Evaluation of HIV treatment outcomes with reduced frequency of clinical encounters and antiretroviral treatment refills: A systematic review and meta-analysis

Noelle Le Tourneau  
*Washington University School of Medicine in St. Louis*

Ashley Germann  
*Johns Hopkins University*

Ryan R. Thompson  
*Johns Hopkins University*

Nathan Ford  
*World Health Organization*

Sheree Schwartz  
*Johns Hopkins University*

See next page for additional authors

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs

Please let us know how this document benefits you.

Recommended Citation

Le Tourneau, Noelle; Germann, Ashley; Thompson, Ryan R.; Ford, Nathan; Schwartz, Sheree; Beres, Laura; Mody, Aaloke; Baral, Stefan; Geng, Elvin H.; and Eshun-Wilson, Ingrid, "Evaluation of HIV treatment outcomes with reduced frequency of clinical encounters and antiretroviral treatment refills: A systematic review and meta-analysis." PLoS Medicine. 19, 3. e1003959 (2022).  
https://digitalcommons.wustl.edu/open_access_pubs/11773

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact vanam@wustl.edu.
Evaluation of HIV treatment outcomes with reduced frequency of clinical encounters and antiretroviral treatment refills: A systematic review and meta-analysis

Noelle Le Tourneau¹*, Ashley Germann², Ryan R. Thompson², Nathan Ford³,⁴, Sheree Schwartz², Laura Beres², Aaloke Mody¹, Stefan Baral⁵, Elvin H. Geng¹,⁵, Ingrid Eshun-Wilson¹

¹ Division of Infectious Diseases, School of Medicine, Washington University in St. Louis, Saint Louis, Missouri, United States of America, ² Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, United States of America, ³ Department of Global HIV, Hepatitis and Sexually Transmitted Diseases, World Health Organization, Geneva, Switzerland, ⁴ Centre for Infectious Disease Epidemiology and Research, School of Public Health and Family Medicine, University of Cape Town, Cape Town, South Africa, ⁵ Center for Dissemination and Implementation, Institute for Public Health, Washington University in St. Louis, Saint Louis, Missouri, United States of America

* letourneau@wustl.edu

Abstract

Background

Global HIV treatment programs have sought to lengthen the interval between clinical encounters for people living with HIV (PLWH) who are established on antiretroviral treatment (ART) to reduce the burden of seeking care and to decongest health facilities. The overall effect of reduced visit frequency on HIV treatment outcomes is however unknown. We conducted a systematic review and meta-analysis to evaluate the effect of implementation strategies that reduce the frequency of clinical appointments and ART refills for PLWH established on ART.

Methods and findings

We searched databases between 1 January 2010 and 9 November 2021 to identify randomized controlled trials (RCTs) and observational studies that compared reduced (6- to 12-monthly) clinical consultation or ART refill appointment frequency to 3- to 6-monthly appointments for patients established on ART. We assessed methodological quality and real-world relevance, and used Mantel–Haenszel methods to generate pooled risk ratios (RRs) with 95% confidence intervals for retention, viral suppression, and mortality. We evaluated heterogeneity quantitatively and qualitatively, and overall evidence certainty using GRADE. Searches yielded 3,955 records, resulting in 10 studies (6 RCTs, 3 observational studies, and 1 study contributing observational and RCT data) representing 15 intervention arms with 33,599 adults (≥16 years) in 8 sub-Saharan African countries. Reduced frequency clinical consultations occurred at health facilities, while reduced frequency ART refills were delivered through facility or community pharmacies and adherence groups. Studies were highly...
pragmatic, except for some study settings and resources used in RCTs. Among studies comparing reduced clinical consultation frequency (6- or 12-monthly) to 3-monthly consultations, there appeared to be no difference in retention (RR 1.01, 95% CI 0.97–1.04, \( p = 0.682 \), 8 studies, low certainty), and this finding was consistent across 6- and 12-monthly consultation intervals and delivery strategies. Viral suppression effect estimates were markedly influenced by under-ascertainment of viral load outcomes in intervention arms, resulting in inconclusive evidence. There was similarly insufficient evidence to draw conclusions on mortality (RR 1.12, 95% CI 0.75–1.66, \( p = 0.592 \), 6 studies, very low certainty). For ART refill frequency, there appeared to be little to no difference in retention (RR 1.01, 95% CI 0.98–1.06, \( p = 0.473 \), 4 RCTs, moderate certainty) or mortality (RR 1.45, 95% CI 0.63–3.35, \( p = 0.382 \), 4 RCTs, low certainty) between 6-monthly and 3-monthly visits. Similar to the analysis for clinical consultations, although viral suppression appeared to be better in 3-monthly arms, effect estimates were markedly influenced by under-ascertainment of viral load outcomes in intervention arms, resulting in overall inclusive evidence. This systematic review was limited by the small number of studies available to compare 12- versus 6-monthly clinical consultations, insufficient data to compare implementation strategies, and lack of evidence for children, key populations, and low- and middle-income countries outside of sub-Saharan Africa.

Conclusions

Based on this synthesis, extending clinical consultation intervals to 6 or 12 months and ART dispensing intervals to 6 months appears to result in similar retention to 3-month intervals, with less robust conclusions for viral suppression and mortality. Future research should ensure complete viral load outcome ascertainment, as well as explore mechanisms of effect, outcomes in other populations, and optimum delivery and monitoring strategies to ensure widespread applicability of reduced frequency visits across settings.

Author summary

Why was this study done?

- Global HIV services have moved toward differentiated service delivery (DSD) models that decrease the frequency of unneeded contact with the health system for patients established on antiretroviral treatment (ART) and ensure alignment between needs and services.

- It remains unclear, however, whether increasing ART clinical consultation intervals beyond 3 months, to 6 or 12 months, results in comparable treatment outcomes to 3-monthly visits.

What did the researchers do and find?

- We conducted a systematic review and meta-analysis to identify studies that compared reduced (6- or 12-monthly) clinical consultation or ART refill appointment frequency to 3- or 6-monthly appointments for patients established on ART.
• We identified 10 studies representing 15 intervention arms with 33,599 adults in 8 sub-Saharan African countries, including 6 randomized controlled trials, 3 cohort studies, and 1 study contributing both observational and randomized data. The methodological quality of the randomized controlled trials was generally high, and cohort data were of high, fair, and poor quality.

• Among 8 studies comparing reduced clinical consultation frequency (6- or 12-monthly) to 3-monthly consultations, there appeared to be no difference in retention in care, and this result was consistent across study design and visit frequency. There was poor viral load outcome ascertainment in reduced frequency study arms; as a result, when including all randomized study participants, viral suppression appeared higher in 3-monthly compared to reduced frequency arms for RCTs, but when analyses were restricted to only patients with viral load measurement, there appeared to be no difference in viral suppression.

• Among studies comparing 6-monthly to 3-monthly ART refill frequency, there similarly appeared to be no difference in retention in care or mortality. Viral suppression also appeared higher in 3-monthly compared to reduced frequency refill arms but showed no difference when analysis was restricted to patients with viral load measurement.

What do these findings mean?
• Extending clinical consultation intervals from 3 months to 6 or 12 months and ART dispensing intervals from 3 months to 6 months may result in similar outcomes for retention in care, with less robust conclusions for viral suppression and mortality.

• Future studies should aim to obtain more complete viral load outcome ascertainment in all study arms.

• To increase the applicability of these findings, further implementation research should explore other population groups, mechanisms of effect, and optimum delivery and monitoring strategies.

Introduction
For people living with HIV (PLWH) in low- and middle-income countries (LMICs) and on antiretroviral treatment (ART), unnecessary clinic and pharmacy appointments impose an avoidable burden on both patients and providers. Efforts to decrease the frequency of unneeded contact with the health system represent a central pillar of the movement toward differentiated service delivery (DSD) models [1,2]. There is, however, concern that reduced frequency might also compromise meaningful clinical and psychosocial interactions with the health system and potentially, paradoxically, increase missed appointments or non-adherence over the long term. Most existing data support the hypothesis that longer intervals reduce obstacles to attending services for PLWH, including structural challenges such as the time and travel required to attend appointments and pick up medications—which result in considerable direct and indirect costs—as well as psychosocial barriers such as stigma [3–7].
Based on these early observations, and consistent with World Health Organization (WHO) recommendations issued in 2016, global HIV programs have moved away from monthly appointments (which were the norm for many years) to 3-monthly appointments. Questions have turned to whether extending appointment intervals beyond 3 months to 6 or 12 months may be safe and effective. Early data syntheses suggesting that outcomes are equivalent for 3-versus 1-monthly appointments [8,9] do not automatically apply to further extensions. Qualitative, survey, and preference data indicate that PLWH value DSD models that include lengthened appointment intervals [4,10,11], but also that some PLWH do prefer more frequent psychosocial support from interactions with their providers [3]. Implementation of reduced visit frequency has accelerated pace in the past 2 years as it has become a practical necessity during the COVID-19 pandemic to decongest health facilities and limit physical contact [12,13]. Understanding how visit frequency impacts HIV treatment outcomes and under what conditions differences in outcomes manifest could further inform the implementation of this DSD strategy.

To support the 2021 update of the WHO service delivery guidelines [14], we undertook a systematic review and meta-analysis, including observational studies and randomized controlled trials (RCTs), in which we explicitly examined outcomes for clinical and ART refill visit intervals of 6 months or greater compared to 3 months. We present the data that contributed to the WHO guideline update as well as recently published studies. Such syntheses regarding further reductions in appointment frequency not only inform normative guidance on evolving DSD approaches but have particular salience for healthcare in the era of a pandemic.

**Methods**

The protocol is registered in PROSPERO (CRD42019128609).

**Eligibility criteria**

We included individual and cluster RCTs, comparative observational studies, cross-sectional studies, and single arm intervention studies without a comparison group. We included studies that enrolled PLWH established on first-line ART in LMICs. The definition of being established on ART varied by study (see Results). We included studies that reported outcomes of retention in care, viral suppression, and mortality. Eligible interventions included those with any component of less frequent clinical consultation or less frequent ART dispensing visits (e.g., 6- or 12-monthly) compared to 3- or 6-month frequency intervals (Table 1). Studies reporting 2-month comparison intervals were classified as 3-monthly for the purposes of this review.

**Search strategy and selection criteria**

We searched MEDLINE (PubMed), Embase (OVID), Cochrane Central Register of Controlled Trials, WHO International Clinical Trials Registry Platform (ICTRP), and ClinicalTrials.gov

Table 1. Eligibility criteria for included studies.

<table>
<thead>
<tr>
<th>PICO criterion</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>People living with HIV established on first-line ART as defined by study, in low- and middle-income countries</td>
</tr>
<tr>
<td>Intervention</td>
<td>Less frequent clinical or ART dispensing appointments (e.g., 6 or 12 months)</td>
</tr>
<tr>
<td>Comparison</td>
<td>3- or 6-monthly clinical or ART dispensing appointments</td>
</tr>
<tr>
<td>Outcome</td>
<td>Retention in care as defined by study; viral suppression as defined by study; mortality</td>
</tr>
</tbody>
</table>

ART, antiretroviral treatment; PICO, population, intervention, comparison, and outcome.

https://doi.org/10.1371/journal.pmed.1003959.t001
from 1 January 2010 through 9 November 2021, as well as Conference on Retroviruses and Opportunistic Infections (CROI) from 2017 to 2021 and International AIDS Society (IAS) conferences from 2016 to 2021. We additionally reviewed references, consulted experts in the field, and reviewed the IAS DSD resources [15].

Data extraction and quality assessment
Abstracts and titles were screened by 2 authors (NL, AG, RRT) in duplicate in Covidence [16], with any discrepancies resolved by a third author (IEW). Data on the study setting and population, intervention, and outcomes of eligible studies were extracted into an online database platform Airtable (https://airtable.com), with quality assurance of data done by a second author (NL or AG). Outcomes were extracted with numerators and denominators, as well as measures of association when possible. Reporting of study outcome quality was also extracted in Airtable. We assessed risk of bias using the Cochrane risk of bias tool (RoB-1) for RCTs and additionally judged risk of bias for each outcome as “low risk,” “high risk,” or “some concerns” [17,18]. We applied the Newcastle–Ottawa Scale to observational studies, with studies categorized with regard to risk of bias as “good quality,” “poor quality,” or “fair quality” [19]. We evaluated heterogeneity qualitatively and quantitatively through stratified analysis, and used GRADE to evaluate overall evidence certainty. We used PRECIS-2 criteria to assess how pragmatic or explanatory included studies were.

Intervention categorization
Increased spacing of clinical assessments and ART refills was frequently a component of broader DSD interventions. We therefore characterized interventions according to the frequency of clinical assessment, the location of clinical assessment, the health worker providing the clinical assessment, the frequency of ART refills, the location of refills, the delivery method of refills, and who was providing refills. We conducted separate analyses to evaluate outcomes associated with (1) reduced clinical appointment frequency and (2) reduced ART refill dispensing frequency.

Outcome definition
The primary outcome was retention in care, defined as the proportion of individuals retained on ART and in care at last available follow-up. Secondary outcomes were documented viral suppression at last available follow-up, and mortality. The viral load threshold for defining viral suppression was determined by the authors of each study reporting viral suppression. Planned secondary outcomes of adherence and morbidity were not assessed due to limited reporting in included studies.

Statistical analysis
We conducted pairwise meta-analysis comparing (1) reduced clinical assessment frequency with either 3-monthly or 6-monthly clinical assessments and (2) reduced ART refill dispensing frequency with 3-monthly ART refill frequency. For studies with more than 1 treatment arm, we split the comparison arm if both treatment arms were included in the pooled estimate. For analysis, we included numerators and denominators reported from individual studies and cluster-adjusted estimates for cluster RCTs based on intraclass correlation coefficients (ICCs) from the literature [20–23], according to methodology outlined in Cochrane guidelines [17], to generate overall risk ratios (RRs) with 95% confidence intervals (CIs) by study design. When adjusted measures of associations were available, we pooled relative data for studies reporting
time-to-event data as hazard ratios (HRs). Data were synthesized using R programming (packages “metafor” and “metabin”) using Mantel–Haenszel methods for pooling and random effects. Subgroup analyses with pooled RRs were performed where appropriate, including for different frequencies of refills or clinical consultations and different delivery strategies.

Results

Search and screening results

Searches yielded 3,955 records after deduplication; these 3,955 records underwent title and abstract screening. In total, 207 full-text articles were assessed for eligibility, and 20 records representing 10 studies with 15 intervention study arms met the criteria for inclusion in our review (Fig 1).

Included studies

The 10 eligible studies including 15 intervention arms included 6 RCTs, 3 comparative observational studies, and 1 study that contributed both observational and randomized data. All studies were conducted in sub-Saharan Africa, with a total of 33,599 participants across all studies. Two studies included PLWH aged ≥16 years, the rest included PLWH aged ≥18 years. We detected substantial clinical heterogeneity, including variability in the intervention components, the contexts in which reduced visit frequency was delivered, and the interval lengths between clinical consultations and ART refills. Eight intervention study arms with 6-monthly clinical consultations and 7 intervention study arms with 12-monthly clinical consultations were included. Four studies had intervention study arms with 6-monthly ART dispensing intervals. Characterization of the intervention strategies to reduce facility contact included 3 studies with ART dispensation at the clinic, 6 studies with ART dispensation in the community, and 4 studies with ART dispensation in the community or clinic. The majority of the studies (n = 8) included an “adherence club” component as part of the reduced appointment frequency strategy, where PLWH met in either a community or clinic setting. Other interventions included ART dispensation in private pharmacies and venues in the community and ART dispensation at home visits. Individual study and intervention details are summarized in Tables 2 and 3. Ten studies contributed to meta-analyses of reduced clinical consultation frequency (with 15 intervention study arms in total), and 4 studies contributed to meta-analyses of reduced ART refill frequency (4 intervention study arms). There was substantial methodological heterogeneity with regard to study design, risk of bias, and outcome assessment measures.

Risk of bias

As assessed by risk of bias tools, data from RCTs were generally judged as having high methodological quality (low risk of bias) or some concerns, and data from the 3 observational studies were judged as having high quality, fair quality, and poor quality (Tables 4 and S3). Data from cohort studies were considered fair or poor quality primarily due to the comparison arm comprising a patient population with different eligibility than the intervention arm (e.g., different levels of “stability”) [32,37]. RCTs were judged as having some concerns when those enrolled did not meet eligibility criteria [25], those eligible for the intervention did not necessarily receive it [20], or there was high withdrawal in the intervention arm [31].

PRECIS-2 score

Overall, studies were highly pragmatic as they were conducted in real-world settings with few additional measures to guarantee adherence to ART beyond what would occur in routine
RCTs were on average less pragmatic than cohort studies. Studies were downgraded when they required extensive expertise or organization to deliver the intervention of reduced frequency clinical assessments or reduced frequency ART refills, such as in the case of reduced frequency clinical assessments or reduced frequency ART refills, such as in the case of...
of home visits or group adherence clubs. Studies were less pragmatic when the intervention was delivered at clinical sites associated with research or at a single site.

Reduced clinical consultation frequency

Retention in care. We identified 9 comparative studies including 2 cohort studies, 6 randomized trials, and 1 study contributing both observational and RCT data (contributing 14 intervention arms in total) that reported retention in care and were included in the pairwise meta-analysis. Retention outcomes were reported by electronic health records and/or chart review, with the definition of retention in care provided by the authors (S1 Table). Among 8 studies with 13 total intervention arms comparing reduced clinic consultation frequency (>3-monthly) to 3-monthly clinical consultations, there appeared to be no difference in retention among all randomized individuals (RR 1.01, 95% CI 0.97–1.04, p = 0.682), consistent in both RCTs and observational data (RR 1.00, 95% CI 0.95–1.04, p = 0.917, and RR 1.02, 95% CI 0.97–1.09, p = 0.434, respectively) (Fig 2A). These findings were also consistent when stratified by 6- or 12-monthly clinical consultations (RR 1.03, 95% CI 0.98–1.08, p = 0.313, and RR 0.99, 95% CI 0.94–1.04, p = 0.672, respectively) (Fig 2B) and delivery strategy (S1 Fig). There was substantial statistical heterogeneity in study design and clinical consultation frequency subgroups, in part due to the inclusion of 1 study with high withdrawal from the intervention...
Among studies rated as high quality or having low risk of bias, heterogeneity remained substantial (S2 Fig). In exploration of the heterogeneity of study-specific definitions of established-on-ART patient populations, there remained substantial heterogeneity.

### Table 3. Summary of intervention characteristics.

<table>
<thead>
<tr>
<th>Study†</th>
<th>Reduced frequency group</th>
<th>Clinical assessment appointment</th>
<th>ART dispensing visit</th>
<th>Extra support</th>
<th>Facility-based comparator group(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Frequency</td>
<td>Location</td>
<td>Clinician</td>
<td>Frequency</td>
</tr>
<tr>
<td>Cassidy 2020 (i)</td>
<td>12 mo</td>
<td>Clinic</td>
<td>ND</td>
<td>6 mo</td>
<td>Community or clinic</td>
</tr>
<tr>
<td>Cassidy 2020 (ii)</td>
<td>~6 mo</td>
<td>Clinic</td>
<td>ND</td>
<td>~2.5 mo (5/ year)</td>
<td>Community or clinic</td>
</tr>
<tr>
<td>Fatti 2020a</td>
<td>12 mo</td>
<td>Clinic</td>
<td>ND</td>
<td>3 mo</td>
<td>Community</td>
</tr>
<tr>
<td>Fatti 2020b</td>
<td>12 mo</td>
<td>Clinic</td>
<td>ND</td>
<td>6 mo</td>
<td>Community</td>
</tr>
<tr>
<td>Fox 2019a</td>
<td>6 mo</td>
<td>Clinic</td>
<td>ND</td>
<td>2–3 mo</td>
<td>Community or clinic</td>
</tr>
<tr>
<td>Fox 2019b</td>
<td>6 mo</td>
<td>Clinic</td>
<td>ND</td>
<td>ND</td>
<td>Community</td>
</tr>
<tr>
<td>Goodrich 2021</td>
<td>12 mo</td>
<td>Clinic</td>
<td>ND</td>
<td>3 mo</td>
<td>Community</td>
</tr>
<tr>
<td>Grimsrud 2016</td>
<td>12 mo</td>
<td>Community</td>
<td>Nurse</td>
<td>~2.5 mo (5/ year)</td>
<td>Community or clinic</td>
</tr>
<tr>
<td>Hoffman 2021 (i)</td>
<td>6 mo</td>
<td>Clinic</td>
<td>Provider</td>
<td>6 mo</td>
<td>Clinic</td>
</tr>
<tr>
<td>Hoffman 2021 (ii)</td>
<td>3 mo</td>
<td>Clinic</td>
<td>Provider</td>
<td>3 mo</td>
<td>Clinic</td>
</tr>
<tr>
<td>Nichols 2021a</td>
<td>6 mo</td>
<td>Clinic</td>
<td>ND</td>
<td>1 mo</td>
<td>Community</td>
</tr>
<tr>
<td>Nichols 2021b</td>
<td>6 mo</td>
<td>Clinic</td>
<td>ND</td>
<td>2–3 mo</td>
<td>Clinic</td>
</tr>
<tr>
<td>Nichols 2021c</td>
<td>6 mo</td>
<td>Clinic</td>
<td>ND</td>
<td>1–3 mo</td>
<td>Home</td>
</tr>
<tr>
<td>Pasipamire 2018 (i)</td>
<td>6 mo</td>
<td>Clinic</td>
<td>ND</td>
<td>1 mo</td>
<td>Community</td>
</tr>
<tr>
<td>Pasipamire 2018 (ii)</td>
<td>3 mo</td>
<td>Clinic</td>
<td>ND</td>
<td>3 mo</td>
<td>Clinic</td>
</tr>
<tr>
<td>Pasipamire 2018 (iii)</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>Community</td>
</tr>
<tr>
<td>Tukei 2020a</td>
<td>12 mo</td>
<td>Clinic</td>
<td>ND</td>
<td>3 mo</td>
<td>Community</td>
</tr>
<tr>
<td>Tukei 2020b</td>
<td>12 mo</td>
<td>Clinic</td>
<td>ND</td>
<td>6 mo</td>
<td>Community</td>
</tr>
<tr>
<td>Woodd 2014</td>
<td>6 mo</td>
<td>Clinic</td>
<td>Clinician</td>
<td>1 mo</td>
<td>Home</td>
</tr>
</tbody>
</table>

AC, adherence club; ART, antiretroviral treatment; CHW, community health worker; CO, clinical officer; HCW, healthcare worker; mo, monthly; ND, not described; pharm, pharmacy; SOC, standard of care.

†For studies with active comparator arms that differ from a facility-based comparator group: (i) reduced frequency arm; (ii, iii) active comparator arms.

https://doi.org/10.1371/journal.pmed.1003959.t003

[community-based care] arm [31]. Among studies rated as high quality or having low risk of bias, heterogeneity remained substantial (S2 Fig). In exploration of the heterogeneity of study-specific definitions of established-on-ART patient populations, there remained substantial...
### Table 4. Summary of risk of bias assessment.

<table>
<thead>
<tr>
<th>Study</th>
<th>Retention in care</th>
<th>Viral suppression among randomized</th>
<th>Viral suppression among analyzed</th>
<th>Mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cassidy 2020</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Some concerns</td>
</tr>
<tr>
<td>Fatti 2020a</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
</tr>
<tr>
<td>Fatti 2020b</td>
<td>Low risk</td>
<td>High risk</td>
<td>Low risk</td>
<td>Low risk</td>
</tr>
<tr>
<td>Fox 2019a</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Not reported</td>
</tr>
<tr>
<td>Fox 2019b**</td>
<td>High quality</td>
<td>High quality</td>
<td>High quality</td>
<td>Not reported</td>
</tr>
<tr>
<td>Goodrich 2021*</td>
<td>Some concerns</td>
<td>High risk</td>
<td>Some concerns</td>
<td>High risk</td>
</tr>
<tr>
<td>Grimsrud 2016*</td>
<td>Fair quality</td>
<td>Fair quality</td>
<td>Fair quality</td>
<td>Not reported</td>
</tr>
<tr>
<td>Hoffman 2021*</td>
<td>Low risk</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Low risk</td>
</tr>
<tr>
<td>Nichols 2021a*</td>
<td>High quality</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>Nichols 2021b*</td>
<td>High quality</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>Nichols 2021c*</td>
<td>High quality</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>Pasipamire 2018**</td>
<td>Poor quality</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Poor quality</td>
</tr>
<tr>
<td>Tukei 2020a</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
</tr>
<tr>
<td>Tukei 2020b*</td>
<td>Low risk</td>
<td>High risk</td>
<td>Low risk</td>
<td>Low risk</td>
</tr>
<tr>
<td>Woodd 2014*</td>
<td>Low risk</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Low risk</td>
</tr>
</tbody>
</table>

*Assessments based on Cochrane risk of bias tool (RoB-1) for randomized controlled trials: high risk (red), some concerns (yellow), or low risk (green).

**Assessments based on Newcastle–Ottawa Scale for cohort studies: poor quality (red), fair quality (yellow), or high quality (green).

https://doi.org/10.1371/journal.pmed.1003959.t004

### Table 5. Summary of PRECIS-2 score.

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design</th>
<th>Eligibility</th>
<th>Recruitment</th>
<th>Setting</th>
<th>Organization</th>
<th>Flexibility: Delivery</th>
<th>Flexibility: Adherence</th>
<th>Follow-up</th>
<th>Primary outcome</th>
<th>Primary analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cassidy 2020</td>
<td>RCT (cluster)</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Fatti 2020</td>
<td>RCT (cluster)</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Fox 2019</td>
<td>Cohort</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Goodrich 2021*</td>
<td>RCT (cluster)</td>
<td>5</td>
<td></td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Grimsrud 2016*</td>
<td>Cohort</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Hoffman 2021</td>
<td>RCT (cluster)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Nichols 2021*</td>
<td>Cohort</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Pasipamire 2018*</td>
<td>Cohort</td>
<td>5</td>
<td></td>
<td>3</td>
<td>4</td>
<td></td>
<td>5</td>
<td></td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Tukei 2020*</td>
<td>RCT (cluster)</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Woodd 2014*</td>
<td>RCT (cluster)</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>

RCT, randomized controlled trial. A value of 5 (dark green) represents a very pragmatic approach, and a value of 1 (yellow) represents a very explanatory approach.

https://doi.org/10.1371/journal.pmed.1003959.t005
Fig 2. Retention in care: Reduced frequency versus 3-monthly clinical consultations among randomized or enrolled individuals. (A) By study design. (B) By clinical consultation frequency. Goodrich 2021 had high withdrawal from the intervention (community-based care) arm; it was unclear if those who withdrew reengaged in care at the health facility. Fox 2019a and Fox 2019b were separated into an RCT and cohort design, respectively, based on the analysis described by the authors, where randomization was not preserved in the intervention arm in Fox 2019b. *Cluster-adjusted RR. 3 MO, 3-monthly; AC, adherence club; Comm, community; e, number of events; n, number of participants; ND, not described; P, pharmacy; RCT, randomized controlled trial; RF, reduced frequency; RR, risk ratio.

https://doi.org/10.1371/journal.pmed.1003959.g002

statistical heterogeneity across subgroups of required time spent on ART (12 months, 6 months, or other) for eligibility (S3 Fig).
Only 1 study compared 12-monthly clinical consultations to 6-monthly consultations [25]; this study reported similar retention in care at 24 months for 12- and 6-monthly clinical consultations (RR 0.99, 95% CI 0.96–1.01, p = 0.363) (Fig 3).

**Viral suppression.** We identified 6 comparative studies—one cohort study, 4 randomized trials, and 1 study contributing both observational and RCT data (contributing 9 total arms for comparison)—that reported viral suppression for inclusion in the pairwise meta-analysis. Viral suppression outcomes were reported by electronic health records and/or chart review using variable thresholds (<400 copies/ml and <1,000 copies/ml) (S1 Table). RCT meta-analysis suggested decreased viral suppression for reduced frequency of clinical consultations compared to 3-monthly clinical consultations (RR 0.74, 95% CI 0.59–0.94, p = 0.015), while cohort studies showed slightly greater viral suppression for reduced frequency of clinical consultations compared to 3-monthly clinical consultations (RR 1.40, 95% CI 0.95–2.08, p = 0.093) (Fig 4A). Within RCT and cohort subgroup analyses, substantial statistical heterogeneity persisted (Fig 4A), markedly influenced by 1 RCT where only 7.3% of the reduced frequency arm received viral load testing and 1 cohort study with substantially higher viral suppression among those in the reduced frequency arm (there were substantial baseline imbalances between study arms in this study, with those receiving reduced visit frequency on ART for longer periods than those in the 3-monthly arm) [27,32] (S2 Table). As estimates differed by study design, overall estimates were not pooled across RCTs and cohort studies in subgroup analyses for clinical consultation frequency (6- or 12-monthly) and delivery strategy (S4 and S6 Figs).

In an available case analysis (including only those who received viral load testing), there appeared to be similar viral suppression in the arms for reduced frequency and 3-monthly clinical consultations among RCTs (RR 1.00, 95% CI 0.92–1.08, p = 0.916), and possible improved viral suppression for reduced frequency clinical consultations compared to 3-monthly clinical consultations among cohort studies (RR 1.44, 95% CI 1.24–1.66, p < 0.001) (Fig 4B). Due to differences in the pooled estimates in RCTs and cohorts, estimates were not pooled across study designs in subgroup analyses for clinical consultation frequency (6- or 12-monthly) (S5 and S7 Figs).

In the single study comparing 12-monthly to 6-monthly clinical consultations, among all individuals randomized, viral suppression was higher in the arm with 12-monthly versus 6-monthly clinical consultations (RR 1.06, 95% CI 1.02–1.10, p = 0.004) (Fig 5A). Among those who received viral load testing, there was no difference in viral suppression for 12-monthly compared to 6-monthly clinical consultations (RR 0.99, 95% CI 0.97–1.01, p = 0.391) (Fig 5B).

**Mortality.** Seven comparative studies (9 comparisons), including 1 cohort study and 6 RCTs, contributed to the mortality meta-analysis. There was no evidence of a difference in mortality between reduced clinical consultations and 3-monthly consultations among these studies (overall RR 1.12, 95% CI 0.75–1.66, p = 0.592) (Fig 6A). This was consistent for 6-monthly

---

**Fig 3. Retention in care: Reduced frequency versus 6-monthly clinical consultations.** Cluster-adjusted RR. 6 MO, 6-monthly; AC, adherence club; Comm, community; e, number of events; n, number of participants; RF, reduced frequency; RR, risk ratio.

https://doi.org/10.1371/journal.pmed.1003959.g003
Evaluation of reduced visit frequency and HIV outcomes

(A) Among those randomized or enrolled.
(B) Among those with viral load testing. Fox 2019a and Fox 2019b were separated into an RCT and cohort design, respectively, based on the analysis described by the authors, where randomization was not preserved in the intervention arm in Fox 2019b. *Cluster-adjusted RR. 3 MO, 3-monthly; AC, adherence club; Comm, community; e, number of events; n, number of participants; ND, not described; P, pharmacy; RCT, randomized controlled trial; RF, reduced frequency; RR, risk ratio.

https://doi.org/10.1371/journal.pmed.1003959.g004

Reduced ART refill dispensing frequency

Retention in care. Among the 4 studies (all cluster RCTs) investigating reduced ART refill frequency, there appeared to be no difference in retention in care between increased (6 month) refill intervals and 3-monthly refill frequency (RR 1.01, 95% CI 0.98–1.06, p = 0.473) (Fig 8). No studies assessed ART refills at intervals greater than 6 months.
Viral suppression.

Among all PLWH enrolled in the 3 RCTs comparing reduced (6-monthly) ART dispensing frequency to 3-monthly dispensing, viral suppression appeared better among those in the 3-monthly dispensing arms (RR 0.60, 95% CI 0.41–0.88, \( p = 0.009 \)) (Fig 9A). This comparison had substantial statistical heterogeneity, influenced by 2 studies with under-ascertainment of viral load in intervention arms [27,38] (S2 Table). The available case analysis, including only PLWH who received viral load testing, showed no difference between intervention arms for 6-monthly refills compared to 3-monthly refills (RR 0.99, 95% CI 0.98–1.00, \( p = 0.235 \)) (Fig 9B).

Mortality.

In the 4 studies (all RCTs) comparing 6-monthly refill frequency to 3-monthly refills, there was no evidence of a difference in mortality between reduced (6-monthly) ART dispensing frequency and 3-monthly refills (RR 1.45, 95% CI 0.63–3.35, \( p = 0.382 \)) (Fig 10).

Certainty of evidence (GRADE)

The certainty of the evidence (a combined assessment of strength of association, methodological quality, heterogeneity, and external validity) for the pooled data for the primary outcomes of retention, viral suppression, and mortality was assessed as very low to moderate quality (Tables 6 and 7). Effect estimates were downgraded due to high risk of bias in the contributing studies and heterogeneity in the contributing effect estimates. For clinical visit frequency comparisons, the certainty of evidence for retention in care was ranked overall as moderate, and viral suppression and mortality were rated as having very low quality evidence, largely due to imprecision and/or high risk of bias for contributing studies. Regarding the overall estimates for ART refill dispensing frequency comparisons, the outcome of retention in care was ranked as having moderate certainty, while viral suppression was ranked as having very low certainty, and mortality was ranked as having low certainty, also due to high risk of bias and imprecision.

These evidence rankings contribute to statements and assumptions that can be made about the evidence contributing to this review. The very low quality evidence for viral load and...
mortality lead to final assessments of insufficient evidence to draw meaningful conclusions for these outcomes.

Discussion

In this systematic review we found among the 10 included studies (6 RCTs, 3 observational studies, and 1 study contributing both observational and RCT data)—with 15 study arms with 33,599 adults in 8 countries in sub-Saharan Africa—that reduced frequency of clinical consultations and ART dispensing appeared to have comparable HIV treatment outcomes to 3-monthly clinical or dispensing visits. For reduced frequency clinical consultations, there was no evidence of a difference in retention in care, when comparing...
reduced frequency (i.e., 6- or 12-monthly) clinical consultations to 3-monthly consultation visits. For clinical consultations, viral load results were inconsistent, and it was not possible to discern the effect of reduced clinical consultation frequency on viral suppression due to marked under-ascertainment of viral load in reduced frequency intervention arms. Similarly, conclusions could not be drawn on the effects on mortality, due to the overall small number of events and very low quality evidence. A single study that compared 12-monthly to 6-monthly clinical consultations showed similar retention in care and viral suppression between study arms. When comparing 6-monthly to 3-monthly ART dispensing frequency, there appeared to be little to no difference in retention in care. For ART refill frequency, evidence quality ratings for viral suppression and mortality were similarly very low; it was therefore not possible to draw conclusions for these outcomes.

Visit frequency was reduced through a variety of implementation strategies: In most cases clinical consultations occurred at the health facility, and ART dispensing was facilitated through adherence clubs at the health facility or in the community, with individual club members, lay staff, or nurses distributing ART. Other community ART delivery strategies included distribution at community venues, private pharmacies, or mobile health units, or directly in the homes of PLWH, though there were relatively few studies to compare across delivery strategies. Overall, included studies were highly pragmatic. There was, however, marked heterogeneity of effects, study designs, risk of bias, implementation strategies, and outcome measurement time points—this contributed to the low-certainty evidence ratings for several outcomes. The definition of the established-on-ART patient population varied by study; however, no studies included data on children or key population groups, or were from outside of the sub-Saharan African region.

Fig 7. Mortality: Reduced frequency versus 6-monthly clinical consultations. *Cluster-adjusted RR. 6 MO, 6-monthly; AC, adherence club; Comm, community; e, number of events; n, number of participants; RF, reduced frequency; RR, risk ratio.

https://doi.org/10.1371/journal.pmed.1003959.g007

reduced frequency (i.e., 6- or 12-monthly) clinical consultations to 3-monthly consultation visits. For clinical consultations, viral load results were inconsistent, and it was not possible to discern the effect of reduced clinical consultation frequency on viral suppression due to marked under-ascertainment of viral load in reduced frequency intervention arms. Similarly, conclusions could not be drawn on the effects on mortality, due to the overall small number of events and very low quality evidence. A single study that compared 12-monthly to 6-monthly clinical consultations showed similar retention in care and viral suppression between study arms. When comparing 6-monthly to 3-monthly ART dispensing frequency, there appeared to be little to no difference in retention in care. For ART refill frequency, evidence quality ratings for viral suppression and mortality were similarly very low; it was therefore not possible to draw conclusions for these outcomes.

Visit frequency was reduced through a variety of implementation strategies: In most cases clinical consultations occurred at the health facility, and ART dispensing was facilitated through adherence clubs at the health facility or in the community, with individual club members, lay staff, or nurses distributing ART. Other community ART delivery strategies included distribution at community venues, private pharmacies, or mobile health units, or directly in the homes of PLWH, though there were relatively few studies to compare across delivery strategies. Overall, included studies were highly pragmatic. There was, however, marked heterogeneity of effects, study designs, risk of bias, implementation strategies, and outcome measurement time points—this contributed to the low-certainty evidence ratings for several outcomes. The definition of the established-on-ART patient population varied by study; however, no studies included data on children or key population groups, or were from outside of the sub-Saharan African region.

Fig 7. Mortality: Reduced frequency versus 6-monthly clinical consultations. *Cluster-adjusted RR. 6 MO, 6-monthly; AC, adherence club; Comm, community; e, number of events; n, number of participants; RF, reduced frequency; RR, risk ratio.

https://doi.org/10.1371/journal.pmed.1003959.g007
While we found overall little difference in clinical outcomes for reduced visit frequency, there are other potential benefits of reduced visits, including decongestion of health facilities, reduced provider workload, prioritization of care for new or clinically unstable PLWH, and reduced transmission of COVID-19 in health centers [5,13,41–44]. Reducing visit frequency has been reported to be one of the easiest DSD models to implement and aligns strongly with the care preferences of PLWH by reducing the economic costs of attending frequent appointments, reducing stigma, and allowing PLWH to normalize HIV [3,4,10,11,45,46]. HIV services, however, need to remain flexible enough to accommodate return to facilities for those who opt back into standard care or when clinical requirements change [27]. Further research is needed to develop strategies that allow for transition between models of care and provide psychosocial support between extended visits (e.g., virtual visits or group models) [3,47]. As many countries, in response to COVID-19, have expanded multi-month dispensing for patients who have not previously been considered established on ART [12], it will be essential

![Fig 9. Viral suppression: Reduced frequency versus 3-monthly antiviral treatment refills, by study design.](https://doi.org/10.1371/journal.pmed.1003959.g009)

(A) Among those randomized. (B) Among those with viral load testing. ∗Cluster-adjusted RR. 3 MO, 3-monthly; AC, adherence club; Comm, community; e, number of events; n, number of participants; P, pharmacy; RF, reduced frequency; RR, risk ratio.

While we found overall little difference in clinical outcomes for reduced visit frequency, there are other potential benefits of reduced visits, including decongestion of health facilities, reduced provider workload, prioritization of care for new or clinically unstable PLWH, and reduced transmission of COVID-19 in health centers [5,13,41–44]. Reducing visit frequency has been reported to be one of the easiest DSD models to implement and aligns strongly with the care preferences of PLWH by reducing the economic costs of attending frequent appointments, reducing stigma, and allowing PLWH to normalize HIV [3,4,10,11,45,46]. HIV services, however, need to remain flexible enough to accommodate return to facilities for those who opt back into standard care or when clinical requirements change [27]. Further research is needed to develop strategies that allow for transition between models of care and provide psychosocial support between extended visits (e.g., virtual visits or group models) [3,47]. As many countries, in response to COVID-19, have expanded multi-month dispensing for patients who have not previously been considered established on ART [12], it will be essential

![Fig 10. Mortality: Reduced frequency versus 3-monthly refills.](https://doi.org/10.1371/journal.pmed.1003959.g010)

While we found overall little difference in clinical outcomes for reduced visit frequency, there are other potential benefits of reduced visits, including decongestion of health facilities, reduced provider workload, prioritization of care for new or clinically unstable PLWH, and reduced transmission of COVID-19 in health centers [5,13,41–44]. Reducing visit frequency has been reported to be one of the easiest DSD models to implement and aligns strongly with the care preferences of PLWH by reducing the economic costs of attending frequent appointments, reducing stigma, and allowing PLWH to normalize HIV [3,4,10,11,45,46]. HIV services, however, need to remain flexible enough to accommodate return to facilities for those who opt back into standard care or when clinical requirements change [27]. Further research is needed to develop strategies that allow for transition between models of care and provide psychosocial support between extended visits (e.g., virtual visits or group models) [3,47]. As many countries, in response to COVID-19, have expanded multi-month dispensing for patients who have not previously been considered established on ART [12], it will be essential
Table 6. Review evidence certainty assessment (GRADE): Reduced clinical appointment frequency (6- or 12-monthly) versus 3-monthly clinical appointments.

<table>
<thead>
<tr>
<th>Number of studies</th>
<th>Certainty assessment</th>
<th>Risk of bias</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other considerations</th>
<th>Reduced clinic appointment frequency</th>
<th>3-monthly clinic appointments</th>
<th>Effect estimate: Risk ratio (95% CI)</th>
<th>Certainty</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Retention in care among all enrolled at longest time point: Any reduced frequency versus 3-monthly clinic appointment frequency</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 RCTs &amp; observational studies</td>
<td>SeriousΔ</td>
<td>SeriousΔ</td>
<td>Not serious</td>
<td>Not serious</td>
<td>None</td>
<td>11,526/12,638 (91.2%)</td>
<td>8,454/9,623 (87.9%)</td>
<td>1.01 (0.97 to 1.04)</td>
<td>⊗⊗⊗⊗ Low</td>
<td></td>
</tr>
<tr>
<td><strong>Retention in care among all enrolled at longest time point: RCTs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 RCTs</td>
<td>Not serious</td>
<td>SeriousΔ</td>
<td>Not serious</td>
<td>Not serious</td>
<td>None</td>
<td>9,892/10,759 (91.9%)</td>
<td>6,932/7,815 (88.7%)</td>
<td>1.00 (0.95 to 1.04)</td>
<td>⊗⊗⊗⊗ Moderate</td>
<td></td>
</tr>
<tr>
<td><strong>Retention in care among all enrolled at longest time point: Cohort studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Observational studies</td>
<td>SeriousΔ</td>
<td>Not serious</td>
<td>Not serious</td>
<td>Not serious</td>
<td>None</td>
<td>1,634/1,879 (87.0%)</td>
<td>1,522/1,808 (84.2%)</td>
<td>1.02 (0.97 to 1.09)</td>
<td>⊗⊗⊗⊗ Very low</td>
<td></td>
</tr>
<tr>
<td><strong>Retention in care among all enrolled at longest time point: 6-monthly versus 3-monthly clinic appointment frequency</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 RCTs &amp; observational studies</td>
<td>SeriousΔ</td>
<td>Not serious</td>
<td>Not serious</td>
<td>Not serious</td>
<td>None</td>
<td>2,602/3,013 (86.4%)</td>
<td>2,262/2,696 (83.9%)</td>
<td>1.03 (0.98 to 1.08)</td>
<td>⊗⊗⊗⊗ Moderate</td>
<td></td>
</tr>
<tr>
<td><strong>Retention in care among all enrolled at longest time point: 12-monthly versus 3-monthly clinic appointment frequency</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 RCTs</td>
<td>Not serious</td>
<td>SeriousΔ</td>
<td>Not serious</td>
<td>Not serious</td>
<td>None</td>
<td>8,924/9,625 (92.7%)</td>
<td>6,196/6,927 (89.3%)</td>
<td>0.99 (0.94 to 1.04)</td>
<td>⊗⊗⊗⊗ Moderate</td>
<td></td>
</tr>
<tr>
<td><strong>Viral suppression among all enrolled at longest time point: RCTs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 RCTs</td>
<td>Very seriousΔ</td>
<td>SeriousΔ</td>
<td>Not serious</td>
<td>Not serious</td>
<td>None</td>
<td>3,424/6,801 (50.3%)</td>
<td>2,777/4,325 (64.2%)</td>
<td>0.74 (0.59 to 0.94)</td>
<td>⊗⊗⊗⊗ Very low</td>
<td></td>
</tr>
<tr>
<td><strong>Viral suppression among all enrolled at longest time point: Cohort studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Observational studies</td>
<td>Not serious</td>
<td>SeriousΔ</td>
<td>Not serious</td>
<td>Not serious</td>
<td>None</td>
<td>1,941/2,345 (82.8%)</td>
<td>2,934/6,729 (43.6%)</td>
<td>1.40 (0.95 to 2.08)</td>
<td>⊗⊗⊗⊗ Very low</td>
<td></td>
</tr>
<tr>
<td><strong>Viral suppression among all who received viral load testing at longest time point: RCTs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 RCTs</td>
<td>Not serious</td>
<td>SeriousΔ</td>
<td>Not serious</td>
<td>Not serious</td>
<td>None</td>
<td>3,424/3,489 (98.1%)</td>
<td>2,777/2,826 (98.3%)</td>
<td>1.00 (0.92 to 1.08)</td>
<td>⊗⊗⊗⊗ Moderate</td>
<td></td>
</tr>
<tr>
<td><strong>Viral suppression among all who received viral load testing at longest time point: Cohort studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Observational studies</td>
<td>Not serious</td>
<td>SeriousΔ</td>
<td>Not serious</td>
<td>Not serious</td>
<td>None</td>
<td>1,941/2,295 (84.6%)</td>
<td>2,934/2,962 (99.1%)</td>
<td>1.44 (1.24 to 1.66)</td>
<td>⊗⊗⊗⊗ Very low</td>
<td></td>
</tr>
<tr>
<td><strong>Mortality among all enrolled at longest time point: Any reduced frequency versus 3-monthly clinic appointment frequency</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 RCTs &amp; observational studies</td>
<td>SeriousΔ</td>
<td>Not serious</td>
<td>Not serious</td>
<td>Very seriousΔ</td>
<td>None</td>
<td>166/11,015 (1.5%)</td>
<td>101/7,810 (1.3%)</td>
<td>1.12 (0.75 to 1.66)</td>
<td>⊗⊗⊗⊗ Very low</td>
<td></td>
</tr>
<tr>
<td><strong>Mortality among all enrolled at longest time point: RCTs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 RCTs</td>
<td>SeriousΔ</td>
<td>Not serious</td>
<td>Not serious</td>
<td>Very seriousΔ</td>
<td>None</td>
<td>163/10,484 (1.6%)</td>
<td>100/7,521 (1.3%)</td>
<td>1.10 (0.73 to 1.65)</td>
<td>⊗⊗⊗⊗ Low</td>
<td></td>
</tr>
<tr>
<td><strong>Mortality among all enrolled at longest time point: Cohort studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Observational studies</td>
<td>SeriousΔ</td>
<td>Not serious</td>
<td>Not serious</td>
<td>Very seriousΔ</td>
<td>None</td>
<td>3/531 (0.6%)</td>
<td>1/289 (0.3%)</td>
<td>1.63 (0.17 to 15.63)</td>
<td>⊗⊗⊗⊗ Very low</td>
<td></td>
</tr>
<tr>
<td><strong>Mortality among all enrolled at longest time point: 6-monthly versus 3-monthly clinic appointment frequency</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 RCTs &amp; observational studies</td>
<td>SeriousΔ</td>
<td>Not serious</td>
<td>Not serious</td>
<td>Very seriousΔ</td>
<td>None</td>
<td>120/1,390 (8.6%)</td>
<td>81/883 (9.2%)</td>
<td>1.05 (0.66 to 1.66)</td>
<td>⊗⊗⊗⊗ Very low</td>
<td></td>
</tr>
<tr>
<td><strong>Mortality among all enrolled at longest time point: 12-monthly versus 3-monthly clinic appointment frequency</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 RCTs</td>
<td>Not serious</td>
<td>Not serious</td>
<td>Not serious</td>
<td>Very seriousΔ</td>
<td>None</td>
<td>46/9,625 (0.5%)</td>
<td>20/6,927 (0.3%)</td>
<td>1.35 (0.61 to 2.98)</td>
<td>⊗⊗⊗⊗ Low</td>
<td></td>
</tr>
</tbody>
</table>

CI, confidence interval; RCT, randomized controlled trial. Explanations
ΔCombination of cohort and RCT data. In addition, contributing observational study was ranked as poor quality.
ΔStatistical heterogeneity.
ΔVariable study quality.
ΔStudies with high risk of bias and/or some concerns of bias contribute substantially to estimate.
ΔDowngraded due to very wide confidence intervals including benefit and harm.
ΔEstimate consists of only 1 study.

https://doi.org/10.1371/journal.pmed.1003959.t006
to explore outcomes in those less “established” on ART, as well as to develop strategies to align and integrate non-communicable disease care with these models and to identify optimum models of care for key populations, other regions, and children, to ensure the utility of these models for all PLWH [5,48].

Ongoing successful scale-up of multi-month scripting and sustainability will depend on well-functioning drug supply chains. To date, 3-monthly ART dispensing visits and 6-monthly clinical consultations have been widely adopted in LMICs, and the COVID-19 epidemic has accelerated the adoption of even longer intervals, with 11 countries providing 6-monthly ART refills and 6 countries providing 12-monthly clinical consults as of June 2021 [12,49–51]. The reliability of supply chains to maintain multi-month dispensing remains a concern, however, with drug stock-outs common, particularly in the sub-Saharan African region [4,52]. At this time, local drug supply chains and pharmacy capacity should be robust to ensure that PLWH do not experience barriers to obtaining at minimum 3-monthly refills [53,54].

In addition to ensuring adequate ART supply, incorporating well-functioning treatment monitoring strategies into differentiated models of care will be crucial. While there were too few studies within subgroups to compare outcomes by delivery strategy, reducing facility visits reduces opportunities for viral load measurement at centralized locations, and viral load monitoring, in particular, appeared to be a challenge in treatment arms providing primarily community-based services with infrequent facility visits. Strengthening facility-based laboratory systems as well as establishing reliable decentralized viral load monitoring strategies (e.g., point-of-care or community-based sample collection) represent further areas for investigation to support reduced clinical and ART dispensing visit frequency [55–57].

### Table 7. Review evidence certainty assessment (GRADE): Reduced ART dispensing (6-monthly) versus 3-monthly ART dispensing.

<table>
<thead>
<tr>
<th>Number of studies</th>
<th>Certainty assessment</th>
<th>Number of patients with outcome/number of patients total (%)</th>
<th>Effect estimate: Risk ratio (95% CI)</th>
<th>Certainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study design</td>
<td>Risk of bias</td>
<td>Inconsistency</td>
<td>Indirectness</td>
<td>Imprecision</td>
</tr>
<tr>
<td>Retention in care among all enrolled at longest time point</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>RCTs</td>
<td>Not serious&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Serious&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Not serious</td>
</tr>
<tr>
<td>Viral suppression among all enrolled at longest time point</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>RCTs</td>
<td>Very serious&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Serious&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Not serious</td>
</tr>
<tr>
<td>Viral suppression among all who received viral load testing at longest time point</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>RCTs</td>
<td>Serious&lt;sup&gt;d&lt;/sup&gt;</td>
<td>Not serious</td>
<td>Not serious</td>
</tr>
<tr>
<td>Mortality among all enrolled at longest time point</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>RCTs</td>
<td>Not serious&lt;sup&gt;e&lt;/sup&gt;</td>
<td>Not serious</td>
<td>Not serious</td>
</tr>
</tbody>
</table>

CI, confidence interval; RCT, randomized controlled trial.
<sup>a</sup>Four RCTs—3 with low risk of bias and 1 with some concerns.
<sup>b</sup>Marked statistical heterogeneity.
<sup>c</sup>Three RCTs—1 with some concerns and 2 with high risk of bias.
<sup>d</sup>Three RCTs—2 with low risk of bias and 1 with some concerns.
<sup>e</sup>Very few events and wide confidence intervals.

https://doi.org/10.1371/journal.pmed.1003959.t007
Limitations and strengths

This synthesis was strengthened by inclusion of a wide range of pragmatic trial data and programmatic observational data providing real-world insights into the effect of reducing dispensing and clinical visit intervals. There were, however, also several limitations of the data included in the review. First, we acknowledge that pooling heterogeneous studies cannot generate one true effect estimate relevant to all contexts; however, such syntheses can give insights into the broader question of whether an intervention results in benefit or harm, which was the overarching goal of this synthesis. Second, there was a lack of evidence for children, key populations, and LMICs outside of sub-Saharan Africa, limiting the generalizability of the findings. Third, few studies contributed to comparisons of 12- versus 6-monthly clinical consultations, and therefore no firm conclusions could be generated on the relative effect of such consultation intervals. Fourth, there were insufficient data to compare and stratify by reduced visit interval implementation strategies. Lastly, there were no data on how increasing visit intervals impacts management of other comorbid illnesses such as diabetes and hypertension.

Conclusion

Based on data from this synthesis, extending clinical consultation intervals beyond 3 months and ART dispensing intervals to 6 months likely results in similar retention in care compared to 3-monthly intervals, with uncertain effects on mortality and viral suppression. As countries shift toward 6-monthly clinical consultations and extended ART dispensation intervals, research should identify which delivery strategies are most efficient in accommodating both patient preferences and pragmatic concerns regarding cost and logistical health system capabilities. Ongoing monitoring of emerging evidence on the scale-up of reduced visit interval strategies will be critical to inform future HIV service delivery guidelines.

Supporting information

S1 Appendix. Deviations from protocol listed in PROSPERO.

S2 Appendix. Search terms.

S1 Fig. Retention in care: Reduced frequency versus 3-monthly clinical consultations, by delivery strategy. Fox 2019a and Fox 2019b were separated into an RCT and cohort design, respectively, based on the analysis described by the authors, where randomization was not preserved in the intervention arm in Fox 2019b. *Cluster-adjusted RR. 3 MO, 3-monthly; AC, adherence club; Comm, community; e, number of events; n, number of participants; ND, not described; P, pharmacy; RF, reduced frequency.

S2 Fig. Retention in care: Reduced frequency versus 3-monthly clinical consultations, by risk of bias assessment. Fox 2019a and Fox 2019b were separated into an RCT and cohort design, respectively, based on the analysis described by the authors, where randomization was not preserved in the intervention arm in Fox 2019b. *Cluster-adjusted RR. 3 MO, 3-monthly; AC, adherence club; Comm, community; e, number of events; n, number of participants; ND, not described; P, pharmacy; RF, reduced frequency.
S3 Fig. Retention in care: Reduced frequency versus 3-monthly clinical consultations, by time on ART for established-on-ART patient population. Fox 2019a and Fox 2019b were separated into an RCT and cohort design, respectively, based on the analysis described by the authors, where randomization was not preserved in the intervention arm in Fox 2019b. *Cluster-adjusted RR. 3 MO, 3-monthly; AC, adherence club; Comm, community; e, number of events; n, number of participants; ND, not described; P, pharmacy; RF, reduced frequency.

(TIF)

S4 Fig. Viral suppression among those enrolled: Reduced frequency versus 3-monthly clinical consultations, by clinical consultation frequency. Fox 2019a and Fox 2019b were separated into an RCT and cohort design, respectively, based on the analysis described by the authors, where randomization was not preserved in the intervention arm in Fox 2019b. *Cluster-adjusted RR. 3 MO, 3-monthly; AC, adherence club; Comm, community; e, number of events; n, number of participants; ND, not described; P, pharmacy; RF, reduced frequency.

(TIF)

S5 Fig. Viral suppression among those with viral load testing: Reduced frequency versus 3-monthly clinical consultations, by clinical consultation frequency. Fox 2019a and Fox 2019b were separated into an RCT and cohort design, respectively, based on the analysis described by the authors, where randomization was not preserved in the intervention arm in Fox 2019b. *Cluster-adjusted RR. 3 MO, 3-monthly; AC, adherence club; Comm, community; e, number of events; n, number of participants; RF, reduced frequency.

(TIF)

S6 Fig. Viral suppression among those enrolled: Reduced frequency versus 3-monthly clinical consultations, by delivery strategy. Fox 2019a and Fox 2019b were separated into an RCT and cohort design, respectively, based on the analysis described by the authors, where randomization was not preserved in the intervention arm in Fox 2019b. *Cluster-adjusted RR. 3 MO, 3-monthly; AC, adherence club; Comm, community; e, number of events; n, number of participants; RF, reduced frequency.

(TIF)

S7 Fig. Viral suppression among those with viral load testing: Reduced frequency versus 3-monthly clinical consultations, by delivery strategy. Fox 2019a and Fox 2019b were separated into an RCT and cohort design, respectively, based on the analysis described by the authors, where randomization was not preserved in the intervention arm in Fox 2019b. *Cluster-adjusted RR. 3 MO, 3-monthly; AC, adherence club; Comm, community; e, number of events; n, number of participants; RF, reduced frequency.

(TIF)

S1 PRISMA Checklist.

(DOCX)

S1 Table. Outcome definitions by study.

(DOCX)

S2 Table. Ascertainment of viral suppression by study.

(DOCX)

S3 Table. Risk of bias.

(DOCX)
Author Contributions


Funding acquisition: Nathan Ford, Elvin H. Geng.


Project administration: Ingrid Eshun-Wilson.

Supervision: Elvin H. Geng, Ingrid Eshun-Wilson.


Writing – original draft: Noelle Le Tourneau, Elvin H. Geng, Ingrid Eshun-Wilson.


References


Evaluation of reduced visit frequency and HIV outcomes


