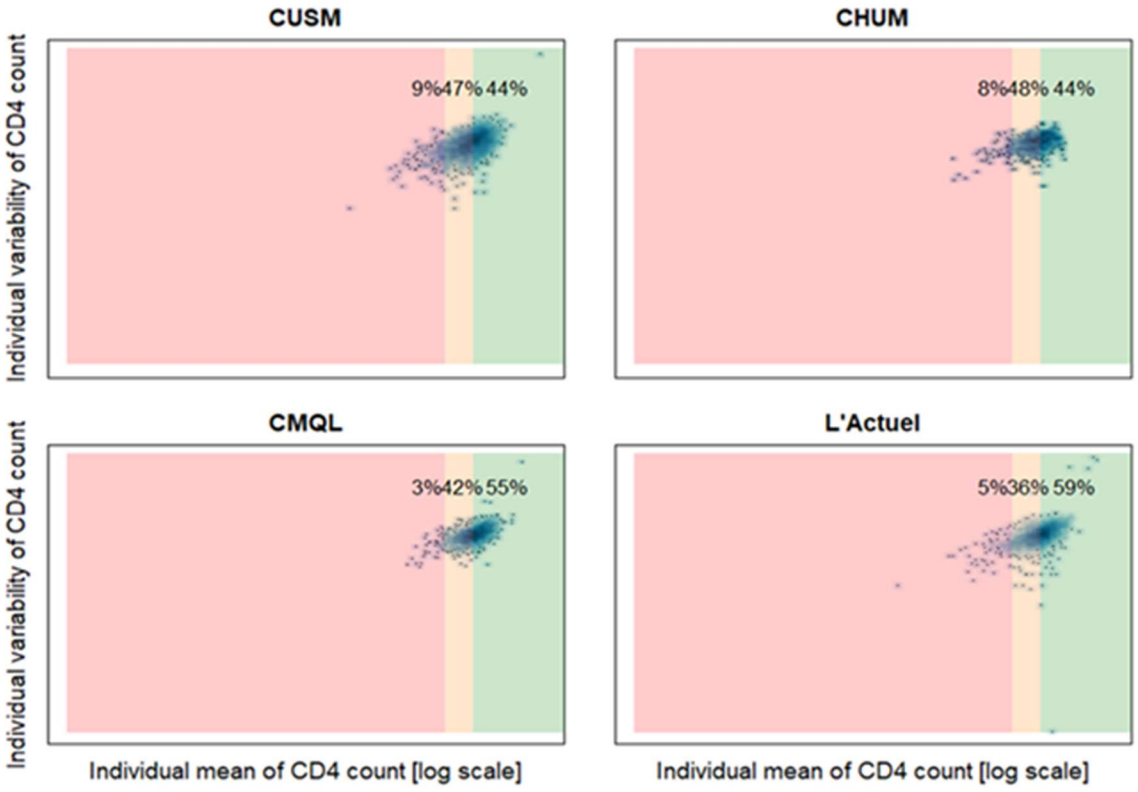

Table 4.4: Proportion of PLHIV according to CD4 count levels.

Variables	CUSM	CHUM	CMQL	L'Actuel
CD4 count <200	9%	8%	3%	5%
CD4 count >=200&<500	47%	48%	42%	36%
CD4 count >=500	44%	44%	55%	59%

HIV Viral Load analysis

PLHIV at the CUSM had primarily undetectable viral loads (<50 copies/mm of blood). Approximately 93% of the PLHIV at CUSM had undetectable viral loads for HIV and this was comparable to CHUM (89%) &CMQL (90%); however, L'Actuel had a relatively lower proportion of PLHIV with undetectable viral load (85%) (Figure 4.9, Table 4.5).

Both the undetectable viral loads and high CD4 counts are indicators for treatment adherence and good quality of care. These results indicate that despite being different in demographics, the clinical characteristics are still comparable across all sites included in CM.

Table 4.5: proportion of PLHIV having undetectable viral loads in CM (2016).

	CUSM	CHUM	CMQL	L'Actuel
Proportion of PLHIV with undetectable viral loads (95% CI)	0.93 (0.89, 0.96)	0.89 (0.84-0.92)	0.90 (0.84-0.94)	0.85 (0.82-0.94)

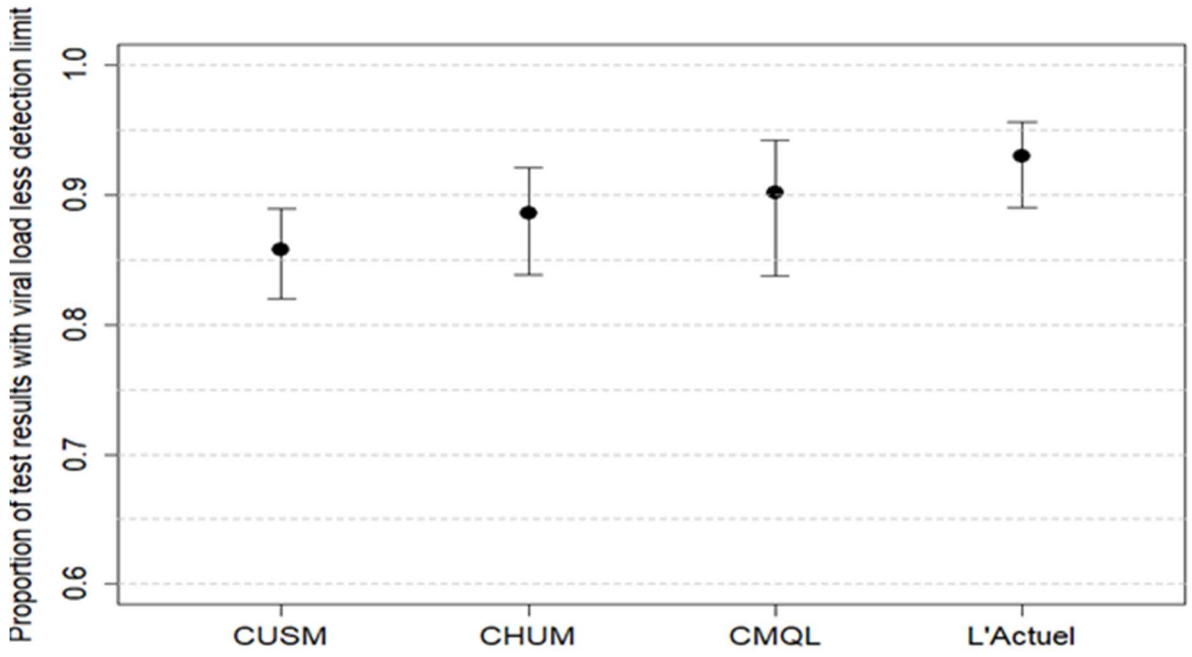

Figure 4.3: 95% confidence intervals for the proportions of test results with viral load less than detection limit.

5 Discussion:

My thesis research investigated the feasibility of a platform trial design for developing an electronic patient-reported outcome-based mHealth application. I first explored the clinical trial literature to understand the ethical, pragmatic, and technical issues associated with the implementation of platform trials. I then determined the demographic and clinical characteristics of the target population (PLHIV followed at the CVIS clinic at the MUHC that is more likely to be in the PT) and compared them to the study population (PLHIV followed up at other HIV clinics included in CM) in 2016. In this section, I will discuss my literature review and data analysis findings in the context of the I-Score study and will recommend avenues for future research.

Feasibility of PT in the context of I-score study

My literature review indicated that various **ethical issues** such as loss of equipoise, injustice, complexities with informed consent are inherent to adaptive platform trial designs [67]. In the context of determining the efficacy of mHealth apps; however, these ethical issues may not seriously affect the implementation of the proposed platform trial. To contextualize, the purpose of the I-Score study is to compare different versions of a mHealth app which is not equivalent to exposing patients to harmful interventions or treatment or withholding an effective treatment. Therefore, loss of equipoise should not substantially influence the outcome of the proposed PT. Likewise, feelings of injustice that may arise because of an unbalanced randomization scheme and withholding effective treatment may not be relevant in the context of evaluating a mHealth app.

Regardless, in a PT of I-Score, all the participants are more likely to receive the best version of the app at some point while being in the trial, which is expected to wave off this feeling of

injustice. I consider that explanation of the trial design and informed consent are of paramount importance in this regard. Besides, the inclusion of a suitable and balanced study population and adequate explanation can feasibly overcome the ethical issues related to the implementation of PT/MAMS/RAR and ultimately, can maximize the attainable benefits of a PT.

Confidentiality of the interim results is an important feature that preserves the **integrity of a trial** [77, 78]. However, in PT conduct, it is often difficult to maintain the confidentiality of the interim results, particularly, when a recruitment arm has to be stopped [77, 78]. For that, the establishment of an independent Data Management Committee (DMC) and a Trial Steering Committee (TSC) has been recommended [76, 78]. DMC (non-blind) can review interim data analyses and can make recommendations for TSC (blind to interim analysis) [76, 79, 138] .TSC's main role is to oversee the trial and have an independent role in deciding how the trial should be proceeded [76, 79]. TSC must approve any ad-hoc decisions (including non-implementation of any modification that was planned [76].

In light of the clinical trial literature [76, 79, 138, 139], I highly recommend the I-Score study to establish independent DMC (nonblind to interim data) and TSC (blind to interim data) to preserve the integrity of PT of I-Score study. The planned modifications should be agreed upon at the beginning of the trial and should not require the permission of TSC unless it is inevitably required like stopping a trial for futility reasons [76]. Regardless, to avoid any ambiguity, the adaptation rules should be clearly defined in the study protocol and charter of DMC and TSC [76]. The sponsors of the PT should ensure that a firewall is set up that prevents undue disclosure of sensitive information to the trial team [76].

Allocation concealment is done in randomized controlled trials to prevent selection bias [89]. However, in trials with adequate allocation concealment, selection bias[†††] can still arise if the recruiters can access the next allocation with more than 50 % probability [140]. Allocation concealment is often not possible in trials using outcome-based randomization because of imbalance in treatment assignments and can lead to selection bias [82, 89].

Authors of the first RAR-based clinical trial indicated the risk of selection bias and temporal drift because the nature of their intervention (i.e. extracorporeal membrane oxygenation) could not preserve the allocation concealment, also, patients might have different severity of illness. These issues later were addressed by concealing the allocation strategy from the clinicians and by selecting comparable disease severity patients across the phases of the study [82, 141]. Learning from this example [141] and in the light of the literature review, I recommend the I-Score study for using a centralized service for randomization allocation [142]. Such service provides an independent verification for allocation that cannot be known to or changed by the investigators [142]. Other recommended allocation concealment methods which may also do not increase the probability of guessing future allocation are simple randomization (which has a little effect on trial randomization), adjusting for the site of recruitment in analysis rather than stratifying it, minimization with a random element or urn modification, etc. [140].

Interim monitoring may also alter the behavior of researchers and other stakeholders in the platform trial which may lead to **investigators bias** [49]. For instance, just speculation of the interim results can affect the behavior of the stakeholders involved in the trial [76]. Likewise, inconsistencies in the conduct of trial across different stages (changes in the care provided or the

[†††] For instance, if the recruiters know the allocation probability they may assign the population of their own interest which effects the randomization and leads to selection bias in randomized trials 140. Kahan, B.C., S. Rehal, and S. Cro, *Risk of selection bias in randomised trials.* Trials, 2015. **16**(1): p. 405.

way the outcome is observed) may lead to operational bias and undermine the validity of the study results [76]. **operational bias/ investigators bias** can be controlled by preplanning and transparency, predefining the sponsor's role in decision making, establishing a competent DMC, blinding the trialist, and appropriate reporting of the results [76, 143, 144].

In light of these recommendations, I consider that the I-score study can feasibly control the bias by preplanning the modifications, predefining the primary and secondary outcome indicators and establishing the DMC and TSC. Planned modifications and blinding of the TSC are anticipated to ensure the integrity of the Platform trial.

Temporal drift is inherent to RARs and is associated with an increased risk of type 1 error [94]. To elaborate, changes in the characteristics (covariates) of the study population over time may lead to increased Type 1 error rate and false-negative error rates, particularly, in a single-arm trial; however, in two-arm trials, the drift decreases the error rate at expense of increasing the sample size [145]. Generally, temporal drift makes the platform trial design less suitable for long term-definitive Phase II and III trials [91].

Nonetheless, J. Ning and X. Huang developed a patient randomization scheme to adjust the covariate imbalance during RAR (for trials with binary outcomes) [146]. Adjustment of covariates in RAR can control the imbalance in characteristics (covariates) between the treatment arms, consequently, prevents the chances of a type 1 error rate [146].

In my opinion, PT of the I-score study is different from Phase II and III drug testing trials. Also, a large heterogeneous population has to be recruited over time as PT of the I-Score study to contextualize, personalize and generalize the app to a diverse population of Montreal. However, adjustment of the covariates during the randomization will minimize the risk for temporal drift

and associated type 1 error rate. Hence, with covariate adjustment, I anticipate that temporal drift may not significantly bias the results of PT of the I-Score.

In addition to various biases that may threaten the internal validity of the study, various associated **statistical issues** have been reported in clinical trial literature. For instance, changes in the sample mean and effect size leads to inaccuracies in confidence interval and p-values, which ultimately leads to biased effect size estimations and erroneous conclusion [53]. Generally, these statistical issues can be controlled by avoiding drastic changes in the inclusion and exclusion criteria and by rigorous statistical analysis [62] and simulations [85].

In my opinion, the PT of I-score can overcome such statistical issues. Since our team has a statistician with expertise in Bayesian statistics (TS) who has designed simulation algorithms for defining the sample size and controlling over the statistical issues and baseline characteristics of the PLHIV have been determined. In addition, we do not anticipate any drastic changes in the characteristics of the population or inclusion or exclusion criteria.

Various **logistical and administrative issues** are important to consider during the conduct of trial [49, 76]. For instance, planning, set-up, adaptation decisions and recruitment of new participants may take longer than expected. These time constraints may affect the efficiency of the trial design [76]. A recently published study also highlighted that further workup is required for improving the operational efficiencies of trial outcomes [147]. For instance, teamwork, use of resources, costing models that can express the funding cost and savings to the funding bodies are desired for the efficiency of the trial [147]. Also, many operator-dependent issues that can be controlled with adequate planning and appropriate investment of the resources.

PT design requires teamwork and frequent communications between researchers, clinicians, statisticians and other key stakeholders [139]. I believe the research coordinators can

play an important role here as they can collaborate with the investigators, sponsors and can boost up teamwork by setting up frequent team meetings required for the iterative process. I consider PT of I-score study can be able to overcome these logistical challenges preplanning and commitment of all stakeholders and enhanced communication between the team.

In MAMS designs, dropping of inefficient treatment arms may lead to **conflicts in power and sample size** [45] i.e., if we drop sample size, the power of the study decreases and there are more chances of type II error [45]. However, in PT accumulating evidence (from interim data) can be helpful in re-estimation of the sample size and allow the experimental arm(s) to leave the trial as soon the interim data permits [139].

In PT with RAR, participants recruited at the end of trial may be different from those recruited at the beginning of the trial, which raises concerns about the generalizability and external validity of the trial results [45]. This may be important in the context of HIV care trials because HIV is a chronic condition with patients having diverse clinical characteristics (i.e. viral loads and CD4 counts) depending on their adherence to ART. Generally, sample size and power conflicts and, in part, generalizability issues can be controlled by increasing sample size [76]. Since the I-Score study aims to conduct the trial in a big cohort of PLHIV in Montreal, a large sample size, hence high statistical power of the study and generalizability of the study results can be anticipated.

Limitations

The literature search was limited to the English language because of my linguistic competence, limited time and resources for translation. Nonetheless, I anticipate that this limitation may not bias the findings of review as some evidence exists in the literature about this [65]. In addition, I had limited time and resources and could not hire another reviewer.

Therefore, the design of choice for my study was the descriptive review of the literature (does not require any quality appraisal or additional reviewer). I explored only two databases (PubMed and EMBASE) and did not include grey literature because my research had a limited budget and I believe it was not worthy to spend the "increased costs of securing these difficult-to-locate studies" (p. 257) [148]. Being a physician with minimal experience and understanding of mathematics, I could not grasp the in-depth statistical concept related to platform trials discussed in the literature. Therefore, I excluded any articles that discussed or proposed complex methods or algorithms. In the future, I recommend an expert statistician to be included in the review team to unfold the complexities of statistical concepts.

Likewise, I identified a few shortcomings in the quantitative analysis relevant to the I-Score study. First, the Cohorte de Montreal is a clinical database extracted from patients' EMR data. It was not collected with a purpose to determine the baseline characteristics of the patients for clinical trials or recruit patients for clinical trials. I encountered various limitations during the analysis. For instance, I could not retrieve any information from the database about patients who had died, were lost to follow-up, or moved away which indicates maybe there may be a loss of information. In addition, several variables were not collected systematically; for example, the social history for the CUSM was completely missing. To prevent loss of information due to missing values, missing data were imputed to minimize potential selection biases. Similarly, injection drug use was recorded as three levels at CMQL and L'Atuel (current consumers, ex-consumers and no-IDU) but as two levels (consumers and non-consumers) at CHUM and was completely not recorded for CUSM. I recommend that the data collection pattern should be uniform across all sites. Second, there was a risk of misclassification of outcome; many patients from L'Actuel clinic were lost to follow-up and/or moved to CMQL that might have

misclassified the L'Actuel population. For instance, prevalence of white ethnicity is lowest (16%) at clinic L'Actuel while it is highest (84%) at clinic CMQL, probably; L'Actuel clinic's patients were misclassified into CMQL clinic.

Implications of findings and recommendations for future research

To my knowledge, this is the first study that summarized a diverse body of clinical trial literature in the context of ethical, pragmatic and technical issues related to platform trials. My research draws the attention of HIV clinicians and trialist towards this underutilized trial design. I anticipate that this study will be beneficial for HIV clinicians, researchers, biostatisticians, and trialists developing mHealth apps. Because this is the first study that describes the PT design in the context of developing health apps for HIV patients and indicates a few research gaps important for future research.

Despite the efficiency and cost-effectiveness of PT /MAMS/RAR adaptive designs, they remain underutilized, particularly, in HIV clinical care [100]. This lack of use can likely be explained by the additional amount of logistical and statistical work required for setting up an adaptive trial [100]. Also, possible increased demand for observing confidentiality and overcoming the ethical issues might discourage the trialist from implementing this trial design in HIV care research [100]. I consider that exploration and understanding of barriers to the implementation of PT designs in the context of HIV care demand a qualitative study in the future.

The descriptive literature review conducted indicates several important aspects of PTs to be considered in the design, planning, and conduct; however, I recommend a systematic review in the future that includes more databases, grey literature, quality appraisal, and more reviewers.

The systematic literature review could identify the inconsistencies in the definitions of the PT designs across the literature and the way they are rolled out. For instance, some PTs were conducted under a master protocol (e.g. Lung-Map [149] (NCT02154490) or NCI-MATCH [150] (NCT02465060)) whereas some used only Bayesian analytic strategy but not a master protocol (e.g. Sepsis-ACT [150](NCT02508649)) [139]. Recently, the Adaptive Platform Trials Coalition committee responded to this knowledge gap and endorsed the term "Adaptive Platform Trials" [139]. This term defines platform trials as a combination of a master protocol (rather than a stand-alone trial) with adaptive design (rather flexible than strict). I acknowledge this effort and recommend future platform trials, including the PT of the I-Score study, to be consistent with the term "Adaptive Platform Trials" which implies both an adaptive design and a master protocol.

Platform trials are different than conventional RCT in terms of frequency of scheduled interim data analyses. I consider it important that there should be separate reporting guidelines (other than CONSORT, Consolidated Standard of Reporting Trials) for publishing results of adaptive platform trials in the future. These reporting guidelines should accommodate the adaptive components (i.e. adaptive randomization) of the trial and provide an overview of the interim results as they emerged over the course of the trial.

6 Conclusion

The platform trial design is efficient, cost-effective, acceptable to stakeholders, operationally feasible, allows preplanning, and can accommodate a heterogeneous study population. The planned platform trial within the I-Score study can be a long-term resource to learn about and improve health and treatment among PLHIV in Montreal. Adequately defined inclusion criteria for the PT can accommodate a diverse and representative population in terms of sociodemographic, average CD4 cell counts and HIV viral loads. Henceforth, it will enable contextualization, personalization, and generalization of the results generated by the platform trial.

I anticipate that platform trials, MAMS and RAR will be feasible in the context of developing and evaluating mHealth apps. Particularly, developing a patient-reported outcome-based mHealth app can allow contextualization and personalization of such apps to identify maximum treatment adherence barriers of PLHIV. However, it remains important that the trial methodologists and statisticians creatively develop and improve the existing methodology for conducting adaptive platform trials to benefit the specialized patient population of PLHIV. Likewise, restriction to standardized definitions of platform trials and their reporting guidelines is of paramount importance for the future adaptive platform trials.

7 References

1. World Health Organization. *HIV/AIDS*. 2018, ; Available from: http://www.who.int/hiv/data/en/.
2. Organization, W.H., *HIV/AIDS Data and Statistics* 2019, World Health organization: Global. p. 1.
3. Haddad N, L.J., Totten S, McGuire M., *HIV in Canada-Surveillance Report, 2017*. 2018. Centre for Communicable Diseases and Infection Control, Public Health Agency of Canada, Ottawa, ON.: PubMed.
4. Gardner, E.M., et al., *The Spectrum of Engagement in HIV Care and its Relevance to Test-and-Treat Strategies for Prevention of HIV Infection.* Clinical Infectious Diseases, 2011. **52**(6): p. 793-800.
5. Sahay, S., K.S. Reddy, and S. Dhayarkar, *Optimizing adherence to antiretroviral therapy.* The Indian journal of medical research, 2011. **134**(6): p. 835.
6. *Canada's source for HIV and hepatitis C information*. Fall 2017; Available from: https://www.catie.ca/en/pif/fall-2017/getting-undetectable-population-inequities-canada.
7. Adefolalu, A.O. and Z.Z. Nkosi, *The complex nature of adherence in the management of HIV/AIDS as a chronic medical condition.* Diseases, 2013. **1**(1): p. 18-35.
8. Laws, M.B., et al., *Provider-patient adherence dialogue in HIV care: results of a multisite study.* AIDS and Behavior, 2013. **17**(1): p. 148-159.
9. Fredericksen, R., *Integrating a web-based, patient-administered assessment into primary care for HIV-infected adults.* Journal of AIDS and HIV Research, 2012. **4**(2).
10. Wilson, I.B., et al., *Provider-Focused Intervention Increases Adherence-Related Dialogue but Does Not Improve Antiretroviral Therapy Adherence in Persons With HIV.* JAIDS Journal of Acquired Immune Deficiency Syndromes, 2010. **53**(3): p. 338-347.
11. Roberts, K.J., *Physician Beliefs about Antiretroviral Adherence Communication.* AIDS Patient Care and STDs, 2000. **14**(9): p. 477-484.
12. Santana, M.-J. and D. Feeny, *Framework to assess the effects of using patient-reported outcome measures in chronic care management.* Quality of Life Research, 2014. **23**(5): p. 1505-1513.
13. Health, U.D.o., et al., *Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance.* Health and Quality of Life Outcomes, 2006. **4**: p. 1-20.

14. Engler, K., et al., *Engaging stakeholders into an electronic patient-reported outcome development study: On making an HIV-specific e-PRO patient-centered.* Health Policy and Technology, 2017. **6**(1): p. 59-66.
15. Engler, K., D. Lessard, and B. Lebouché, *A review of HIV-specific patient-reported outcome measures.* The Patient-Patient-Centered Outcomes Research, 2017. **10**(2): p. 187-202.
16. Greenhalgh, J., *The applications of PROs in clinical practice: what are they, do they work, and why?* Quality of Life Research, 2009. **18**(1): p. 115-123.
17. Deshpande, P.R., et al., *Patient-reported outcomes: A new era in clinical research.* Perspectives in clinical research, 2011. **2**(4): p. 137-144.
18. *How to Design and Develop a Mobile Health Application*. [Newsletter] 2017 May 12, 2017.; Available from: (https://mhealthintelligence.com/features/how-to-design-and-develop-a-mobile-health-application).
19. Cooper, V., et al., *mHealth Interventions To Support Self-Management In HIV: A Systematic Review.* The open AIDS journal, 2017. **11**: p. 119-132.
20. Behar, R., *76 percent of Canadians owned a smartphone in 2016, according to StatsCan*. 4th Nov, 2017.
21. Infoway, C.H., *One third of Canadians use Mobile Apps to Track Health: Study* 2017: Canada Health Infoway.
22. Cox, A., *Mobile Healthcare Opportunities Interactive Forecast Suite.* Smartphone Applications, Monitoring & mHealth Strategies, 2011. **2016**.
23. Franko, O.I. and T.F. Tirrell, *Smartphone app use among medical providers in ACGME training programs.* Journal of medical systems, 2012. **36**(5): p. 3135-3139.
24. Robustillo Cortés, M.d.l.A., et al., *High quantity but limited quality in healthcare applications intended for HIV-infected patients.* Telemedicine and e-Health, 2014. **20**(8): p. 729-735.
25. Lee, J.-A., et al., *Effective behavioral intervention strategies using mobile health applications for chronic disease management: a systematic review.* BMC medical informatics and decision making, 2018. **18**(1): p. 12-12.
26. *GPs to 'prescribe' apps for patients*.
27. Deeks, S.G., S.R. Lewin, and D.V. Havlir, *The end of AIDS: HIV infection as a chronic disease.* The Lancet, 2013. **382**(9903): p. 1525-1533.
28. Cortés, M.d.l.A.R. and R.M. Verdugo, *Quality of health websites adressed to patients with HIV+.* European journal of clinical pharmacy: atención farmacéutica, 2013. **15**(1): p. 13-21.
29. Amico, K.R., J.J. Harman, and B.T. Johnson, *Efficacy of antiretroviral therapy adherence interventions: a research synthesis of trials, 1996 to*

2004. JAIDS Journal of Acquired Immune Deficiency Syndromes, 2006. **41**(3): p. 285-297.
30. Perera, A.I., et al., *Effect of a smartphone application incorporating personalized health-related imagery on adherence to antiretroviral therapy: a randomized clinical trial.* AIDS patient care and STDs, 2014. **28**(11): p. 579-586.
31. Pellowski, J.A. and S.C. Kalichman, *Recent advances (2011-2012) in technology-delivered interventions for people living with HIV.* Current HIV/AIDS Reports, 2012. **9**(4): p. 326-334.
32. Muessig, K.E., et al., *A systematic review of recent smartphone, Internet and Web 2.0 interventions to address the HIV continuum of care.* Current HIV/AIDS Reports, 2015. **12**(1): p. 173-190.
33. Knobel, H., et al., *Validation of a simplified medication adherence questionnaire in a large cohort of HIV-infected patients: the GEEMA Study.* Aids, 2002. **16**(4): p. 605-613.
34. Walsh, J.C., S. Mandalia, and B.G. Gazzard, *Responses to a 1 month self-report on adherence to antiretroviral therapy are consistent with electronic data and virological treatment outcome.* Aids, 2002. **16**(2): p. 269-277.
35. Muessig, K.E., et al., *Mobile phone applications for the care and prevention of HIV and other sexually transmitted diseases: a review.* Journal of medical Internet research, 2013. **15**(1).
36. Engler, K., et al., *Barriers to antiretroviral therapy adherence in developed countries: a qualitative synthesis to develop a conceptual framework for a new patient-reported outcome measure.* AIDS care, 2018: p. 1-12.
37. Lessard, D., et al., *Evaluation of a project to engage patients in the development of a patient-reported measure for HIV care (the I-Score Study).* Health Expectations, 2018.
38. Toupin, I., et al., *Developing a patient-reported outcome measure for HIV care on perceived barriers to antiretroviral adherence: assessing the needs of HIV clinicians through qualitative analysis.* Quality of Life Research, 2018. **27**(2): p. 379-388.
39. Unhelkar, B. and S. Murugesan, *The Enterprise Mobile Applications Development Framework.* IT Professional, 2010. **12**(3): p. 33-39.
40. Chen, H., et al., *Dynamic software updating using a relaxed consistency model.* IEEE Transactions on Software Engineering, 2011. **37**(5): p. 679-694.
41. Saville, B.R. and S.M. Berry, *Efficiencies of platform clinical trials: a vision of the future.* Clinical Trials, 2016. **13**(3): p. 358-366.
42. Bajpai, R. and J. Car, *Could adaptive research designs be useful in designing an effective ehealth intervention? A methodological analysis.*

Trials. Conference: 4th International Clinical Trials Methodology Conference , ICTMC and the 38th Annual Meeting of the Society for Clinical Trials. United Kingdom, 2017. **18**(Supplement 1).
43. Lee, J.J., N. Chen, and G. Yin, *Worth adapting? Revisiting the usefulness of outcome-adaptive randomization.* Clinical Cancer Research, 2012.
44. Lawler, M. and R. Sullivan, *Personalised and Precision Medicine in Cancer Clinical Trials: Panacea for Progress or Pandora's Box?* Public Health Genomics, 2015. **18**(6): p. 329-37.
45. Guidance, F.D., *Adaptive design clinical trials for drugs and biologics*. 2010.
46. Lin, J. and V. Bunn, *Comparison of multi-arm multi-stage design and adaptive randomization in platform clinical trials.* Contemporary Clinical Trials, 2017. **54**: p. 48-59.
47. *Adaptive trial designs*.
48. Berry, S.M., J.T. Connor, and R.J. Lewis, *The platform trial: An efficient strategy for evaluating multiple treatments.* JAMA - Journal of the American Medical Association, 2015. **313**(16): p. 1619-1620.
49. Renfro, L.A. and S.J. Mandrekar, *Definitions and statistical properties of master protocols for personalized medicine in oncology.* Journal of Biopharmaceutical Statistics, 2017: p. 1-12.
50. Berry, S.M., J.T. Connor, and R.J. Lewis, *The platform trial: an efficient strategy for evaluating multiple treatments.* Jama, 2015. **313**(16): p. 1619-1620.
51. Berry, S.M., et al., *A response adaptive randomization platform trial for efficient evaluation of Ebola virus treatments: A model for pandemic response.* Clinical Trials, 2016. **13**(1): p. 22-30.
52. Lee, J.J., N. Chen, and G. Yin, *Worth adapting? Revisiting the usefulness of outcome-adaptive randomization.* Clinical Cancer Research, 2012. **18**(17): p. 4498-4507.
53. Korn, E.L. and B. Freidlin, *Adaptive Clinical Trials: Advantages and Disadvantages of Various Adaptive Design Elements.* Journal of the National Cancer Institute, 2017. **109**(6).
54. Sampson, A.R. and M.W. Sill, *Drop-the-losers design: normal case.* Biometrical Journal: Journal of Mathematical Methods in Biosciences, 2005. **47**(3): p. 257-268.
55. Wason, J.M. and L. Trippa, *A comparison of Bayesian adaptive randomization and multi-stage designs for multi-arm clinical trials.* Statistics in medicine, 2014. **33**(13): p. 2206-2221.
56. Kaplan, R., *The FOCUS4 design for biomarker stratified trials.* Chinese clinical oncology, 2015. **4**(3): p. 35.

57. Carthon, B.C. and E.S. Antonarakis, *The STAMPEDE trial: Paradigm-changing data through innovative trial design.* Translational Cancer Research, 2016. **5**(Supplement3): p. S485-S490.
58. Junker, A., *I-SPY 2 project with adaptive randomization: First data on neratinib and veliparib in breast carcinoma.* Krankenhauspharmazie, 2017. **38**(4): p. 212-213.
59. Papadimitrakopoulou, V., et al., *Battle-2 program: A biomarker-integrated targeted therapy study in previously treated patients with advanced non-small cell lung cancer (NSCLC).* Journal of Thoracic Oncology, 2013. **2)**: p. S367.
60. Mehta, A.D.S.a.S. *Adaptive Platform Trials: The Clinical Trial of the Future?* 2017; Available from: https://www.hbs.edu/faculty/Pages/item.aspx?num=53315.
61. Berry, D.A., *Bayesian clinical trials.* Nature reviews Drug discovery, 2006. **5**(1): p. 27.
62. Connor, J.T., J.J. Elm, and K.R. Broglio, *Bayesian adaptive trials offer advantages in comparative effectiveness trials: An example in status epilepticus.* Journal of Clinical Epidemiology, 2013. **66**(8 SUPPL.8): p. S130-S137.
63. Cellamare, M., et al., *Bayesian adaptive randomization in a clinical trial to identify new regimens for MDR-TB: the endTB trial.* Int J Tuberc Lung Dis, 2016. **20**(12): p. 8-12.
64. Kruse, S.D. and A. Warbel, *Developing a comprehensive literature review: An inquiry into method*. 2010.
65. Moher, D., et al., *The inclusion of reports of randomised trials published in languages other than English in systematic reviews.* Health Technol Assess, 2003. **7**(41): p. 1-90.
66. London, A.J., *Learning health systems, clinical equipoise and the ethics of response adaptive randomisation.* J Med Ethics, 2017.
67. Laage, T., et al., *Ethical considerations in adaptive design clinical trials.* Therapeutic innovation & regulatory science, 2017. **51**(2): p. 190-199.
68. Freedman, B., *Equipoise and the ethics of clinical research.* New England Journal of Medicine,(th July), 1987.
69. Thall, P.F., *Ethical issues in oncology biostatistics.* Statistical Methods in Medical Research, 2002. **11**(5): p. 429-448.
70. Saxman, S.B., *Ethical considerations for outcome-adaptive trial designs: a clinical researcher's perspective.* Bioethics, 2015. **29**(2): p. 59-65.
71. Stallard, N., et al., *Adaptive designs for confirmatory clinical trials with subgroup selection.* Journal of biopharmaceutical statistics, 2014. **24**(1): p. 168-187.

72. *Revising the Declaration of Helsinki*. 2013: online world report. p. 1532–33.
73. Richard A Wagner, M., PhD. *Informed consent* Available from: https://www.emedicinehealth.com/informed_consent/article_em.htm#what_is_informed_consent.
74. Begg, C.B., *Ethical concerns about adaptive randomization.* Clinical Trials, 2015. **12**(2): p. 101.
75. Buyse, M., E.D. Saad, and T. Burzykowski, *Adaptive randomization of neratinib in early breast cancer.* New England Journal of Medicine, 2016. **375**(16): p. 1591-1592.
76. Pallmann, P., et al., *Adaptive designs in clinical trials: why use them, and how to run and report them.* BMC medicine, 2018. **16**(1): p. 29.
77. He, W., et al., *Addressing challenges and opportunities of "less well-understood" adaptive designs.* Therapeutic Innovation & Regulatory Science, 2017. **51**(1): p. 60-68.
78. Calis, K.A., et al., *Recommendations for data monitoring committees from the Clinical Trials Transformation Initiative.* Clinical Trials, 2017. **14**(4): p. 342-348.
79. Sanchez-Kam, M., et al., *A practical guide to data monitoring committees in adaptive trials.* Therapeutic Innovation & Regulatory Science, 2014. **48**(3): p. 316-326.
80. Laage, T., et al., *Ethical considerations in adaptive design clinical trials.* Therapeutic Innovation and Regulatory Science, 2017. **51**(2): p. 190-199.
81. Hey, S.P. and J. Kimmelman, *Are outcome-adaptive allocation trials ethical?* Clinical Trials, 2015. **12**(2): p. 102-106.
82. Gupta, S., et al., *A framework for applying unfamiliar trial designs in studies of rare diseases.* Journal of Clinical Epidemiology, 2011. **64**(10): p. 1085-1094.
83. Berry, D.A., *Adaptive clinical trials: the promise and the caution*. 2011.
84. Gallo, P., *Confidentiality and trial integrity issues for adaptive designs.* Drug Information Journal, 2006. **40**(4): p. 445-450.
85. Cook, J.D., *The Effect of Population Drift on Adaptively Randomized Trials.* 2007.
86. Korn, E.L. and B. Freidlin, *Outcome-adaptive randomization: Is it useful?* Journal of Clinical Oncology, 2011. **29**(6): p. 771-776.
87. Simon, R., *Adaptive treatment assignment methods and clinical trials.* Biometrics, 1977: p. 743-749.
88. Cellamare, M., et al., *A Bayesian response-adaptive trial in tuberculosis: The endTB trial.* Clinical Trials, 2017. **14**(1): p. 17-28.
89. hub, P.S. *Allocation Concealment* [cited 2019 16 Aug]; Available from: https://psychscenehub.com/psychpedia/allocation-concealment/.

90. Korn, E.L. and B. Freidlin, *Adaptive clinical trials: advantages and disadvantages of various adaptive design elements.* JNCI: Journal of the National Cancer Institute, 2017. **109**(6).
91. Faderl, S., et al., *A randomized study of clofarabine versus clofarabine plus low-dose cytarabine as front-line therapy for patients aged 60 years and older with acute myeloid leukemia and high-risk myelodysplastic syndrome.* Blood, 2008. **112**(5): p. 1638-1645.
92. Simon, R. and N.R. Simon, *Using randomization tests to preserve type I error with response adaptive and covariate adaptive randomization.* Statistics & probability letters, 2011. **81**(7): p. 767-772.
93. GUIDANCE, D., *Adaptive Designs for Clinical Trials of Drugs and Biologics.* Center for Biologics Evaluation and Research (CBER), 2018.
94. Graf, A.C. and P. Bauer, *Maximum inflation of the type 1 error rate when sample size and allocation rate are adapted in a pre-planned interim look.* Statistics in medicine, 2011. **30**(14): p. 1637-1647.
95. Chow, S.-C., M. Chang, and A. Pong, *Statistical consideration of adaptive methods in clinical development.* Journal of biopharmaceutical statistics, 2005. **15**(4): p. 575-591.
96. Feng, H., J. Shao, and S.-C. Chow, *Adaptive group sequential test for clinical trials with changing patient population.* Journal of biopharmaceutical statistics, 2007. **17**(6): p. 1227-1238.
97. Villar, S.S., J. Bowden, and J. Wason, *Response-adaptive designs for binary responses: How to offer patient benefit while being robust to time trends?* Pharmaceutical statistics, 2018. **17**(2): p. 182-197.
98. Lipsky, A.M. and S. Greenland, *Confounding due to changing background risk in adaptively randomized trials.* Clinical Trials, 2011. **8**(4): p. 390-397.
99. FDA. *US Food and Drug administration. Adaptive designs for medical device clinical studies: guidance for industry and Food and Driug Administration staff.* 2016; Available from: https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm446729.pdf.
100. Moore, C.L., et al., *Multi-arm, multi-stage randomised controlled trials for evaluating therapeutic HIV cure interventions.* The Lancet HIV, 2019. **6**(5): p. e334-e340.
101. Esserman, L.J. and J. Woodcock, *Accelerating identification and regulatory approval of investigational cancer drugs.* Jama, 2011. **306**(23): p. 2608-2609.
102. Steuer, C., et al., *Innovative clinical trials: the LUNG-MAP study.* Clinical Pharmacology & Therapeutics, 2015. **97**(5): p. 488-491.

103. Saville, B. and P.D.S. Berry, *Tailoring Therapies in Adaptive Platform Clinical Trials*. 2015.
104. Yuan, Y. and G. Yin, *BAYESIAN PHASE I/II ADAPTIVELY RANDOMIZED ONCOLOGY TRIALS WITH COMBINED DRUGS.* Ann Appl Stat, 2011. **5**(2a): p. 924-942.
105. Ajaz, M., et al., *Current and investigational drug strategies for glioblastoma.* Clin Oncol (R Coll Radiol), 2014. **26**(7): p. 419-30.
106. Bleck, T., et al., *The established status epilepticus trial 2013.* Epilepsia, 2013. **54**(SUPPL. 6): p. 89-92.
107. Fieten, K.B., et al., *Effectiveness of alpine climate treatment for children with difficult to treat atopic dermatitis: Results of a pragmatic randomized controlled trial (DAVOS trial).* Clin Exp Allergy, 2018. **48**(2): p. 186-195.
108. van Haren, F., *Personalised fluid resuscitation in the ICU: Still a fluid concept?* Critical Care, 2017. **21 (Supplement 3) (no pagination)**(313).
109. Collins, S.P., et al., *Bayesian adaptive trial design in acute heart failure syndromes: moving beyond the mega trial.* American heart journal, 2012. **164**(2): p. 138-145.
110. Ellis, R.J., et al., *Randomized trial of central nervous system-targeted antiretrovirals for HIV-associated neurocognitive disorder.* Clin Infect Dis, 2014. **58**(7): p. 1015-22.
111. Morris, A., et al., *Use of rosuvastatin in HIV-associated chronic obstructive pulmonary disease.* Aids, 2017. **31**(4): p. 539-544.
112. Choko, A.T., et al., *HIV self-testing alone or with additional interventions, including financial incentives, and linkage to care or prevention among male partners of antenatal care clinic attendees in Malawi: An adaptive multi-arm, multi-stage cluster randomised trial.* PLoS Med, 2019. **16**(1): p. e1002719.
113. Ellis, R.J., et al., *Randomized trial of central nervous system-targeted antiretrovirals for HIV-associated neurocognitive disorder.* Clinical Infectious Diseases, 2014. **58**(7): p. 1015-1022.
114. Blake Linthwaite, N.S., Helen Trottier, Jean-Guy Baril, Madeleine Durand, Marina Klein, Nimâ Machouf, Costas Pexos, Réjean Thomas, Cécile Tremblay, Anne-Fanny Vassal, and Joseph Cox, *In-care HIV cascades for the city of Montreal: Data from the Cohorte Montréalaise.*
115. *Montréal signe la Déclaration de Paris visant à l'éradication du VIH/SIDA*. Available from: http://ville.montreal.qc.ca/portal/page?_pageid=5798,42657625&_dad=portal&_schema=PORTAL&id=29427.
116. AIDS, G.i.a.e.o.H.a. *GLOBAL HIV TARGETS*. 2018 26 March 2019; Available from: https://www.avert.org/global-hiv-targets.

117. Setia, M.S., *Methodology Series Module 3: Cross-sectional Studies.* Indian journal of dermatology, 2016. **61**(3): p. 261-264.
118. Control, C.f.D. and Prevention, *Characteristics associated with HIV infection among heterosexuals in urban areas with high AIDS prevalence---24 cities, United States, 2006-2007.* MMWR. Morbidity and mortality weekly report, 2011. **60**(31): p. 1045.
119. Cherry, K. *Cross-Sectional Research Method: How Does It Work? Advantages and Challenges*
2019 June 29, 2019 [cited 2019 7/15]; Available from: https://www.verywellmind.com/what-is-a-cross-sectional-study-2794978.
120. Gaille, B. *15 Cross Sectional Study Advantages and Disadvantages*. 2018 13 Aug, 2018; Available from: https://brandongaille.com/15-cross-sectional-study-advantages-and-disadvantages/.
121. *Cross-sectional or Prevalence Studies*. 2019; Available from: https://cursos.campusvirtualsp.org/mod/tab/view.php?id=34149&forceview=1.
122. Services, U.d.a.H.a.H. *CD4 Counts* Available from: https://aidsinfo.nih.gov/understanding-hiv-aids/glossary/822/cd4-count.
123. Affairs, U.D.o.V. *CD4 count (or T-cell count)*. Available from: https://www.hiv.va.gov/patient/diagnosis/labs-CD4-count.asp.
124. services, U.D.o.H.a.H. *Viral load test*. Available from: https://aidsinfo.nih.gov/understanding-hiv-aids/glossary/743/viral-load-test.
125. services, U.D.o.H.a.h., *Undetectable viral loads*
126. Services, U.D.o.H.a.H. *Viral supression* Available from: https://aidsinfo.nih.gov/understanding-hiv-aids/glossary/1650/viral-suppression.
127. Affairs, U.D.o.V. *HIV viral load (or 'HIV RNA')*. [cited 2019 7th June]; Available from: https://www.hiv.va.gov/patient/diagnosis/labs-viral-load.asp.
128. Medicine, U.N.L.o. *Injection Drug Use* 2019 June 7, 2019; Available from: https://aidsinfo.nih.gov/understanding-hiv-aids/glossary/1585/injection-drug-use.
129. Boris Garber, J.G. *Complications of Injection Drug Abuse*. 2015; Available from: https://www.reliasmedia.com/articles/136655-complications-of-injection-drug-abuse.
130. (PHAC), P.H.A.o.C., *People from countries where HIV in Endemic*. 2009.
131. William C. Shiel Jr., M., FACP, FACR. *Medical Definition of Blood transfusion*. 2018 12/4/2018 [cited 2019 7th June]; Available from: https://www.medicinenet.com/script/main/art.asp?articlekey=2491.

132. William C. Shiel Jr., M., FACP, FACR. *Medical Definition of Vertical transmission*. 2018 12/27/2018 [cited 2019 7th June]; Available from: https://www.medicinenet.com/script/main/art.asp?articlekey=7982.
133. Dictionary.com. *Homosexuality*. Available from: dictionary.com/browse/homosexuality.
134. Dictionary.com. *Heterosexuality*. Available from: https://www.dictionary.com/browse/heterosexuality?s=t.
135. William C. Shiel Jr., M., FACP, FACR *Medical Definition of Bisexual*. 2018 12/27/2018; Available from: https://www.medicinenet.com/script/main/art.asp?articlekey=10794.
136. *R Core Team (2019). R: A language and environment for statistical computing.* . Available from: https://www.R-project.org/.
.
137. Stef van Buuren, K.G.-O., *mice: Multivariate Imputation by Chained Equations in R.* Journal of Statistical Software 2011. **45 (2011)** (Issue 3): p. 1-67.
138. DeMets, D.L. and S.S. Ellenberg, *Data monitoring committees—expect the unexpected.* New England Journal of Medicine, 2016. **375**(14): p. 1365-1371.
139. Angus, D.C., et al., *Adaptive platform trials: definition, design, conduct and reporting considerations.* 2019.
140. Kahan, B.C., S. Rehal, and S. Cro, *Risk of selection bias in randomised trials.* Trials, 2015. **16**(1): p. 405.
141. Bartlett, R.H., et al., *Extracorporeal circulation in neonatal respiratory failure: a prospective randomized study.* Pediatrics, 1985. **76**(4): p. 479-487.
142. *BLINDING AND ALLOCATION CONCEALMENT*. 2019; Available from: https://www.sealedenvelope.com/randomisation/blinding/.
143. Gallo, P., *Operational challenges in adaptive design implementation.* Pharmaceutical Statistics: The Journal of Applied Statistics in the Pharmaceutical Industry, 2006. **5**(2): p. 119-124.
144. Chow, S.-C., R. Corey, and M. Lin, *On the independence of data monitoring committee in adaptive design clinical trials.* Journal of biopharmaceutical statistics, 2012. **22**(4): p. 853-867.
145. Tang, H., et al., *Comparison of error rates in single-arm versus randomized phase II cancer clinical trials.* Journal of Clinical Oncology, 2010. **28**(11): p. 1936.
146. Ning, J. and X. Huang, *Response-adaptive randomization for clinical trials with adjustment for covariate imbalance.* Statistics in medicine, 2010. **29**(17): p. 1761-1768.

147. Schiavone, F., et al., *This is a platform alteration: a trial management perspective on the operational aspects of adaptive and platform and umbrella protocols.* Trials, 2019. **20**(1): p. 264.
148. Conn, V.S., et al., *Grey Literature in Meta-Analyses.* Nursing Research, 2003. **52**(4): p. 256-261.
149. Steuer, C., et al., *Innovative clinical trials: the LUNG-MAP study.* Clinical Pharmacology & Therapeutics, 2015. **97**(5): p. 488-491.
150. Conley, B.A., et al., *NCI-MATCH (Molecular Analysis for Therapy Choice)–a national signal finding trial*. 2016, American Society of Clinical Oncology.

8 Appendices

Appendix A

Table 1. Baseline Demographic and Clinical Characteristics of 49 Participants in Primary Intent-to-Treat Analysis

Characteristic	CNS Targeted	Comparison	*P* Value[a]
No.	26	23	
Age, y, mean (SD)	44.9 (8.7)	43.6 (11.6)	.79
Sex, male	22 (85%)	17 (74%)	.48
Race/ethnicity			.18
White, non-Hispanic	10 (38%)	9 (39%)	
Black	12 (46%)	13 (57%)	
Hispanic	4 (15%)	0 (0%)	
Other	0 (0%)	1 (4%)	
Education, y, mean (SD)	12.1 (2.6)	12.4 (1.6)	.68
Current CD4 count, cells/μL, median (range)	214 (5–964)	306 (3–1224)	.27
Prior nadir CD4 count <200 cells/μL	16 (67%)	8 (38%)	.08
Plasma HIV RNA[b], median (range)	4.2 (1.7–5.9)	3.5 (1.7–6.2)	.71
Plasma HIV RNA undetectable	7 (27%)	6 (26%)	>.99
CSF HIV RNA[b], median (range)	3.1 (1.7–4.6)	3.1 (1.7–5)	.52
CSF HIV RNA undetectable	7 (27%)	7 (30%)	>.99
Prior AIDS-defining illness	11 (46%)	7 (32%)	.75
Prior ARV treatment	17 (65%)	17 (74%)	.55
HCV coinfection	9 (35%)	3 (13%)	.10
Baseline GDS, mean (SD)	0.88 (0.6)	0.92 (0.6)	.84

Figure 1: Characteristics of PLHIV participating in PT for ART in neuro cognitive disorders. [110].

Baseline demographics for the cohort and by treatment group.

	All (N=22) Median (Q1, Q3) n (%)	Placebo (N=11) Median (Q1, Q3) n (%)	Rosuvastatin (N=11) Median (Q1, Q3) n (%)
Age (years)	50.3 (47.1, 53.2)	49.2 (48.5, 52.9)	51.3 (45, 54.4)
Female, n (%)	7 (31.8)	3 (27.3)	4 (36.4)
African-American, n (%)	12 (54.5)	6 (54.5)	6 (54.5)
Hispanic, n (%)	1 (4.5)	1 (9.1)	0 (0)
Body mass index (kg/m^2)	24.5 (20.9, 27.6)	25.7 (21.3, 26.6)	21.7 (20.6, 27.9)
Ever smoked cigarettes, n (%)	19 (86.4)	10 (90.9)	9 (81.8)
Total pack years smoked	12.1 (4.8, 27.4)	11.7 (10, 20.9)	13 (4, 27.1)
CD4 cell count (cells/μl)	630 (526, 784)	660 (539.5, 866)	624 (505.5, 756.5)
HIV viral level < 40 copies/ml, n (%)	17 (77.3)	9 (81.8)	8 (72.7)
ART use, n (%)	21 (95.5)	11 (100)	10 (90.9)
hs-CRP (mg/l)	0.27 (0.13, 0.6)	0.18 (0.1, 0.37)	0.40 (0.17, 1.07)
Total cholesterol (mg/dl)	164 (148, 204.3)	160 (144.5, 175.5)	176 (157.5, 209)
HDL (mg/dl)	45 (36.3, 66.5)	41 (36.5, 58.5)	47 (36.5, 72.5)
Total cholesterol/HDL	3.7 (2.5, 4.8)	3.7 (3.4, 4.2)	4.7 (2.2, 5.2)
LDL (mg/dl)	89 (69, 105)	89 (81, 98.5)	101.5 (69, 133.5)

Abbreviations: ART, antiretroviral therapy; hs-CRP, high sensitivity C-reactive protein; HDL, high density lipoprotein; LDL, low density lipoprotein; SD, standard deviation; Q, quartile.

Figure 2: Characteristics of PLHIV participating in a PT for evaluating rosuvastatin in pulmonary disease in PLHIV [111].

Table 1: Search Strategy for Embase Classic and Embase on 5 February 2018

#	Search statement	Results
1	Platform trial*.mp.	115
2	Multi-arm multi-stage.mp	73
3	Multiarm multistage.mp	7
4	Multistage multiarm.mp	1
5	Adaptive randomisation*.mp	12
6	Adaptive randomization*.mp	264
7	1 or 2 or 3 or 4 or 5 or 6	455
8	limit 7 to English	446

Table 2: Search strategy for PubMed on 5 February 2018

#	Search statement	Results
`1	Search platform trial*[Text Word]	72
2	Search adaptive randomization[Text Word]	190
3	Search adaptive randomisation[Text Word]	13
4	Search ((multi-arm[Text Word] OR multiarm[Text Word]) AND (multi-stage[Text Word] OR multistage[Text Word]))	55
5	#1OR #2 OR#03 OR#4	**320**
6	Search Chronic disease[MeSH Terms]	246055
7	Search HIV infections[MeSH Major Topic]	221906
8	Search (HIV[Text Word]) OR Chronic disease[Text Word]	585640
9	Search (#6 OR #7 OR #8 OR)	618671

10	Search (#9 AND #5)	**419**

Table 3: Search strategy for PubMed on 2 Nov 2018:

#	Search statement	Results
1	Search ((Search[All Fields] AND (((platform trial[Text Word] OR platform trials[Text Word]) OR Multi-arm multi-stage[Text Word]) OR (Multi arm[All Fields] OR multistage[Text Word])) OR (Multistage[All Fields] OR multi arm[Text Word])) AND Clinical Trial[ptyp] AND Humans[Mesh] AND English[lang]) Filters: Clinical Trial; Humans; English	277
2	Search ((((Patient reported outcome measures*[MeSH Terms]) OR Patient reported outcomes*[Text Word]) OR PROS[Text Word]) OR Electronic patient reported outcomes*[Text Word]) OR ePROS[Text Word] Filters: Clinical Trial; Humans; English	1422
3	Search #1 AND #2 Filters: Clinical Trial; Humans; English	**0**

Table 4: Search strategy for PubMed on 2^{4th} April 2019:

#	Search statement	Results
1	Search ((Search[All Fields] AND (((platform trial[Text Word] OR platform trials[Text Word]) OR Multi-arm multi-stage[Text Word]) OR (Multi arm[All Fields] OR multistage[Text Word])) OR (Multistage[All Fields] OR multi arm[Text Word])) AND Clinical Trial[ptyp] AND Humans[Mesh] AND English[lang]) Filters: Clinical Trial; Humans; English	378
2	Search ((((Patient reported outcome measures*[MeSH Terms]) OR Patient reported outcomes*[Text Word]) OR PROS[Text Word]) OR Electronic patient reported outcomes*[Text Word]) OR ePROS[Text Word] Filters: Clinical Trial; Humans; English	19371
3	Search ((Search (#1) AND #2 Filters: English))	0

Table 5: Subject headings and key words searched in PubMed.

Searched Terms for different concepts:	Concept #1	Concept #2	Concept #3	Concept #4
Subject Heading 1			Chronic Disease	Patient reported outcome measures*
OR Subject Heading 2		HIV infections*		
OR Subject Heading 3		HIV		
Etc.		HIV AIDS		
OR Keyword 1	Platform trial*	HIV	Chronic disease	Patient reported outcomes*
OR Keyword 2	Multi-arm multi-stage	HIV AIDS		PROS

OR Keyword 3	Multiarm multistage	AIDS		Electronic patient reported outcomes*
Etc.	Multistage multiarm			ePROS
	Adaptive randomi?ation*			

Data Extraction Charts
The following date extraction charts

a)-Chart 1 comprehensive literature review b) Chart 2a pragmatic, technical and operational challenges c) Chart 2b ethical issues are available at the following link.

https://onedrive.live.com/?id=98F715E0BE963BA0%211021&cid=98F715E0BE963BA0

Appendix B

The following permissions and approvals are available on the link

https://onedrive.live.com/?id=98F715E0BE963BA0%211021&cid=98F715E0BE963BA0

a) REB approval for quantitative data analysis (CM) data b) permission to access Cohortd de Montreal data email (pdf) by Dr. Jean Guy Brail c) permission from Harvard university d) Permission to reproduce the copy right diagram of definitions of Platform trials.

Visualization of missing data

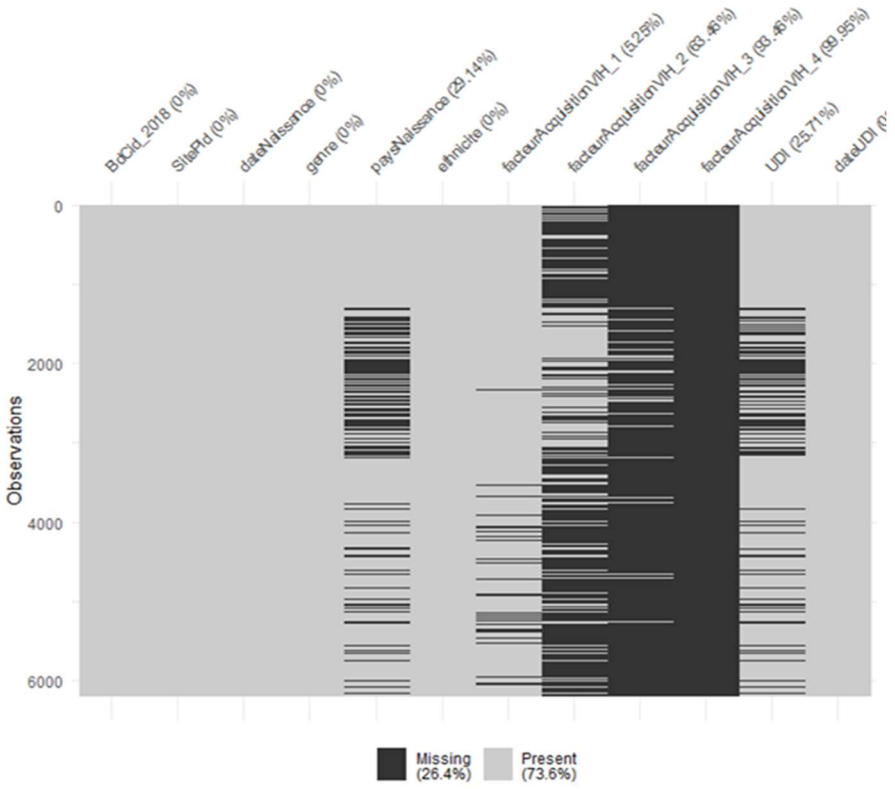

Figure 3: Missing data visualization for demographic dataset

Figure 4: Missing data visualization for anti HCV antibody status dataset

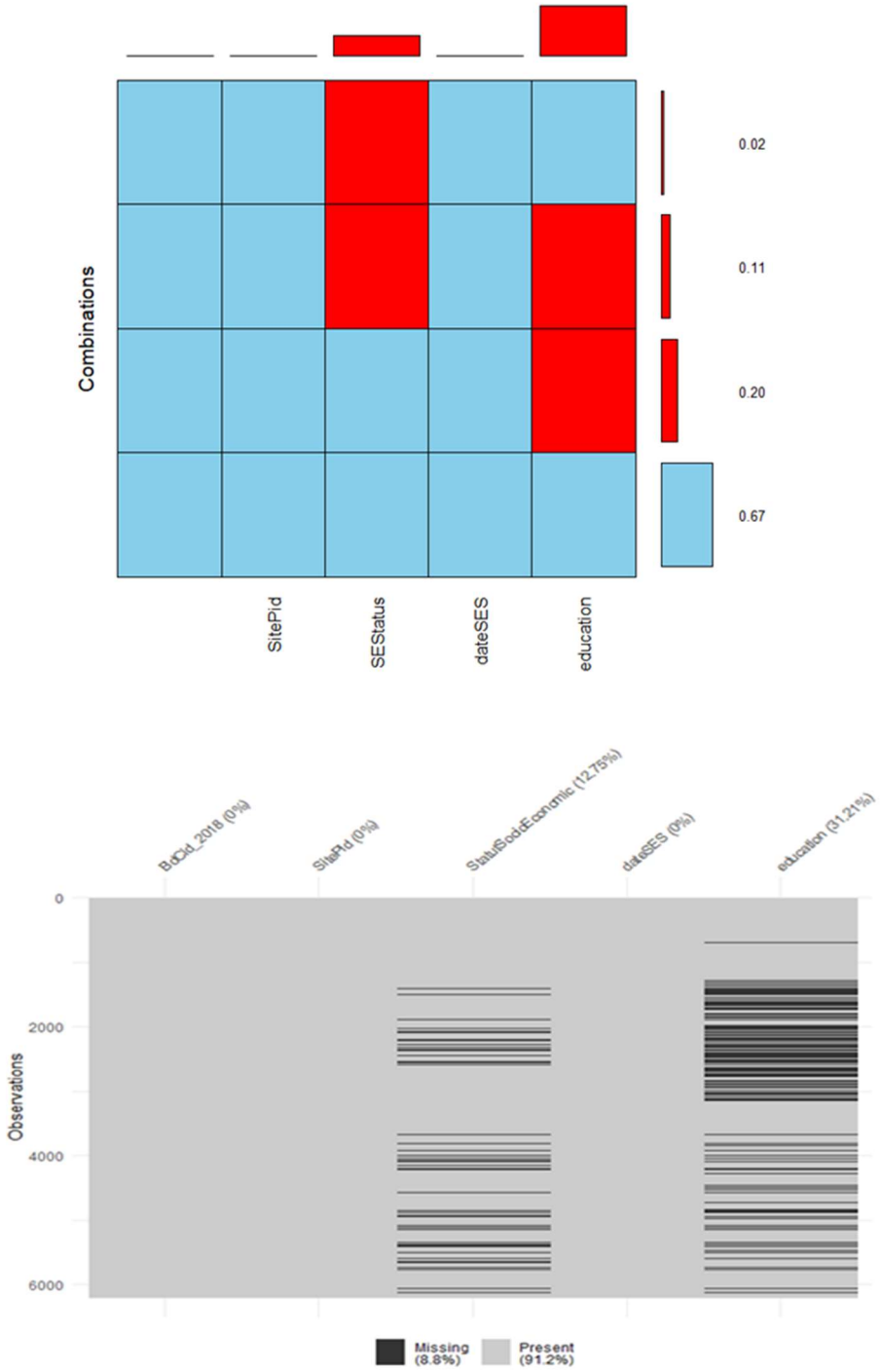

Figure 5: Missing data visualization for socioeconomic dataset

Data sets used in analysis (translated to English):

The following data set used for analysis are available on this link.

https://onedrive.live.com/?id=98F715E0BE963BA0%211021&cid=98F715E0BE963BA0

a). Demographic data 9un-CDM data-D.csv b) CDMdata-HCV-un.csv c) CDM data CD4 (2), d) Viral load data .csv e) CDM data-Visits

Note: the attached datasets are not the raw data instead it is translated, sorted and cleaned version of the data.

R script for quantitative analysis:

Codes or R script (pdf and R files) used for quantitative data analysis (phase 2) are available on the https://onedrive.live.com/?id=98F715E0BE963BA0%211021&cid=98F715E0BE963BA0