HD-OLDER — Statistical Analysis Plan (SAP)

Study title:	HemoDiafiltration versus HemoDialysis in OLDER people: a randomized, multicenter, crossover, pragmatic clinical trial
Short title / Acronym:	HD-OLDER
SAP version/date:	v1.0 — 02-Sep-2025
Protocol reference:	HD-OLDER, Protocol v3 (17-Jul-2025)
Sponsor:	Fondation AUB SANTÉ
Principal Investigator:	Dr Mabel AOUN

---

1. Background and rationale

Hemodiafiltration (HDF) and high-flux hemodialysis (HD) are standard dialysis modalities. Very elderly patients (≥85 years) are under-represented in prior trials, and Dialysis Recovery Time (DRT) is a practical, patient-centred endpoint for tolerability. HD-OLDER uses a two-period crossover (HDF↔HD) to enable precise within-patient comparisons in a small, frail population. A two-week low-flux HD washout is implemented at study start and between periods to limit carryover. Pragmatic conduct preserves usual-care practices and enhances external validity.

---

2. Objectives and estimand

2.1 Primary objective

To compare the within-period change in DRT (minutes) between HDF and high-flux HD over each 3-month period.

2.2 Secondary objectives

To compare, between modalities, changes in: fatigue (SONG-Fatigue 0–9), functional performance (STS-30), health-related quality of life (SF-12 PCS/MCS); and counts per period of symptomatic intradialytic hypotension and circuit clotting; changes in laboratory markers and dialysis adequacy (Kt/V); and safety (deaths, hospitalizations).

2.3 Estimand (primary)

- **Population:** All randomized patients ≥85 years who enter the treatment periods.
- **Variable:** Period-level change in DRT: mean(end-of-period window) − mean(start-of-period window), minutes.
- **Intercurrent events: Treatment-policy** strategy (analyse as randomized regardless of adherence, hospitalizations, etc.).
- **Summary measure:** Adjusted mean difference (HDF − HD) from a linear mixed-effects model (LMM) with fixed effects for modality, period, sequence, centre; random intercept for patient; and pre-specified covariates.
- **Population-level interpretation:** Average causal effect of allocating HDF vs HD on within-period change in DRT, in the intended pragmatic setting.

---

3. Hypothesis framework and error control

Confirmatory framework: Superiority for the primary endpoint only.
Null (H0): Mean change in DRT is the same under HDF and HD ($\beta_1 = 0$).
Alternative (H1): HDF reduces DRT ($\beta_1 < 0$).
Type I error: Two-sided $\alpha = 0.05$; report 95% CIs.
Multiplicity: None for secondaries (estimation-focused with 95% CIs and two-sided p-values, descriptive interpretation). No hierarchical testing.

---

4. Study design overview

- **Design:** Randomized (1:1) two-sequence, two-period crossover; sequences A (HDF→HD) and B (HD→HDF).
- **Periods:** Each 3 months.
- **Washout:** Two weeks of low-flux HD at study start and between periods.
- **Blinding:** Not feasible for patients/clinicians; statisticians analyse anonymised treatment codes where possible.
- **Centres:** Ten dialysis units in France.

- **Planned sample size:** 62 randomized (targeting ≥56 evaluable).

---

5. Analysis populations

- **Intention-to-treat (ITT) — primary:** All randomized, analysed as randomized and per period modality, irrespective of adherence or deviations.
- **Per-protocol (PP) — sensitivity:** ITT subset completing both periods; ≥80% of attended sessions on allocated modality in each period; valid start and end windows for the endpoint in the period; exclude major deviations (e.g., inadequate washout, cross-over contamination).
- **Safety set:** All randomized with ≥1 post-randomization dialysis session; safety attributed to modality received within each period (as-treated by session aggregation).

---

6. Outcomes and derivations

6.1 Primary endpoint — DRT (minutes)

- **Question:** "How long did it take you to recover from your last dialysis session?"
- **Schedule:** Start (~week 0), Mid (~6 weeks), End (~12 weeks) in each period. At each window, obtain up to three consecutive sessions.
- **Window rule:** Compute the window mean if ≥2 of 3 session values are available; otherwise the window is missing.
- **Derivation:** Period-level change = mean(End) − mean(Start). Negative values indicate faster recovery (improvement).

6.2 Secondary endpoints

- **SONG-Fatigue (0–9):** Sum of three items (all required). Change = End − Start per period.
- **STS-30:** Mean repetitions over three consecutive sessions per window; Change = End − Start.
- **SF-12 PCS/MCS:** Baseline and End of each period; Change = End − Start.
- **Counts per period:** Total symptomatic intradialytic hypotension; total circuit clotting.
- **Laboratories & adequacy:** Haemoglobin, albumin, potassium, phosphate, Kt/V as priority; other routine labs exploratory. Change = month 3 − baseline month.
- **Safety:** Deaths and hospitalizations per period; exposure-adjusted incidence rates.

7. Sample size and information size

- **Planned N:** 62 randomized, anticipating ≥56 evaluable.

Justification: Prior data in ≥85-year-olds indicate large between-modality differences in mean DRT with wide SDs. The current N achieves ≈80% power (two-sided α=0.05) under conservative parallel-group assumptions. The crossover design is expected to yield greater precision than this conservative calculation by exploiting within-patient comparisons. No interim analysis or re-estimation is planned.

8. General analysis conventions

- **Data cut:** Final database lock after Period-2 ends for the last participant and all queries are resolved.
- **Units:** DRT in minutes (convert hours to minutes).
- **Transformations:** Primary on the minute scale for interpretability; log[DRT+1] sensitivity if residual skewness persists.
- **Rounding:** Descriptives to one decimal where appropriate; effect estimates to one decimal (minutes) and 2–3 significant figures for ratios.
- **Covariates (pre-specified):** Age, sex, dialysis vintage, access type (AVF/AVG vs CVC), baseline DRT at period start.

9. Primary analysis

9.1 Model

Unit of analysis: **patient–period**.
For patient *i* in period *p* (p=1,2):

$$\textbf{Change}_{i,p} = \beta_0 + \beta_1 \cdot \textbf{Modality}_{i,p} + \beta_2 \cdot \textbf{Period2}_{p} + \beta_3 \cdot \textbf{Sequence}_{i} + \beta_4 \cdot \textbf{Centre}_{i} + \textbf{X}'_{i,p}\gamma + u_{i} + \varepsilon_{i,p},$$

with $u_{i} \sim N(0, \tau^2)$ (random intercept), $\varepsilon_{i,p} \sim N(0, \sigma^2)$.
- **Modality**: 1 = HDF, 0 = high-flux HD.
- **Period2**: 1 if period 2, else 0.
- **Sequence**: 1 = HDF→HD, 0 = HD→HDF.
- **Centre**: fixed categorical effect.
- **X**: pre-specified covariates listed in §8.

Target parameter: β_1 = adjusted mean difference in within-period DRT change (HDF − HD), minutes; $\beta_1 < 0$ favours HDF.
Estimation: REML. Report β_1 with 95% CI and two-sided p-value.

9.2 Decision rule

Conclude HDF superiority if $\beta_1 < 0$ and the 95% CI excludes 0. Otherwise, no superiority is concluded.

9.3 Carryover assessment and fallback

- **Definition:** Residual influence of Period-1 modality on Period-2 outcomes despite washout.
- **Screen:** Test **Modality × Period** interaction (diagnostic threshold p<0.10).
- **Fallback if flagged:** Primary inference switches to **Period-1-only** ITT **parallel-groups** ANCOVA of Period-1 change (factors: modality, sequence; covariate: baseline DRT at Period-1 start). Provide Wilcoxon two-sample sensitivity.
- **Rationale:** Preserves unbiased inference if washout is insufficient.

9.4 Diagnostics and robustness

Inspect residuals (fitted vs residuals, Q–Q) and influence (e.g., Cook's distance). If assumptions are violated, use robust (sandwich) SEs and confirm with a robust LMM or rank-based analysis; include the log-scale sensitivity if skewness persists.

---

10. Secondary analyses

10.1 Continuous change endpoints (SONG-Fatigue, STS-30, SF-12 PCS/MCS, labs, Kt/V)

LMM as in §9.1 (fixed: modality, period, sequence, centre; random: patient intercept; covariates per §8). Report adjusted mean difference (HDF − HD) with 95% CI.

10.2 Counts per period (hypotension, clotting)

Mixed-effects negative binomial model at patient–period level (fixed: modality, period, sequence, centre; random: patient intercept; offset = log[number of attended sessions]). If NB fails or over-dispersion negligible, use Poisson with robust SEs or a GEE sensitivity. Report rate ratio (HDF vs HD) with 95% CI.

10.3 Safety (SAEs: deaths, hospitalizations)

Present exposure-adjusted incidence rates (EAIRs) per 100 patient-months by modality/period with 95% CIs. Optionally fit NB/Poisson models with offset = log(time on study in the period). No confirmatory testing.

10.4 Subgroups (exploratory)

Pre-specified interactions of modality with: sex; age (85–89 vs ≥90); dialysis vintage (<2 vs ≥2 years); access type (AVF/AVG vs CVC); baseline fatigue tertiles; centre (e.g., enrolment size); convective volume tertiles (HDF only); dialysate buffer. Report interaction p-values (two-sided). Display forest plot. Interpretation exploratory (no multiplicity adjustment).

---

11. Missing data and exposure

- **Windows:** Require **≥2/3** session values to compute a window mean; otherwise missing.
- **Change endpoints:** Compute period-level change only when both Start and End windows are available; otherwise set period-level change to missing for that period.
- **Model-based handling:** Primary LMMs/NB models use all available patient–period records under MAR.
- **Multiple imputation (robustness):** Impute patient–period level changes using chained equations including modality, period, sequence, centre, baseline covariates, and auxiliary variables (e.g., session attendance, hospitalization indicator). Pool estimates via Rubin's rules.
- **Exposure adjustment:** For counts, include offsets as specified; sensitivity excluding periods with <27 attended sessions (~75% of planned 36 sessions).

---

12. Protocol deviations, adherence, and data cleaning

12.1 Major protocol deviations

- Incorrect eligibility after randomization (e.g., age <85; dialysis vintage <3 months).
- Washout not applied or <14 days at study start or between periods.
- Modality misclassification >20% of attended sessions in any period.
- Missing both Start or End windows for an endpoint in a period.
- Use of non-permitted procedures that materially affect endpoints.

12.2 Adherence summaries

- Sessions attended per period; % sessions on allocated modality; mean convective volume (HDF); session duration; UF rate; temperature.
- Listings of crossovers and reasons (clinical need, access failure, etc.).

12.3 Analysis by adherence

- **Primary:** ITT.
- **Sensitivity:** PP per above.
- **Exploratory:** As-treated using session-level exposure aggregated to period (e.g., ≥80% HDF vs ≥80% HD); IPTW sensitivity if meaningful confounding is suspected.

---

13. Interim analyses and data monitoring

No interim analyses or formal stopping rules are planned. Safety is monitored by the DSMB per protocol.

---

14. Software, reproducibility, and deliverables

- **Software:** Primary analyses in R (v4.3+).
- **Reproducibility:** Analysis code (R scripts, versions, sessionInfo) and TFL generation pipelines will be archived in the study repository with date/time stamps at database lock.
- **Deliverables:** TFLs in PDF/RTF; analyzable datasets and metadata (Analysis Data Model) retained per sponsor policy.

---

15. Tables, Figures, and Listings (TFLs) — overview

- **Participant flow and disposition** (CONSORT-style for crossover).
- **Baseline characteristics** by sequence and overall (Period-1 baseline).
- **Treatment exposure/adherence** per period and modality.
- **Primary endpoint:** Model summary (β_1, 95% CI, p), adjusted means by modality, diagnostic plots.
- **Secondary endpoints:** Change summaries and model outputs; counts with EAIRs.
- **Subgroups:** Forest plot of interaction effects.
- **Safety:** Listings of SAEs by modality/period; EAIR tables.
