Statistical Analysis Plan (SAP)

Exploring the Hidden Burden of Living with Glanzmann Thrombasthenia: A Qualitative Study

Sponsor: EquiPath Analytics Limited
Funder: Hemab ApS
PIs: Amy Owen-Wyard & Stacey McGeown
Protocol Version: 1.0 (15 December 2025)
SAP Version: 1.0
SAP Date: *27.01.2027*
Study Type: Non-interventional, qualitative, observational

1. Purpose of the SAP

This Statistical Analysis Plan describes the qualitative data analysis methods and procedures that will be followed for the study. Although no statistical testing will be performed, this SAP ensures:

- Transparency of analytical processes
- Consistency across coders
- Rigorous qualitative methodology
- Protection against analytic bias
- Compliance with regulatory, ethical, and sponsor requirements

This SAP aligns fully with the protocol and does not supersede it. In the event of discrepancies, the protocol prevails.

2. Study Objectives

Primary Objective

To explore the emotional, psychological, social, and practical experiences of adults living with Glanzmann Thrombasthenia (GT).

Secondary Objectives

- Describe demographic characteristics of participants (age range, gender identity, region, caregiver role).
- Understand caregiver burden, emotional labour, coping, and role strain.
- Identify unmet support needs and systemic barriers.

- Generate insights to inform patient-centred interventions and communication materials.

3. Analysis Populations

3.1 Qualitative Analysis Set (QAS)

All interviews with adequate audio quality or sufficiently detailed notes will be included, provided informed consent was obtained.

3.2 Exclusion Criteria for Analysis

Interviews will be excluded if:

- Consent was withdrawn *before* anonymisation
- The recording is technically unusable and notes are insufficient
- Ethical or safeguarding concerns require exclusion (extremely rare)

A log of excluded interviews will be maintained.

4. Data Sources for Analysis

- Semi-structured interview transcripts (primary data source)
- Interviewer field notes
- Reflexive journals
- Demographic information (age range, gender identity, caregiver status, region)
- Analytic memos from coding meetings

No quantitative clinical data will be analysed.

5. Data Handling and Preparation

5.1 Transcription

- Verbatim transcription by GDPR-compliant service
- Checked against audio by researcher
- Anonymised immediately
- Audio deleted post-verification

5.2 Anonymisation Procedures

- Removal of names, exact locations, hospitals, schools, workplaces
- Removal of rare contextual identifiers that may risk re-identification
- Replacement with neutral markers (e.g., “[city]”, “[hospital]”)
- Assignment of unique study IDs (e.g., GT001, CG010)

5.3 Data Management Tools

- MaxQDA (or equivalent) for coding
- Encrypted server storage
- MFA access protection

6. Analysis Approach

This is a **thematic analysis** following Braun & Clarke’s six-phase framework. The analysis will be inductive (data-driven), with optional deductive structuring for research objectives.

6.1 Step-by-Step Analysis Plan

Phase 1: Familiarisation

Researchers will:

- Read each transcript against audio
- Document initial impressions
- Note emotional tone, contradictions, and emerging ideas
- Update reflexive journal entries

Output: Initial analytic memos.

Phase 2: Generating Initial Codes

- Line-by-line coding in MaxQDA
- Codes created inductively (e.g., “fear of work absences,” “bleeding unpredictability”)
- Coding performed independently by trained analysts
- Minimum 20% of transcripts double-coded to assess consistency

Coding Rules:

- Codes applied to smallest meaningful units of text
- Overlapping codes permitted
- New codes added iteratively

Output: Preliminary codebook.

Phase 3: Searching for Themes

Codes are grouped into candidate themes by examining:

- Conceptual similarities
- Emotional patterns
- Shared social or systemic challenges
- Differences between GT patients and caregivers

Potential themes may include:

- "Living with constant uncertainty"
- "Caregiver vigilance and emotional exhaustion"
- "Healthcare system mistrust"

Output: Thematic map (Version 1).

Phase 4: Reviewing Themes

Themes are refined at two levels:

1. **Within-theme coherence**
   – Do extracts align and tell a consistent story?
2. **Across-theme distinctiveness**
   – Are themes overlapping or redundant?

Inclusion/exclusion decisions will be documented.

Output: Thematic map (Version 2).

Phase 5: Defining and Naming Themes

Each theme will include:

- Clear definition

- Rationale for inclusion
- Subthemes (if relevant)
- Boundary statements describing what is *not* included
- Illustrative example quotes (anonymised)

Output: Final codebook + theme definitions.

Phase 6: Producing the Final Report

The final narrative will:

- Present each theme with evidence
- Integrate participant quotations
- Draw distinctions between patients and caregivers
- Identify cross-cutting patterns
- Compare findings with existing literature
- Explicitly answer the research question

Output: Final Qualitative Report & Summary Deliverables.

7. Reliability and Validity Procedures

7.1 Inter-coder Agreement

- 20% of transcripts double-coded
- Discrepancies discussed and consensus reached
- Not quantified statistically (qualitative study)

7.2 Audit Trail

Maintained for:

- Codebook iterations
- Reflexive journals
- Thematic maps
- Decision logs
- Coding meeting minutes

7.3 Reflexivity

All analysts:

- Keep ongoing reflexive journals
- Reflect on assumptions, positionality, emotional reactions
- Discuss insights during analysis meetings

7.4 Data Saturation

Saturation defined as:

- No new meaningfully different codes emerging in later interviews
- Themes stable and well-supported across interviews
  Documented through saturation grid.

8. Handling Demographic Information

Demographic variables will be summarised descriptively:

- Age ranges (not exact ages)
- Gender identity distribution
- Participant type (GT adult, caregiver)
- Broad geographic region

Reported using frequency counts and percentages only (no statistical testing).

Used solely to contextualise qualitative findings.

9. Deviations From the SAP

Any deviations will be:

- Documented
- Justified
- Approved by the PI
- Included in the final report

Examples may include expansion of sample size if saturation not achieved.

10. Data Integration and Reporting

10.1 Outputs

- Codebook
- Thematic map
- Interim findings slide deck
- Final qualitative report
- Plain-language summary
- Manuscript for publication

10.2 Quotation Use

- Fully anonymised
- Only representative, non-identifiable excerpts
- Checked for re-identification risk in a rare disease context

11. Archiving and Retention

- All anonymised analytic data stored for **10 years**
- Audio recordings destroyed after transcription quality checks
- Master ID file stored separately and encrypted
- Destruction performed according to Sponsor policy

12. Roles and Responsibilities

Principal Investigators

- Final approval of analytic approach and outputs
- Oversight of data integrity and reflexivity practices

Qualitative Analysis Team

- Coding
- Thematic development
- Validation and triangulation
- SAP adherence

Sponsor Data Governance Lead

- Oversight of data security
- No involvement in analysis or interpretation

Funder (Hemab ApS)

- Receives only anonymised outputs
- No role in analysis or interpretation

13. SAP Approval

To be signed by:

- Principal Investigator
- Sponsor Representative