

Documentation

Table of contents

1	MW-PrEP Visit Log	3
1.1	Visit Log Overview	3
1.1.1	Form Purpose	3
1.1.2	Context of the Form	4
1.1.3	Visit Log in the System of Programme Tools	4
1.2	How Data Flows	4
1.3	Form Summary	5
1.4	How to Use This Documentation	5
2	Indicator Definitions and Calculations	7
2.1	Routine Indicators	7
2.1.1	Detailed Definitions	8
2.2	Custom Indicators	14
2.3	Programme Cascade Indicators	16
2.4	DQA Flag Handling	17
3	Data Dictionary	18
3.1	Summary	18
3.2	Identifiers	18
3.3	Client Details	19
3.4	Visit Details	21
3.5	Readiness Assessment	22
3.6	STI Screening	23
3.7	Side Effects	25
3.8	PrEP Given	27
3.9	Condoms Given	27
3.10	Contraception Assessment	29
3.11	Visit Outcome	31
3.12	Administrative & Quality Control Fields	33
3.13	OCR Models Reference	35
3.14	Verification Levels	35
4	Data Quality Assessment	37
5	Part 1 — Client-Side Data Validation	38
5.1	Row Eligibility	38
5.2	Active Checks — Page 1	38
5.3	Active Checks — Page 2	39
5.4	Check Types Explained	39
5.5	Fields Not Validated Client-Side	39
6	Part 2 — Server-Side Data Validation	41
6.1	Record Drop Logic	41

6.2	Error Classification	41
6.3	Full Check Catalogue — Missing Values and Invalid Numbers	42
6.4	Full Check Catalogue — Invalid Dates and Codes	42
6.5	Full Check Catalogue — Multiple / Missing Answers	43
6.6	Full Check Catalogue — ID Transcription and Cross-Record Checks	43
6.7	Full Check Catalogue — Clinical Logic	44
6.8	Summary by Category	44
6.9	Check Tag Reference	45

1 MW-PrEP Visit Log

Note

This is live documentation. It is automatically regenerated each time modifications are made to the source XLSForm or pipeline configuration. The content on every tab always reflects the currently deployed form version. Data is hosted and managed in compliance with Malawi national health data protection requirements.

1.1 Visit Log Overview

1.1.1 Form Purpose

The **MW-PrEP Visit Log** is the primary clinical encounter record for the Malawi Ministry of Health Pre-Exposure Prophylaxis (PrEP) programme. At each client visit, a health worker completes one row of the register, capturing the client's identity, clinical readiness indicators, the PrEP products dispensed, contraception and STI screening results, and the visit outcome. The completed log is then photographed and digitised automatically by the ScanForm platform.

The form supports three operational objectives:

1. **Continuity of care** — providing a longitudinal record of each client's PrEP journey, including product type, adherence intent, side effects, and next appointment date.
2. **Programme monitoring** — generating the routine visit, outcome, and product dispensing counts required for PEPFAR MER reporting and Malawi MoH monthly summaries.

Programme Metadata

Field	Value
Ministry	Ministry of Health, Malawi
Programme	MW-PrEP — Pre-Exposure Prophylaxis
Form type	Visit Log (tabular register, 15 rows per spread)
Version	v3.2
Released	2026-06-09
Paper size	840 × 297 mm (A3 landscape, two-page spread)
Pages	Page 1 (client details, visit details, readiness assessment) · Page 2 (treatment, outcomes, c
Rows per form	15 client rows per physical form
OCR platform	QED ScanForm
Pipeline	dbt · ClickHouse · Metabase dashboard

3. **Clinical safety** — flagging clients with organ risk factors, side effects, or HIV positive outcomes for clinical review and appropriate action.

1.1.2 Context of the Form

The MW-PrEP Visit Log is deployed across PrEP service delivery points in Malawi, including OPD clinics, ANC services, STI clinics, family planning units, mobile outreach sites, and community distribution points. It is part of a broader suite of PrEP programme registers commissioned by the Malawi Ministry of Health and implemented with support from implementing partners.

The form was designed as a ScanForm paper register — completed by hand in the facility, then digitised on-site by a field worker using a smartphone camera. The ScanForm platform performs OCR on the photographed pages and applies automated data quality checks before the data enters the analytics pipeline.

The Visit Log (v3.2) is the tabular version of the PrEP encounter form. A companion **Visit Card** format exists for individual client tracking at higher-volume sites; both formats feed the same analytics pipeline and are documented together in the DQA and Indicators tabs.

1.1.3 Visit Log in the System of Programme Tools

The MW-PrEP Visit Log sits at the centre of a linked system of programme registers and forms:

Linked Form	Relationship
PrEP Register	Source of PrEP Registration ID; records client initiation and initial clinical assessment
HTS Initial Register	Records the HIV testing event that referred the client to PrEP. Visit Log rows carry forward HTS results
Pregnancy Registry	Supplementary form for pregnant PrEP clients. Linked by PrEP Registration ID. Includes pregnancy outcomes
PrEP Daily Inventory	Clinic-room CAB-LA vial stock record. Not linked by client ID — linked by facility
PrEP Pharmacy Inventory	Pharmacy-level weekly vial stock record. Used alongside the clinic-room inventory

1.2 How Data Flows

Completed Visit Log pages are photographed in the field using the ScanForm mobile application. The image is transmitted to the ScanForm server, where OCR models extract digit box values and oval bubble selections for each of the 15 client rows. Automated in-phone validation checks run immediately on the OCR output — if a required field is missing or a single-select bubble has multiple marks, the field worker is prompted to correct the paper form and retake the photo before submission. Once submitted, the data passes through the dbt pipeline, which applies server-side cross-record consistency and clinical logic checks, and produces the clean and refinery tables that power the Metabase dashboard and PEPFAR reporting outputs.

- **Deployment:** Physical register books are pre-printed and distributed to PrEP service delivery points by implementing partners. Each book contains multiple pages; each page is one double-sided A3 spread (Visit Log pages 1 and 2).

Form Sections at a Glance

Section	Fields	Page	Key Variables
Identifiers	2	Page 1	prep_reg_id, hts_link_id
Client Details	3	Page 1	client_name, physical_address, phone_number (handwritten)
Visit Details	5	Page 1	visit_date, year, delivery_point, age, sex_disagg
Readiness Assessment	6	Page 1	kidney_risk, liver_risk, bp_sys, bp_dias, height, weight
STI Screening	1	Page 2	sti_screening
Side Effects	1	Page 2	side_effects
PrEP Given	5	Page 2	prep_injectable_option, tdf_3tc_given, oral_prep_intention
Condoms Given	3	Page 2	condom_m_given, condom_f_given, lubes_given
Contraception Assessment	1	Page 2	contraception_assess
Visit Outcome	2	Page 2	visit_outcome, appointment_date
Administrative / QC	9	Both	discard_row, book_number, page_number, implementation_

- **Completion:** Health workers (nurses, clinical officers, or trained lay providers) complete one row per client visit using ballpoint pen. Instructions for correcting mistakes are printed on Page 1 of the form.
- **Digitisation:** Field workers photograph each completed page spread using the ScanForm mobile application, typically at the end of each clinic session or working day. Images are uploaded over mobile data or Wi-Fi.

1.3 Form Summary

1.4 How to Use This Documentation

Tip

Navigate using the tabs at the top of this page. Each tab covers a different aspect of the form:

- **Data Dictionary** — all form fields, their types, required status, constraints, OCR model configuration, and identifier linkage rules. Start here to understand what each variable means and how it is captured on paper.
- **Data Quality Assessment** — two-layer DQA documentation: client-side validation rules applied on the device immediately after photographing (before submission), and server-side checks applied after upload to the central database. Use this tab to understand what errors are caught automatically and which checks affect which records.
- **Indicators** — formal definitions of all reported indicators, including PEPFAR MER routine indicators, Malawi MoH programme indicators, and custom operational metrics. Each indicator includes its exact numerator, denominator, disaggregation dimensions, and the source dbt model column names.

For data managers: The DQA tab documents the exact check messages displayed in the Metabase DQA dashboard, so you can match dashboard alerts to the underlying rule logic without reading SQL.

For programme officers: The Indicators tab uses the same terminology as the PEPFAR MER and Malawi MoH reporting templates, with explicit notes where the pipeline definition deviates from the standard.

For clinical supervisors: The Data Dictionary tab documents all clinically relevant fields — including delivery point codes, PrEP product options, side effects recording, and visit outcome sub-reasons — using the exact labels printed on the physical form.

Generated automatically from MW-PrEP Visit Log v3.2 source files. Last updated: 2026-06-09.

2 Indicator Definitions and Calculations

Indicators are computed from the clean and refinery dbt layers of the MW-PrEP pipeline. This page defines every indicator derived from the **MW-PrEP Visit Log v3.2**, including its numerator, denominator, disaggregation, and source variables.

i Note

Indicator definitions follow **PEPFAR MER 2.7** and **Malawi Ministry of Health PrEP programme protocols** where applicable. Deviations from standard definitions are noted explicitly under each indicator. All indicators exclude records with at least one **DQA Error** in the clean dataset unless otherwise noted.

2.1 Routine Indicators

Standard indicators reported to PEPFAR and the Malawi Ministry of Health on a monthly or quarterly basis.

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Indicator	Short Name	Frequency	Source Model
PrEP New	PrEP NEW	Quarterly	mw prep program indicators report
PrEP Continuing	PrEP CT	Quarterly	mw prep program indicators report
PrEP Visit	PrEP VISIT	Monthly	mw prep program indicators report
PrEP Outcome: Continuing at Facility	PrEP CONT	Monthly	mw prep program indicators report
PrEP Outcome: Transfer Out	PrEP TO	Monthly	mw prep program indicators report
PrEP Outcome: Stop PrEP	PrEP STOP	Monthly	mw prep program indicators report
HIV Seroconversion on PrEP	HIV SEROCONV	Monthly	mw prep program indicators report
CAB-LA Initiation	CABLA INIT	Quarterly	mw prep program indicators report
CAB-LA Continuation	CABLA CONT	Monthly	mw prep program indicators report
PEPFAR PrEP New + Returning (CAB-LA)	PEPFAR CABLA	Quarterly	mw prep pepfar init and returning clients on cabla indicators

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Indicator	Numerator	Disaggregation
PrEP New	prep initiation = 1 and register prep reg type = first assesement	Sex, Age group (PEPFAR), Facility, District, Quarter (PEPFAR)
PrEP Continuing	prep visit = 1 and register prep reg type != first assesement	Sex, Age group (PEPFAR), Facility, District, Quarter (PEPFAR)
PrEP Visit	prep visit = 1	Sex, Age group (PEPFAR), Delivery point, Facility, District, Month
PrEP Outcome: Continuing at Facility	outcome = start/continue prep at facility	Sex, Age group, Facility, District, Month
PrEP Outcome: Transfer Out	prep transfer out = 1	Sex, Age group, Facility, District, Month
PrEP Outcome: Stop PrEP	outcome includes any noprep value	Stop reason, Sex, Age group, Facility, Month
HIV Seroconversion on PrEP	hiv seroconversion = 1	PrEP type at seroconversion, Facility, District, Month
CAB-LA Initiation	cab la registration = 1 and register prep reg type = first assesement	Sex, Age group, Facility, District, Quarter
CAB-LA Continuation	cab la continuation is non-null and non-empty	Continuation status, Facility, District, Follow-up month
PEPFAR PrEP New + Returning (CAB-LA)	total clients initiated on cabla; total clients returning in one month second visit	Report end date

2.1.1 Detailed Definitions

2.1.1.1 PrEP New

Field	Value
Full name	PrEP New — Individuals initiating PrEP for the first time
Short name	PrEP_NEW
Definition	Count of individuals who received PrEP for the first time during the reporting period, recorded as a first assessment in the PrEP Register.
Numerator	COUNT(*) where prep_initiation = 1 AND register_prep_reg_type = 'first assesement'
Denominator	N/A
Disaggregation	sex_pepfar, agegroup_pepfar, location_name, district, quarter_pepfar
Inclusion criteria	Records with a linked PrEP Register entry; visit date on or after site activation date.
Exclusion criteria	Records with DQA Error severity; transfer-in registrations; reassessments.
Reporting frequency	Quarterly (PEPFAR MER)
Source variables	prep_initiation, register_prep_reg_type, sex_pepfar, agegroup_pepfar, quarter_pepfar, location_name

Field	Value
Reference	PEPFAR MER 2.7 § PrEP_NEW

Transfer-in clients (`register_prep_reg_type = 'transfer in'`) are excluded from PrEP_NEW and counted separately. Reassessments (`register_prep_reg_type = 'reassessment'`) are also excluded. The `prep_initiation_flag` column in the source model documents edge cases in cohort construction.

2.1.1.2 PrEP Continuing

Field	Value
Full name	PrEP Continuing — Clients continuing PrEP in the reporting period
Short name	PrEP_CT
Definition	Count of individuals who received at least one PrEP visit during the reporting period and had previously initiated PrEP (i.e. not a new initiate).
Numerator	COUNT(*) where <code>prep_visit = 1</code> AND <code>register_prep_reg_type != 'first assessement'</code>
Denominator	N/A
Disaggregation	<code>sex_pepfar</code> , <code>agegroup_pepfar</code> , <code>location_name</code> , <code>district</code> , <code>quarter_pepfar</code>
Inclusion criteria	Must have a prior PrEP initiation record; visit outcome = start/continue.
Exclusion criteria	Records with DQA Error; first-assessment visits; defaulters without a return visit in period.
Reporting frequency	Quarterly (PEPFAR MER)
Source variables	<code>prep_visit</code> , <code>register_prep_reg_type</code> , <code>sex_pepfar</code> , <code>agegroup_pepfar</code> , <code>quarter_pepfar</code>
Reference	PEPFAR MER 2.7 § PrEP_CT

The `prep_visit_flag` column documents records with ambiguous visit classification. Defaulters who return within the same reporting quarter are counted as continuing.

2.1.1.3 PrEP Visit

Field	Value
Full name	Total PrEP visits in reporting period
Short name	PrEP_VISIT

Field	Value
Definition	Total number of valid PrEP client visit records in the reporting period, regardless of registration type or outcome.
Numerator	COUNT(*) where prep_visit = 1
Denominator	N/A
Disaggregation	sex_pepfar, agegroup_pepfar, delivery_point_name, location_name, district, month_calendar
Inclusion criteria	All valid, non-discarded visit records with a parseable visit date.
Exclusion criteria	DQA Error records; discarded rows; implementation error rows.
Reporting frequency	Monthly
Source variables	prep_visit, visit_card_visit_date, delivery_point_name, location_name
Reference	Malawi MoH PrEP programme protocol

2.1.1.4 PrEP Outcome: Continuing at Facility

Field	Value
Full name	PrEP clients continuing at facility — visit outcome
Short name	PrEP_CONT
Definition	Number of visits where the recorded outcome was start/continue PrEP at this facility.
Numerator	COUNT(*) where outcome = 'start/continue prep at facility' (column: prep_here)
Denominator	N/A
Disaggregation	sex_pepfar, agegroup_pepfar, location_name, district, month_calendar
Inclusion criteria	Valid visit records only.
Exclusion criteria	DQA Error records.
Reporting frequency	Monthly
Source variables	prep_here, outcome, visit_card_visit_date, location_name
Reference	Malawi MoH PrEP programme protocol

2.1.1.5 PrEP Outcome: Stop PrEP

Field	Value
Full name	Clients stopping PrEP — by stop reason
Short name	PrEP_STOP
Definition	Number of clients who discontinued PrEP during the reporting period, disaggregated by reason for stopping.
Numerator	COUNT(*) where prep_outcome is in stop categories: noprep_refuse, noprep_side_effect, noprep_low_risk, noprep_hiv_positive, noprep_hiv_positive_outcome
Denominator	N/A
Disaggregation	stop_reasons, sex_pepfar, agegroup_pepfar, location_name, district, month_calendar
Inclusion criteria	Valid visit records with a stop outcome.
Exclusion criteria	DQA Error records.
Reporting frequency	Monthly
Source variables	stop_reasons, noprep_refuse, noprep_side_effect, noprep_low_risk, noprep_hiv_positive, visit_card_visit_date
Reference	Malawi MoH PrEP programme protocol; PEPFAR MER 2.7

Stop reasons map directly to the visit_outcome bubble options on the physical form. The stop_reasons_flag column documents invalid or multi-selected stop reason combinations.

2.1.1.6 HIV Seroconversion on PrEP

Field	Value
Full name	HIV seroconversion while on PrEP
Short name	HIV_SEROCONV
Definition	Number of PrEP clients who tested HIV positive during a follow-up visit while enrolled in PrEP, indicating a seroconversion event.
Numerator	COUNT(*) where hiv_seroconversion = 1
Denominator	N/A
Disaggregation	prep_type_when_seroconverted, hiv_seroconversion_on_cab_la, location_name, district, seroconversion_follow_up_month
Inclusion criteria	Clients with a prior clean PrEP initiation record; HIV positive outcome confirmed in visit record or HTS initial register.
Exclusion criteria	Clients already HIV positive at initiation; DQA Error records.
Reporting frequency	Monthly

Field	Value
Source variables	hiv_seroconversion, hiv_seroconversion_on_cab_la, prep_type_when_seroconverted, seroconversion_follow_up_month
Reference	Internal programme protocol; PEPFAR MER 2.7

The `hiv_seroconversion_flag` column marks records with ambiguous seroconversion status (e.g. missing HIV result, or HIV positive outcome without a confirmed test date). Seroconversions on CAB-LA are separately tracked via `hiv_seroconversion_on_cab_la`.

2.1.1.7 CAB-LA Initiation

Field	Value
Full name	Cabotegravir LA injectable PrEP — new initiates
Short name	CABLA_INIT
Definition	Number of clients starting injectable PrEP (Cabotegravir LA) for the first time during the reporting period.
Numerator	COUNT(*) where <code>cab_la_registration = 1 AND register_prep_reg_type = 'first assesement'</code>
Denominator	N/A
Disaggregation	sex_pepfar, agegroup_pepfar, cab_la_enrolment_type, location_name, district, quarter_pepfar
Inclusion criteria	Valid PrEP Register and Visit Card/Log linked records; CAB-LA prescribed and injection given.
Exclusion criteria	Oral-only initiates; clients switching from oral to CAB-LA (counted separately as <code>started_cab_after_first_visit</code>); DQA Error records.
Reporting frequency	Quarterly
Source variables	cab_la_registration, cab_la_enrolment_type, started_cab_first_visit, register_prep_reg_type
Reference	PEPFAR MER 2.7; Malawi MoH CAB-LA rollout protocol

Clients switching from oral PrEP to CAB-LA after first visit are tracked via `started_cab_after_first_visit`. The `cab_la_specific_registration_flag` column identifies records with ambiguous CAB-LA registration status.

2.1.1.8 CAB-LA Continuation

Field	Value
Full name	Cabotegravir LA injectable PrEP — continuation
Short name	CABLA_CONT
Definition	Number of CAB-LA clients returning for continuation injections, disaggregated by continuation timing status (on time, late, defaulted).
Numerator	COUNT(*) where cab_la_continuation is non-null and non-empty
Denominator	N/A
Disaggregation	cab_la_continuation_status, cab_la_continuation_timing, cab_la_follow_up_month, location_name, district
Inclusion criteria	Clients with at least one prior CAB-LA visit.
Exclusion criteria	Clients who stopped CAB-LA; DQA Error records.
Reporting frequency	Monthly
Source variables	cab_la_continuation, cab_la_continuation_status, cab_la_continuation_timing, cab_la_follow_up_month, master_registration_date
Reference	PEPFAR MER 2.7; Malawi MoH CAB-LA rollout protocol

Follow-up months are calculated from the master registration date. The cab_la_continuation_flag documents records where continuation status cannot be determined.

2.1.1.9 PEPFAR CAB-LA New and Returning Clients

Field	Value
Full name	PEPFAR — Clients initiated on CAB-LA and returning within one month
Short name	PEPFAR_CABLA
Definition	PEPFAR standard counts: (1) total clients initiated on CAB-LA; (2) clients returning for second visit within one month; (3) clients returning for any visit within one month.
Numerator	total_clients_initiated_on_cabla; total_clients_returning_in_one_month_second_visit; total_clients_returning_in_one_month_any_visit
Denominator	N/A

Field	Value
Disaggregation	report_end_date
Inclusion criteria	All valid CAB-LA initiation records.
Exclusion criteria	DQA Error records.
Reporting frequency	Quarterly
Source variables	total_clients_initiated_on_cabla, total_clients_returning_in_one_month_second_visit, total_clients_returning_in_one_month_any_visit, report_end_date
Reference	PEPFAR MER 2.7 § PrEP_NEW (CAB-LA sub-indicator)

2.2 Custom Indicators

Programme-specific indicators not part of the standard PEPFAR or MoH reporting framework but used for operational monitoring and programme management.

Indicator	Frequency	Source Model
Active clients - current status	Daily	mw prep current client status
Defaulters (up to 120 days)	Daily	mw prep refinery defaulters
Clients late for appointment	Daily	mw prep refinery clients late for appointment
Upcoming appointments (7-day window)	Daily	mw prep refinery appointment schedule
Sex-age mismatch rate	Monthly	mw prep mismatch sex and age stats
Orphan rate (visit log to PrEP register)	Monthly	mw prep orphans aggregated
Orphan rate (visit log to HTS register)	Monthly	mw prep orphans aggregated
Submission timeliness	Monthly	mw prep delays analysis
Upload delay (days)	Monthly	mw prep delays analysis
Book stock status	Daily	mw prep active books per location
Pregnant client PrEP surveillance	Daily	mw prep path2scale pregnancy report
CAB-LA vial inventory - clinic room	Daily	mw prep clinic room inventory summary
CAB-LA vial inventory - pharmacy	Weekly	mw prep refinery pharmacy
STI positivity rate among PrEP clients	Monthly	mw prep refinery standard summary
Side effects reporting rate	Monthly	mw prep refinery standard summary
Condoms dispensed	Monthly	mw prep refinery standard summary
Family planning coverage	Monthly	mw prep refinery standard summary

Indicator	Numerator	Denominator
Active clients - current status	is active = 1	N/A
Defaulters (up to 120 days)	defaulted = true (within 120 days)	N/A
Clients late for appointment	is late for appointment = true and not defaulted	N/A
Upcoming appointments (7-day window)	next appointment within 7 days, not late, not defaulted	N/A

(continued)

Indicator	Numerator	Denominator
Sex-age mismatch rate	mismatched sex = true OR mismatched age greater than 1 year = true	All records with a valid HTS link
Orphan rate (visit log to PrEP register)	orphan direction tag = visit card to reg and orphan status = orphan	All valid visit log records
Orphan rate (visit log to HTS register)	orphan direction tag = visit card to hts and orphan status = orphan	visit card records
Submission timeliness	on time = true	Total pages in period
Upload delay (days)	AVG(upload delay in days)	N/A
Book stock status	months remaining from first pages remaining / pages used per day	N/A
Pregnant client PrEP surveillance	active pregnancy records where has preg outcome = false	N/A
CAB-LA vial inventory - clinic room	recorded closing stock; expected closing stock	N/A
CAB-LA vial inventory - pharmacy	recorded weekly closing stock; expected weekly closing stock	N/A
STI positivity rate among PrEP clients	sti screening = pos	sti screening is not null and not ND
Side effects reporting rate	side effects = yes	side effects is not null and not ND
Condoms dispensed	SUM(male condoms given) + SUM(female condoms given)	N/A
Family planning coverage	fp is not null and not missing/invalid	Eligible visits (non-male clients)

Indicator	Disaggregation	Rationale
Active clients - current status	PrEP type, Facility, District	Supports facility-level management of active client caseloads.
Defaulters (up to 120 days)	Days late, PrEP type, Facility	Supports tracing and re-engagement of lost clients.
Clients late for appointment	Days late, PrEP type, Facility	Supports proactive outreach before clients become defaulters.
Upcoming appointments (7-day window)	PrEP type, Facility, Days to appointment	Supports daily clinic scheduling and preparation.
Sex-age mismatch rate	Facility, District	Monitors data consistency across PrEP and HTS registers.
Orphan rate (visit log to PrEP register)	Facility, District, Month	Monitors ID linkage quality and flags systematic transcription problems.
Orphan rate (visit log to HTS register)	Facility, District, Month	Monitors cross-form linkage integrity between PrEP Visit Log and HTS Initial Register.
Submission timeliness	Facility, District, Submitted by	Monitors data submission discipline; supports program management follow-up.
Upload delay (days)	Facility, District, Submitted by	Identifies sites with systematic delays in scanning and uploading forms.
Book stock status	Register type, Facility, District, Region	Supports logistics planning and prevents data collection gaps.
Pregnant client PrEP surveillance	Passed EDD, Due next month, Current PrEP type, Facility	Supports clinical monitoring of pregnant women on PrEP.
CAB-LA vial inventory - clinic room	Facility, Week, Month	Enables real-time stock monitoring and detection of discrepancies.

(continued)

Indicator	Disaggregation	Rationale
CAB-LA vial inventory - pharmacy	Facility, Week, Month	Enables pharmacy-level stock reconciliation and supply chain management.
STI positivity rate among PrEP clients	Sex, Age group, Facility, District, Month	Monitors STI burden in PrEP clients; informs integrated prevention programming.
Side effects reporting rate	PrEP type, Sex, Age group, Facility, Month	Monitors tolerability of PrEP regimens; triggers pharmacovigilance review.
Condoms dispensed	Condom type (M/F), Facility, District, Month	Monitors commodity distribution at PrEP service points.
Family planning coverage	FP outcome category, Facility, District, Month	Monitors integration of family planning services into PrEP delivery.

2.3 Programme Cascade Indicators

These indicators together describe the PrEP care cascade and are used for high-level programme performance reviews.

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Step	Cascade Stage	Source Column	Definition
1	Initial PrEP Screening	prep screening	Client presented for initial PrEP eligibility screening.
2	PrEP Eligible	prep eligible	Client assessed as eligible for PrEP initiation.
3	PrEP Initiation	prep initiation	Client initiated PrEP (first visit with prescription).
4	PrEP Registration	prep registration	Client registered in the PrEP Register.
5	PrEP Visit (Follow-up)	prep visit	Client attended at least one follow-up PrEP visit.
6	PrEP Outcome Recorded	prep outcome	A visit outcome was recorded (continuing, transferred, stopped).
7	CAB-LA Registration	cab la registration	Client registered as a CAB-LA client.
8	CAB-LA Visit	cab la visit	Client attended a CAB-LA injection visit.
9	CAB-LA Continuation	cab la continuation	Client returned for continuation injection on schedule.

All cascade steps are available in `mw_prep__program_indicators_report` as the `indicator / value / flag` array-joined columns. Steps 7–9 are only populated for clients enrolled on CAB-LA.

2.4 DQA Flag Handling

Warning

Records with DQA **Errors** are excluded from all indicator calculations by default. The table below documents which specific DQA checks affect which indicators and how.

DQA Check	Indicators Affected	Effect
prep reg id missing	All indicators	Record excluded from numerator and denominator - no PrEP ID means no cohort linkage possible.
PrEP ID mod-223 transcription error	All indicators using prep reg id linkage	Record excluded unless orphan matching finds a verified alternative.
HTS Link ID Damm transcription error	Sex-age mismatch rate; HIV seroconversion; Cross-register sex consistency	Record excluded from cross-register linked indicators; included in non-linked indicators.
visit date invalid or out of range	All time-series indicators (PrEP VISIT, PrEP NEW, PrEP CT, monthly summaries)	Record excluded from numerator - cannot be assigned to a reporting period.
appointment date invalid or before visit date	Upcoming appointments; Clients late for appointment; Defaulters	Record excluded from appointment schedule - next visit date cannot be computed.
Sex/Pregnancy multi-select or missing	All sex-disaggregated indicators (PrEP NEW, PrEP CT, PEPFAR summary)	Record excluded from sex-disaggregated numerators; counted under missing/invalid sex category.
Visit Outcome multi-select or missing	PrEP STOP, PrEP CONT, PrEP Transfer Out, HIV seroconversion	Record excluded from outcome-based indicators; cannot be assigned a stop reason or continuation status.
PrEP option multi-select or missing	CAB-LA Initiation, CAB-LA Continuation, PrEP Dispensed	Record excluded from regimen-based indicators; counted under prescribe invalid or prescribe missing.
Age invalid integer	All age-disaggregated PEPFAR indicators	Record excluded from age-band numerators; counted under age invalid.
HIV positive client prescribed PrEP	PrEP NEW, PrEP CT, CAB-LA Initiation	Record excluded from all active-prescription indicators as clinically ineligible.
PrEP not prescribed but continued at facility	PrEP CT, PrEP CONT	Record excluded - outcome and prescription are contradictory.
Multiple visit entries same day	PrEP VISIT, PrEP CT, all time-series counts	Duplicate records excluded - only one visit per client per date is counted.
PrEP ID not unique	All indicators aggregated by prep reg id	Affected records flagged; excluded from cohort-level indicators until resolved.
HIV positive client has subsequent visit	PrEP CT, CAB-LA Continuation, active client status	Post-seroconversion visits excluded from active PrEP continuation counts.
Missing appointment date for continuing client	Upcoming appointments; Clients late for appointment; Defaulters	Record cannot enter the appointment tracking pipeline; excluded from scheduling indicators.
PrEP register not linked	All indicators requiring register-visit card join (PrEP NEW, sex consistency)	Record treated as orphan; excluded from linked indicators; included in orphan rate calculation.
CAB-LA injection inconsistency	CAB-LA Initiation, CAB-LA vial inventory reconciliation	Warning only - record retained but flagged; highlighted in DQA dashboard for clinical review.

3 Data Dictionary

This section defines all variables recorded on the **MW-PrEP Visit Log v3.2** — their names, labels, types, formats, constraints, and ScanForm OCR configurations. Variables are organised by XLSForm group, matching the physical layout of the two-page register.

3.1 Summary

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Table 3.1: Variable Summary — MW-PrEP Visit Log v3.2

Category	Count
Total variables (excluding system metadata)	38
Required	22
Optional	16
OCR-processed (digit boxes / oval bubbles)	33
Handwritten only (scanned but not extracted)	5

3.2 Identifiers

! Important

IDs link PrEP visit records across forms and across time. Errors in IDs break linkage. Both identifier fields are mandatory and use restricted character sets with check-digit positions to reduce transcription errors.

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Table 3.2: Identifier Fields — Overview

Field	Full Name	Format	Entered By
-------	-----------	--------	------------

prep reg id	PrEP Registration ID	PRP-XXXXXX-XXX-XX (13 chars)	Copied from PrEP register or client Health Passport
hts link id	HTS Link ID	HTS-XXXXXX-XXX-XX- X (14 chars)	Copied from integrated initial testing register or client Health Passport

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Table 3.3: Identifier Fields — Details

Field	Check Digit	Links To	Notes
prep reg id	Positions 12-13 use string restricted BFGHJPVWXZ	PrEP register	Pre-printed prefix PRP-. Valid chars: 0-9 and BEFGHJPRSTVWXZ. OCR model enforces restricted character set on final 2 boxes.
hts link id	Position 14 uses string restricted BFGHJPVWXZ	Integrated initial testing register	Pre-printed prefix HTS-. Valid chars: 0-9 and BEFGHJPRSTVWXZ. row int model used for positions 12-13.

i Note

ID format note: Both IDs are printed on the form with pre-filled letter prefixes (PRP- and HTS-). Only the variable character boxes are OCR-processed. The restricted OCR model `string_restricted_BFGHJPVWXZ` is applied to the final check-digit position(s), limiting recognised characters to B F G H J P V W X Z to reduce confusion between visually similar characters (e.g. 0/O, 1/I). Verification level is **verified** — the verifier is shown these fields only when OCR confidence is below threshold.

3.3 Client Details

Handwritten client contact information recorded once per client row. These fields are printed on Page 1 of the register. ScanForm scans but does **not** perform OCR extraction on these fields — they are for visual reference only.

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Table 3.4: Client Detail Fields

Variable	Label	Type	Required	Constraint	OCR Model	Notes
client name	Name	text	No	Max 35 alpha characters	Handwritten	Banded = false. Not OCR-processed. For visual reference by verifier.
physical address	Physical Address	text	No	Max 35 alpha characters	Handwritten	Banded = false. Not OCR-processed.
phone number	Phone Number	text	No	Max 10 digits	Handwritten	Banded = false. Not OCR-processed.

3.4 Visit Details

Core visit information recorded on Page 1. Captures the visit date, facility/community delivery point, client age, and sex/pregnancy status.

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Table 3.5: Visit Detail Fields — Overview

Variable	Label	Type	Format	Required
visit date	Visit Date	text	dd/mm (4 digits)	Yes
Year	Year	text	yyyy (4 digits)	Yes
delivery point	Delivery Point	text	2 digit boxes	Yes
age	Age	text	2 digit boxes	Yes
sex disagg	Sex/Pregnancy	select one	Oval bubbles	Yes

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Table 3.6: Visit Detail Fields — OCR and Notes

Variable	OCR Model	Notes
visit date	date dd, date dd, date mm, date mm	Year is captured separately in the Year field. Verification: verified.
Year	int, int, date nearby yy, date nearby yy	Pre-printed 20. Last 2 digits use date nearby yy model (current year +/-1). Verification: verified.
delivery point	location int	Codes: 11=OPD, 2=ANC, 4=STI, 20=Family Planning, 12=Other (facility); 15=Mobile, 21=Outreach, 17=Other (community). Verification: certain.
age	int	Age in years. Verification: certain.
sex disagg	select one or zero	Options: M (Male), F (Female), NP (Not Pregnant), FP (Family Planning), FBf (Feeding/Breastfeeding). Verification: certain.

i Note

Delivery Point codes are printed in the header of Page 1 for reference:

Code	Facility	Code	Community
11	OPD	15	Mobile
2	ANC	21	Outreach
4	STI	17	Other
20	Family Planning		

3.5 Readiness Assessment

Clinical readiness checks recorded on Page 1, including risk factor screening and vital signs. All fields are mandatory.

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Table 3.7: Readiness Assessment Fields — Overview

Variable	Label	Type	Format	Required
kidney risk	Kidney Risk Factors	select one	Oval bubbles	Yes
liver risk	Liver Risk Factors	select one	Oval bubbles	Yes
bp sys	BP: Systolic	integer	3 digit boxes (sys)	No
bp dias	BP: Diastolic	integer	3 digit boxes (dias)	No
height	Height	text	3 digit boxes (cm)	Yes
weight	Weight	text	4 digits incl. decimal (kg)	Yes

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Table 3.8: Readiness Assessment Fields — Details

Variable	Constraint	OCR Model	Notes
kidney risk	Single selection only	select one or zero	Options: Yes / No / ND. Verification: certain.
liver risk	Single selection only	select one or zero	Options: Yes / No / ND. Verification: certain.
bp sys	$60 \leq \text{value} \leq 250$	int	Systolic blood pressure in mmHg. Verification: certain.
bp dias	$40 \leq \text{value} \leq 140$	int	Diastolic blood pressure in mmHg. Verification: certain.
height	Max 3 digits	int	Height in centimetres. Verification: certain.
weight	Max 4 digits including decimal	int	Weight in kilograms with one decimal place. Pre-printed decimal point. Verification: certain.

3.6 STI Screening

Single-field section on Page 2 capturing the result of STI screening at the visit.

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Table 3.9: STI Screening Field

Variable	Label	Type	Format	Required	Constraint	OCR Model	Notes
sti screening	STI Screening	select one	Oval bubbles	Yes	Single selection only	select one or zero	Options: Pos / Neg / ND. Verification: certain.

3.7 Side Effects

Records whether the client reported any side effects at this visit. If **Yes** is selected, the clinician must document details in the Comments field and submit a Suspected Adverse Drug Reaction Form.

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Table 3.10: Side Effects Field

Variable	Label	Type	Required	Constraint	OCR Model	Notes
side effects	Any Side Effects	select one	Yes	Single selection only	select one or zero	Options: Yes / No / ND. If Yes: details go in Comments field; ADR form must be submitted. Verification: certain.

3.8 PrEP Given

Records the type and quantity of PrEP products dispensed at the visit on Page 2. Covers injectable PrEP, oral TDF/3TC tablets, Lenacapavir tablets, and oral PrEP intention.

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Table 3.11: PrEP Given Fields — Overview

Variable	Label	Type	Format	Required
prep injectable option	PrEP Injectable	select one	Oval bubbles	No
tdf 3tc given	PrEP Tabs (TDF/3TC)	text	3 digit boxes	No
oral prep intention	Oral PrEP Intention	select one	Oval bubbles	No
len tabs given	Lenacapavir Tablets	text	2 digit boxes	No
none tabs given	None (Tablets)	select one	Oval bubble	No

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Table 3.12: PrEP Given Fields — Details

Variable	Constraint	OCR Model	Notes
prep injectable option	Single selection only	select one or zero	Options: CAB-LA (Cabotegravir LA) / LEN (Lenacapavir). Verification: certain.
tdf 3tc given	Max 3 digits	int	Number of TDF/3TC tablets dispensed. Verification: certain.
oral prep intention	Single selection only	select one or zero	Options: Continuous / Event-Driven (ED) / None. Verification: certain.
len tabs given	Max 2 digits	int	Number of Lenacapavir (LEN) tablets dispensed. Verification: certain.
none tabs given	Single selection only	select one or zero	Select if no tablets dispensed. Verification: certain.

3.9 Condoms Given

Records the quantities of male condoms, female condoms, and lubricants dispensed at the visit. All three fields are mandatory.

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Table 3.13: Condoms Given Fields

Variable	Label	Type	Format	Required	Constraint	OCR Model	Notes
condom m given	Condoms Male	text	3 digit boxes	Yes	Max 3 digits	int	Number of male condoms dispensed. Verification: certain.
condom f given	Condoms Female	text	3 digit boxes	Yes	Max 3 digits	int	Number of female condoms dispensed. Verification: certain.
lubes given	Lubricants	text	3 digit boxes	Yes	Max 3 digits	int	Number of lubricants dispensed. Verification: certain.

3.10 Contraception Assessment

Single-field section on Page 2 recording the outcome of the family planning / contraception assessment for the visit.

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Table 3.14: Contraception Assessment Field

Variable	Label	Type	Format	Required	Constraint	OCR Model	Notes
contraception assess	Contraception Assessment	select one	Oval bubbles	Yes	Single selection only	select one or zero	Options: On effective FP / No FP need / FP provided here today / FP needed but not provided. Verification: certain.

3.11 Visit Outcome

Records the clinical outcome of the visit and the next appointment date. The visit outcome field includes stop-PrEP sub-reasons printed on the form for clinician guidance.

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Table 3.15: Visit Outcome Fields — Overview

Variable	Label	Type	Required	Constraint
visit outcome	Visit Outcome	select one	Yes	Single selection only
appointment date	Next Appointment Date	date	Yes	Valid date

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Table 3.16: Visit Outcome Fields — Details

Variable	Format	OCR Model	Notes
visit outcome	Oval bubbles	select one or zero	Options: At This Facility / Transfer Out (T.O.) / Quit (Stop PrEP). Stop PrEP sub-reasons: no more need despite anticipated substantial risk; client decided not to continue; due to contraindications/side-effects/suspected AHI; Low Risk; HIV Positive; Died. Verification: certain.
appointment date	DD/MM/YY (6 digit boxes)	date dd, date dd, date mm, date mm, date nearby yy, date nearby yy	Next scheduled appointment. date nearby yy model used for year (current year +/-1). Verification: verified.

3.12 Administrative & Quality Control Fields

Fields used for scan management, data quality control, and record traceability. These fields appear on both pages of the form and are processed at the page and row level.

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Table 3.17: Administrative and QC Fields — Overview

Variable	Label	Page	Type	Required
discard left and right	Discard Right & Left Page	Page 1	trigger	No
photo taken	Photo Taken	Page 1	trigger	No
book number page 1	Book Number	Page 1	integer	No
page number page 1	Page Number	Page 1	integer	No
discard row	Discard Row	Page 2	trigger	No
book number	Book Number	Page 2	integer	No
page number page 2	Page Number	Page 2	integer	No
implementation error	Implementation Error	Page 2	select one	No
comment	Comments	Page 2	text	No

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Table 3.18: Administrative and QC Fields — Details

Variable	Format	OCR Model	Notes
discard left and right	Single oval bubble	select one or zero	Marks the entire two-page spread for discard. Applied before scanning. Verification: certain.
photo taken	Single oval bubble	select one or zero	Confirms that a photo of this page was captured. Verification: certain.
book number page 1	5 mm digit line	int	Register book number on Page 1. Verification: certain.
page number page 1	3 mm digit line	int	Page number within the book on Page 1. Verification: certain.
discard row	Single oval bubble	select one or zero	Marks an individual client row for discard. Verification: certain.
book number	5 mm digit line	int	Register book number on Page 2. Should match Page 1. Verification: certain.
page number page 2	3 mm digit line	int	Page number within the book on Page 2. Should match Page 1. Verification: certain.
implementation error	Single oval bubble	select one or zero	Flags a row-level implementation error for follow-up. Verification: certain.
comment	Handwriting line 45 mm	Handwritten	Free-text comments, e.g. side-effect details. Banded = false. Not OCR-processed.

3.13 OCR Models Reference

The table below summarises all OCR models used in this form and the character sets they detect.

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Table 3.19: OCR Models Used in MW-PrEP Visit Log v3.2

OCR Model	Characters Detected	Used For
int	Digits 0-9	Age, BP, height, weight, tablet counts, condom counts, book/page numbers
date dd	2-box integer 1-31	Day component of visit date and appointment date
date mm	2-box integer 1-12	Month component of visit date and appointment date
date nearby yy	2-box integer: current year +/-1	Year component of appointment date (last 2 digits)
two digits 20 / int	Fixed 20 then OCR digits	Year field: first 2 digits pre-printed, last 2 OCR
select one or zero	Oval: X or no X; max one X per question	All select one fields (sex, outcomes, risk factors, STI, side effects, etc.)
string	Letters A-Z	Character positions in PrEP Registration ID and HTS Link ID
string restricted BFGHJPVWXZ	Letters B F G H J P V W X Z only	Check-digit positions in PrEP Registration ID and HTS Link ID
location int	Integer constrained to valid location codes	Delivery Point code
row int	Integer constrained to valid row values	Positions 12-13 of HTS Link ID

3.14 Verification Levels

ScanForm uses two verification levels to determine when a field is shown to a human verifier:

Level	Meaning	Fields in This Form
verified	Field is shown to the verifier only when OCR confidence is <i>below</i> threshold (uncertain result). When OCR is confident, the field is hidden automatically.	prep reg id, hts link id, visit date, Year, appointment date

Level	Meaning	Fields in This Form
certain	Field is hidden from the verifier when OCR is certain. Most structured fields use this level.	All remaining OCR-processed fields

 Tip

Why do IDs and dates use *verified* instead of *certain*? These fields have the highest linkage impact. Even when OCR confidence is high, a verifier reviews them to catch errors that might otherwise go undetected and break cross-form record linkage.

Generated automatically from MW-PrEP Visit Log v3.2 source files. Last updated: 2026-06-09.

4 Data Quality Assessment

Data Quality Assessment (DQA) checks are implemented at two levels to ensure data completeness, consistency, and accuracy across all MW-PrEP Visit Log records.

Two-layer architecture: Client-side validation catches errors at the point of collection — before the image is submitted. Server-side validation catches errors that are only visible at the population level, or that were skipped by the field worker on the device.

5 Part 1 — Client-Side Data Validation

When a field worker photographs a form page, ScanForm immediately runs these validation checks against the OCR results on the device. If any check fails, the app **instantly prompts the worker to correct the paper form and retake the photo** — before the data is submitted. There is no human verifier involved at this stage; validation is fully automated and real-time.

Important constraints of client-side validation:

- Validates OCR-extracted field values directly on the device — no network connection required.
- Field workers **can skip** warnings and proceed with submission; client-side checks are advisory, not blocking.
- Does **not** have access to the full dataset — cross-record and longitudinal checks must be done server-side.
- Administrative fields (book number, page number, discard flags, photo taken) are **not** validated client-side.

Because field workers can bypass checks, server-side validation is always required as a second layer.

5.1 Row Eligibility

A row is considered **active** (i.e. validation rules apply) when:

At least **one** of the following key fields has ≥ 1 filled box: `visit_date`, `prep_reg_id` (≥ 3 filled boxes), `sex_disagg`, `age`, `visit_outcome` (≥ 2 filled boxes) — **AND** the `discard_row` bubble for that row is **not** marked.

This logic prevents spurious validation alerts on intentionally blank rows. The visit log contains 15 client rows per page spread; on any given day, many rows may be empty. Without this eligibility gate, every unfilled row would trigger a cascade of false “required field missing” alerts.

In addition, a **page-level** check validates that the **Year** field is filled if any row on the page is active. The year field is shared across all rows on the page and must always be present.

5.2 Active Checks — Page 1

Table 5.1: Client-Side Active Checks — Page 1

Field	Label	Check Type	Rule
Year	Year	enough filled	≥ 1 box filled (page-level, not per row)
visit date	Visit Date	enough filled	At least 1 box must be filled
delivery point	Delivery Point	enough filled	At least 1 box must be filled
prep reg id	PrEP Registration ID	enough filled	≥ 1 box filled (active if ≥ 3 boxes filled)

hts link id	HTS Link ID	enough filled	≥ 1 box filled (active if ≥ 3 boxes filled)
sex disagg	Sex/Pregnancy	exactly one	Exactly one bubble must be selected
age	Age	enough filled	At least 1 box must be filled
height	Height	enough filled	At least 1 box must be filled
weight	Weight	enough filled	At least 1 box must be filled
kidney risk	Kidney Risk Factors	exactly one	Exactly one bubble must be selected

5.3 Active Checks — Page 2

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Table 5.2: Client-Side Active Checks — Page 2

Field	Label	Check Type	Rule
liver risk	Liver Risk Factors	exactly one	Exactly one bubble must be selected
sti screening	STI Screening	exactly one	Exactly one bubble must be selected
side effects	Any Side Effects	exactly one	Exactly one bubble must be selected
condom m given	Condoms Male	enough filled	At least 1 box must be filled
condom f given	Condoms Female	enough filled	At least 1 box must be filled
contraception assess	Contraception Assessment	exactly one	Exactly one bubble must be selected
visit outcome	Visit Outcome	exactly one	Exactly one bubble (active if ≥ 2 filled)
appointment date	Next Appointment Date	enough filled	≥ 1 box filled (active if ≥ 2 filled boxes)

5.4 Check Types Explained

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Table 5.3: Client-Side Validation Check Types

Check Type	Description
enough filled	At least one box or bubble in the field group is filled. Prevents blank mandatory fields. Used for digit-box fields (IDs, dates, counts) and single-bubble triggers.
exactly one	Exactly one bubble is filled. Enforces single-select fields. Implemented as a pair of checks: enough answers (≥ 1 selected) AND not too many answers (≤ 1 selected).
not too many	No more than one bubble may be filled. Caps a field at one selection. Used alone when the field is optional but cannot be multi-selected.

5.5 Fields Not Validated Client-Side

The following fields exist on the form but have **no** client-side validation rule. Errors in these fields are caught by server-side DQA instead.

Table 5.4: Fields Not Validated Client-Side

Field	Label	Reason
bp sys	BP: Systolic	Optional clinical measurement; value range enforced server-side.
bp dias	BP: Diastolic	Optional clinical measurement; value range enforced server-side.
sex disagg sub-options	Sex/Pregnancy detail options	Individual checkbox states are not validated individually; the group is validated as exactly one.
prep injectable option	PrEP Injectable	Optional field; CAB-LA injection consistency is checked server-side.
len tabs given	Lenacapavir Tablets	Optional dispensing count; validated server-side.
tdf 3tc given	PrEP Tabs (TDF/3TC)	Optional dispensing count; integer validity checked server-side.
oral prep intention	Oral PrEP Intention	Optional; multi-option consistency checked server-side.
none tabs given	None (Tablets)	Optional single-select; no client-side rule defined.
lubes given	Lubricants	Optional dispensing count; integer validity checked server-side (Warning).
discard left and right	Discard Right & Left Page	Administrative trigger; not subject to data validation.
photo taken	Photo Taken	Administrative trigger; not subject to data validation.
book number page 1	Book Number (Page 1)	Administrative; validated by page-level linkage checks server-side.
page number page 1	Page Number (Page 1)	Administrative; validated by page-level linkage checks server-side.
book number page 2	Book Number (Page 2)	Administrative; validated by page-level linkage checks server-side.
page number page 2	Page Number (Page 2)	Administrative; validated by page-level linkage checks server-side.
discard row	Discard Row	Used as a discard criterion to suppress validation, not a validated field itself.
implementation error comment	Implementation Error Comments	Administrative flag; not validated client-side. Free-text handwriting zone; OCR not attempted, no validation possible.
client name	Name	Handwriting zone; OCR not attempted.
physical address	Physical Address	Handwriting zone; OCR not attempted.
phone number	Phone Number	Handwriting zone; OCR not attempted.

6 Part 2 — Server-Side Data Validation

Server-side DQA checks run automatically each time the dbt pipeline executes against the full submitted dataset. Unlike client-side validation, these checks can compare across records and across forms — catching errors that are only visible at the population level.

Capabilities beyond client-side:

- Cross-references longitudinal records (e.g. detecting the same PrEP ID assigned to two different clients)
- Validates values against external registers (PrEP register, HTS initial register)
- Enforces clinical plausibility rules that require full cohort context
- Applies to all submitted records, including those where client-side warnings were dismissed
- Separates checks into **raw-data-based** (run before type casting) and **clean-data-based** (run after cleaning and cross-form joins)

6.1 Record Drop Logic

Warning

Records are **excluded from the clean dataset** — not deleted. Excluded records remain visible in the DQA dashboard and can be corrected and re-entered.

A record is **dropped from the clean layer** when:

- It has one or more checks with severity **Error**.
- It is structurally unverified, discarded, a duplicate, or missing (these are excluded at the raw model stage before checks run).
- The `discard_row` bubble is marked — such rows are excluded at raw model ingestion.
- Records where the page 2 scan is missing are flagged via the `page_2_entry_url` linkage check.

Note

Pipeline separation: Checks are split into two streams — **Visit Card** and **Visit Log** — which follow the same logic but are applied to their respective register types. The `register_type` column distinguishes between the two in the combined output.

6.2 Error Classification

```
Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.
```

Table 6.1: Severity Levels

Severity	Meaning	Action
Error	The record violates a hard rule — a value is missing, unparseable, clinically implausible, or contradicts another linked record. Records with ≥ 1 Error are excluded from the clean dataset.	Exclude from analysis. Flag for correction and re-entry in DQA dashboard.
Warning	The record is unusual or internally inconsistent but may still be valid. The record is retained in the clean dataset but flagged for human review.	Keep in analysis. Flag for review. No automatic exclusion.

6.3 Full Check Catalogue — Missing Values and Invalid Numbers

Table 6.2: Server-Side Checks: Missing Values and Invalid Numbers

Check	Field	Rule	Severity
prep reg id missing	prep reg id	PrEP Registration ID must not be empty or blank.	Error
visit date missing	visit date	Visit Date must not be empty (value '/'/'20' treated as blank).	Error
age missing	age	Age must not be empty.	Error
hts link id missing	hts link id	HTS Link ID must be present for start/continue PrEP outcomes. Downgraded to Warning if client refused PrEP.	Error / Warning
missing appointment date for continuing client	appointment date	Next appointment date must be present for clients continuing PrEP at facility (excluding event-driven oral PrEP).	Error
age invalid integer	age	Age OCR value cannot be parsed as an integer.	Error
height invalid integer	height	Height OCR value cannot be parsed as an integer.	Error
tabs given invalid integer	tdf 3tc given	PrEP Tabs OCR value cannot be parsed as an integer.	Error
condom m given invalid integer	condom m given	Condoms Male OCR value cannot be parsed as an integer.	Error
condom f given invalid integer	condom f given	Condoms Female OCR value cannot be parsed as an integer.	Error
delivery point invalid integer	delivery point	Delivery Point OCR value cannot be parsed as an integer.	Warning
lubes given invalid integer	lubes given	Lubricants OCR value cannot be parsed as an integer.	Warning
bmi invalid float	bmi	BMI value cannot be parsed as a decimal number.	Warning
creatinine concentration invalid float	crea concentration	Creatinine concentration cannot be parsed as a decimal number.	Warning

6.4 Full Check Catalogue — Invalid Dates and Codes

Table 6.3: Server-Side Checks: Invalid Dates and Codes

Check	Field	Rule	Severity
visit date invalid date	visit date	Visit Date OCR value cannot be parsed as a valid calendar date.	Error
appointment date invalid date	appointment date	Next appointment date OCR value cannot be parsed as a valid calendar date.	Error
update date hiv pos invalid date	update date hiv pos	Update date (HIV positive) cannot be parsed as a valid date.	Error
update date death invalid date	update date death	Update date (death) cannot be parsed as a valid date.	Error

Table 6.3: Server-Side Checks: Invalid Dates and Codes (*continued*)

Check	Field	Rule	Severity
update date side effects invalid date	update date side effects	Update date (side effects) cannot be parsed as a valid date.	Error
visit date before ScanForm implementation	visit date	Visit date is earlier than the ScanForm implementation date (2024-03-18) or is in the future.	Error
appointment date before visit date	appointment date	Next appointment date is earlier than the visit date on the same record.	Error
visit date before register initiation date	visit date	Visit date in the visit card is earlier than the initiation date recorded in the PrEP register.	Error
delivery point invalid code	delivery point	Delivery point code is not one of the valid values: 2, 4, 11, 12, 15, 17, 20, 21.	Warning

6.5 Full Check Catalogue — Multiple / Missing Answers

Table 6.4: Server-Side Checks: Multiple / Missing Answers

Check	Field	Severity
kidney risk no answer	kidney risk	Error
kidney risk too many answers	kidney risk	Error
liver risk no answer	liver risk	Error
liver risk too many answers	liver risk	Error
sti screening no answer	sti screening	Error
sti screening too many answers	sti screening	Error
PrEP option no answer	prep given group	Error
PrEP option too many answers	prep given group	Error
Outcome no answer	visit outcome	Error
Outcome too many answers	visit outcome	Error
sex disagg no answer	sex disagg	Error
sex disagg too many answers	sex disagg	Error
side effects no answer	side effects	Warning
side effects too many answers	side effects	Warning
contraception assess no answer	contraception assess	Warning
contraception assess too many answers	contraception assess	Warning
oral prep intention no answer	oral prep intention	Warning
oral prep intention too many answers	oral prep intention	Warning

6.6 Full Check Catalogue — ID Transcription and Cross-Record Checks

Table 6.5: Server-Side Checks: ID Transcription and Cross-Record Consistency

Check	Field	Rule	Severity
PrEP ID mod-223 check	prep reg id	PrEP ID fails Mod-223 check-digit algorithm. Suppressed if matched via orphan matching.	Error
HTS Link ID Damm check	hts link id	HTS Link ID fails Damm check-digit algorithm. Suppressed if matched via orphan matching.	Error
PrEP ID orphan match suggestion	prep reg id	No exact PrEP register match; orphan matching identified a likely correct PrEP ID.	Error
HTS Link ID orphan match suggestion	hts link id	No exact HTS register match; orphan matching identified a likely correct HTS Link ID.	Error
PrEP ID not unique	prep reg id	Same PrEP ID appears on same visit date linked to clients with different sex or age.	Error

Table 6.5: Server-Side Checks: ID Transcription and Cross-Record Consistency (*continued*)

Check	Field	Rule	Severity
PrEP ID reused on consecutive pages	prep reg id	Same PrEP ID appears on two consecutive pages of the same book — possible copy error.	Warning
sex inconsistency with PrEP register	sex disagg	Sex on visit card/log differs from sex in the linked PrEP register record.	Error
sex inconsistency with HTS initial register	sex disagg	Sex on visit card/log differs from sex in the linked HTS initial register.	Warning
age differs from HTS by > 1 year	age	Age on visit card/log is more than 1 year greater than age in the linked HTS initial register.	Error
age differs from HTS by > 5 years	age	Age on visit card/log is more than 5 years greater than age in the linked HTS initial register.	Error
visit date not equal to HTS initial visit date	visit date	Visit date differs from the HTS initial register by more than 3 days.	Error
multiple visit entries same day	prep reg id	A client has more than one visit entry on the same visit date.	Error
PrEP register not linked	prep reg id	Visit card/log has no linked PrEP register record and the visit month is not the current month.	Error

6.7 Full Check Catalogue — Clinical Logic

Table 6.6: Server-Side Checks: Clinical Logic

Check	Field(s)	Rule	Severity
pregnant female on effective FP	sex disagg / contraception assess	Female pregnant client also recorded as being on effective family planning — clinically inconsistent.	Warning
HIV positive client prescribed PrEP	visit outcome / prep given group	Client with an HIV positive result has been prescribed PrEP.	Error
PrEP not prescribed but continued at facility	prep given group / visit outcome	Visit outcome is start/continue PrEP at facility but prescription is recorded as None.	Error
low risk client given PrEP	prep given group	PrEP prescribed or given to a client referred from HTS with risk category Low.	Warning
HIV positive client has subsequent visit	visit outcome	Client with a stop: HIV positive outcome has a subsequent visit record.	Error
CAB-LA not given despite prescription	prep injectable option	CAB-LA injection was prescribed but not recorded as given.	Warning
CAB-LA given without prescription	prep injectable option	CAB-LA injection was recorded as given but no CAB-LA prescription is recorded.	Warning
kidney risk not assessed when PrEP prescribed	kidney risk	Kidney risk factor was marked as Not Done but PrEP was prescribed.	Warning
liver risk not assessed when PrEP prescribed	liver risk	Liver risk factor was marked as Not Done but PrEP was prescribed.	Warning

6.8 Summary by Category

Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be resized.

Table 6.7: Server-Side Checks by Category

Category	Errors	Warnings	Total
Missing Mandatory Value	5	0	5

Invalid Number	5	4	9
Invalid Date	7	0	7
Invalid Code	0	1	1
Multiple / Missing Answer	10	8	18
ID Transcription Check	4	0	4
Cross-Record Consistency	8	2	10
Clinical Logic	4	6	10
Total	43	21	64

6.9 Check Tag Reference

Table 6.8: Check Tag Reference

Tag	Meaning
missing-value	A required field is empty or contains only placeholder characters.
invalid-number-checks	An OCR-extracted value expected to be numeric cannot be parsed as an integer or float.
invalid-date-checks	An OCR-extracted date value cannot be parsed as a valid calendar date, or falls outside an expected date range.
multiple-answer-selected-check	More than one bubble is marked for a single-select question.
prep-reg-id-transcription-check	The PrEP Registration ID fails the Mod-223 check-digit algorithm.
hts-link-id-transcription-check	The HTS Link ID fails the Damm check-digit algorithm.
orphaned-matching-suggestion	No exact register match found; orphan matching has suggested a probable correct ID.
reused-prep-reg-id-check	The same PrEP ID appears on consecutive pages of the same book.
sex-inconsistency-check	Sex recorded on visit card/log is inconsistent with a linked register or HTS record.
age-difference	Age on visit card/log differs substantially from age in a linked HTS initial register.
mismatched visit date checks	Visit date on visit card/log differs from the date in a linked HTS initial register.
invalid visit date checks	Visit date is in the future or pre-dates ScanForm implementation (2024-03-18).
invalid next appointment date	Next appointment date is earlier than the visit date.
missing next appointment date	Next appointment date is absent for a client continuing PrEP.
orphaned-visit-card	Visit card/log has no linked PrEP register record.
ineligibility-for-prep-prescription	PrEP prescribed to a client who is ineligible (HIV positive or low risk).
cab-la-injection-inconsistency	CAB-LA injection given/not given inconsistently with the prescription record.
prep-id-not-unique	PrEP ID assigned to more than one distinct client on the same visit date.
hiv pos client with next visit	HIV positive outcome client has a subsequent visit record.
multiple-records-used-for-same-visit	A client has more than one visit entry on the same visit date.
pregnant female on fp checks	Pregnant female client is recorded as being on effective family planning.
prep given to low risk client checks	PrEP given to a client with HTS risk category Low.
invalid visit-card delivery-point	Delivery point code is not in the list of valid codes.

Generated automatically from MW-PrEP Visit Log v3.2 source files. Last updated: 2026-06-09.