# *A disease never seen here: Smallpox and family reconstruction at the Cape, 1713–1714*: supplementary appendices*

Johan Fourie†

Sections, figures, tables, and equations numbered without a letter prefix are in the main paper.

## A Census definitions and sensitivity

### A.1 Source version and analytical units

The analysis starts from the Cape household project's July 2026 version 1 release, specifically its cleaned tax-unit file and sheet and schema registries. The wider release contains several districts and centuries. The present analysis selects only the early Cape and Stellenbosch sources, retains their source-row identifiers, and uses observed sheets between 1708 and 1718 for the plotted comparisons. The saved extract also preserves available observations from 1700 to 1720. The registries identify these as administrative tax units; residential household membership is not verified. No person-level linkage across years is assumed. The project's experimental name-linked family material is not used to identify deaths, migration, or relationships.

The eight population columns are mapped to settler men, women, sons, and daughters and enslaved men, women, boys, and girls. The source labels are historical administrative categories. They do not provide exact ages, infection histories, verified kinship to the head, or a complete enumeration of the district population. Company employees and the Company's own enslaved population are not added to these private returns. The census indexes credit Hans Heese and give archival references; the broader transcription history is described by Fourie et al. (2024), published online in 2025.

The shared release omitted the historical heading *Meijsies* in mapping enslaved girls. The analysis restores this field from the underlying local workbook where needed, preserving blank cells as missing. The row-level override ledger contains 3,332 records, including blanks. A reconciliation of numeric sums against the local workbooks finds one other difference in the plotted period: the canonical 1708 Stellenbosch extraction counts a second male entry associated with the Du Plaisi

*This is the supplementary material for Fourie, Johan. 2026. "A disease never seen here: Smallpox and family reconstruction at the Cape, 1713–1714." Working Paper, Department of Economics, Stellenbosch University. This paper was created with the help of several models of Anthropic's Claude and OpenAI's Codex. Cite the paper, not this supplement.

†Department of Economics, Stellenbosch University. Email: johanf@sun.ac.za.

row. The series retains that documented canonical treatment, adding one adult man relative to the simpler local extraction. Neither correction changes a shared source file.

The analysis comprises 6,029 tax-unit observations over the plotted district-years. It reports sums of entered numeric counts and divides them by recorded units only where the denominator is explicitly a tax-unit count. An empty cell is never an imputed zero in the stored analytical data. Counting units with a positive child entry establishes documented presence, not the complement of childlessness. The focal 1712 and 1714 sheets have no explicit numeric zeros in the eight population fields; their blanks consequently deserve particular caution. Two uncertain sons entries in the 1714 Stellenbosch sheet remain missing rather than becoming zero or guessed counts.

Table A1: Recorded population levels in the plotted district-years

| District | Year | Units | Settler adults | Children | Enslaved adults | Children |
|---|---|---|---|---|---|---|
| Cape | 1708 | 312 | 475 | 393 | 1028 | 113 |
| Cape | 1709 | 298 | 455 | 409 | 1053 | 136 |
| Cape | 1711 | 289 | 438 | 445 | 1116 | 163 |
| Cape | 1712 | 290 | 456 | 414 | 1098 | 143 |
| Cape | 1713 | 189 | 277 | 213 | 565 | 58 |
| Cape | 1714 | 271 | 404 | 274 | 1095 | 108 |
| Cape | 1716 | 377 | 533 | 361 | 1365 | 145 |
| Cape | 1717 | 319 | 437 | 327 | 1082 | 91 |
| Cape | 1718 | 353 | 504 | 365 | 1329 | 135 |
| S–D | 1708 | 317 | 459 | 490 | 432 | 58 |
| S–D | 1709 | 382 | 538 | 530 | 526 | 81 |
| S–D | 1710 | 291 | 441 | 612 | 522 | 88 |
| S–D | 1711 | 291 | 444 | 619 | 509 | 88 |
| S–D | 1712 | 366 | 539 | 558 | 628 | 89 |
| S–D | 1713 | 307 | 429 | 494 | 722 | 54 |
| S–D | 1714 | 271 | 398 | 412 | 631 | 68 |
| S–D | 1716 | 349 | 474 | 485 | 849 | 75 |
| S–D | 1717 | 371 | 518 | 567 | 900 | 92 |
| S–D | 1718 | 386 | 534 | 609 | 919 | 98 |

Source: author's extraction from the household-project release, with the documented girls-column correction. Adults and children denote the administrative category sums. S–D denotes Stellenbosch–Drakenstein. Missing years are not interpolated. The Cape 1718 return is incomplete: the transcribers note that two folios of the Hague copy were not photographed, so its 1718 levels are understated. The full sex-disaggregated series is supplied as a CSV.

## A.2 What could produce a recorded decline?

Equation 1 makes the identification problem explicit. Births and arrivals can offset deaths; departure, movement into other categories and increased omission can produce a decline without death. The comparison below therefore examines competing mechanisms rather than imposing an arbitrary migration correction.

For an explicit sensitivity calculation, let $R_{c,t}$ and $R_{a,t}$ denote recorded children and adults.

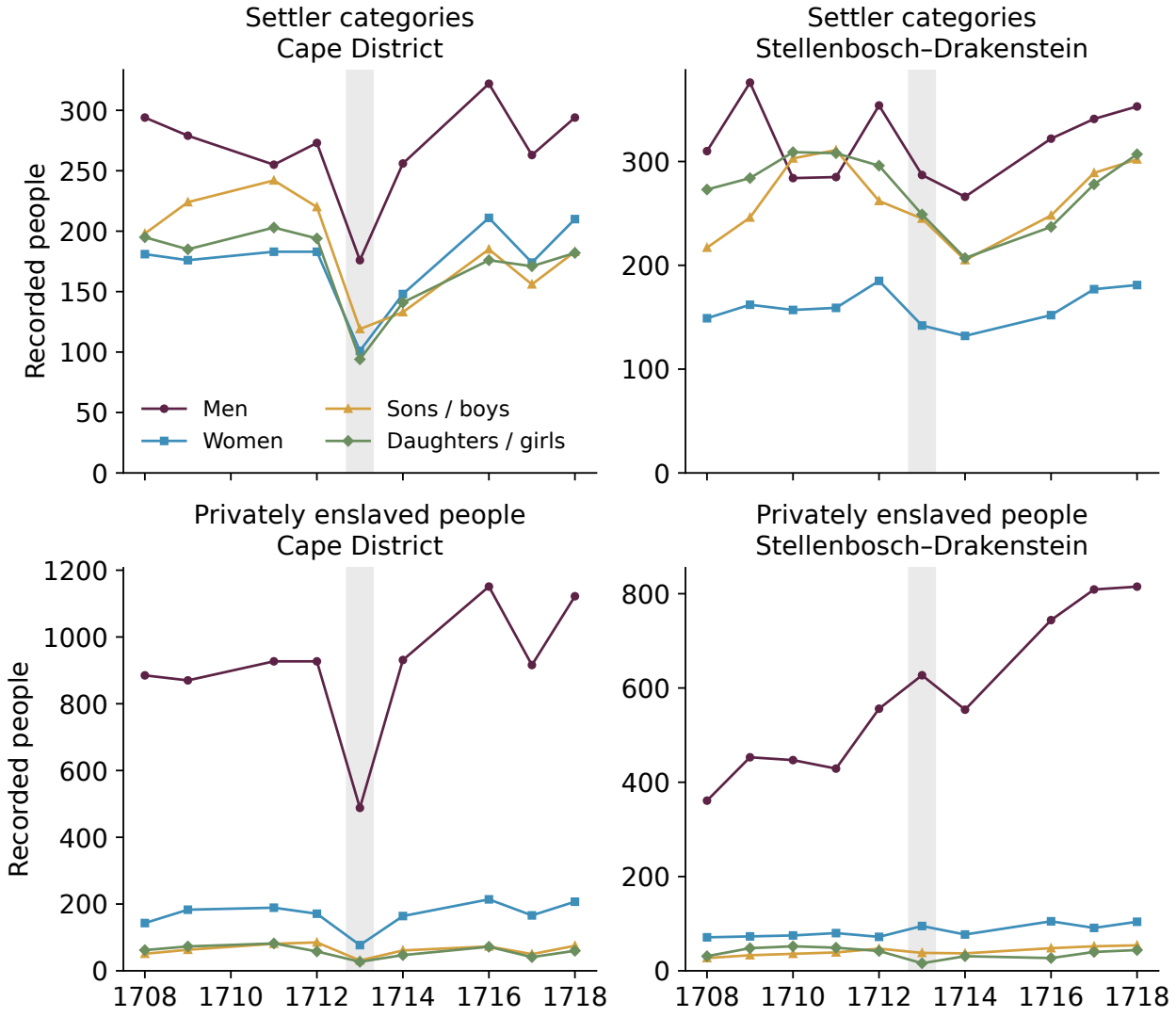



Figure A1: Recorded counts by sex and source category, 1708–1718. The lines join observed returns; they are visual guides across gaps, not interpolated observations. The shaded strip marks 1713.

Source: data as in Table A1. Alt text: Settler adult and child counts both decline between 1712 and 1714. Privately enslaved boys and girls decline in both districts while the much larger adult counts change little. The four panels show different count scales.

Table A2: Missing population cells in the focal returns

| Category | Cape 1712 | Cape 1714 | S–D 1712 | S–D 1714 |
|---|---|---|---|---|
| Settler men | 19/290 | 23/271 | 13/366 | 17/271 |
| Settler women | 107/290 | 123/271 | 184/366 | 139/271 |
| Settler sons | 187/290 | 199/271 | 243/366 | 174/271 |
| Settler daughters | 187/290 | 197/271 | 248/366 | 175/271 |
| Enslaved men | 137/290 | 118/271 | 256/366 | 163/271 |
| Enslaved women | 185/290 | 178/271 | 319/366 | 227/271 |
| Enslaved boys | 251/290 | 246/271 | 337/366 | 249/271 |
| Enslaved girls | 260/290 | 247/271 | 342/366 | 249/271 |

Entries are missing cells divided by all tax-unit rows, shown as counts. They include the two uncertain sons entries. Numeric totals remain the sums of observed counts. The table measures recording patterns and does not determine how many people were omitted.

Table A3: Competing explanations for the recorded changes

| Mechanism | Observable implication and remaining limitation |
|---|---|
| Movement between the two districts | Pooling cancels a transfer only when both declarations are recorded consistently. It does not cancel movement beyond the two districts. |
| Loss of declarations | Tax-unit counts fall in both districts, particularly S–D. Children per unit distinguish part of this compositional change, but do not follow the same families. |
| Ageing and births | Transitions can reduce the child category and increase the adult category without a death. No exact age boundary, birth series or linked cohort resolves their contribution. |
| Age-specific omission | Blank counts and uncertain entries expose limitations. The returns do not establish a constant proportion of unrecorded children. |
| Purchases and transfers | These could offset losses among privately enslaved adults. The Company's documented additions do not establish equivalent private flows. |
| Epidemic mortality | Independently attested in the journal and contemporary estate activity, but not separately identified within the tax-stock changes. |

The difference

$$Q = R_{c,1712}\frac{R_{a,1714}}{R_{a,1712}} - R_{c,1714}$$

measures how many fewer children were recorded than if their count had changed proportionally with adults. It is a descriptive benchmark. It would identify excess child losses only under strong restrictions on differential births, ageing, movement, and recording that cannot be established here.

Table A4: Child shortfall relative to proportional adult change

| District | Settler children | Enslaved children |
|---|---|---|
| Cape District | 92.8 | 34.6 |
| Stellenbosch–Drakenstein | 0.0 | 21.4 |
| Pooled | 97.5 | 56.0 |

Positive values mean fewer recorded children than under the adult-proportional benchmark. Values are rounded; the S–D settler value is 0.03 before rounding. Neither this quantity nor the raw decline is a death estimate.

Available two-year changes before and after the epidemic are preserved in the replication file. They provide a check on the scale and variability of recorded changes rather than a causal counterfactual. Missing sheets and changing administrative coverage preclude a balanced annual panel. No trend is fitted through the missing years and no confidence interval is attached to a census sum. The uncertainty concerns coverage and interpretation, not sampling error from a representative population sample.

# B Estate documents, dates, and family cases

The extraction parses each MOOC8 document division with a dated heading, retaining the source file, archival identifier, heading date, named parties, opening and closing text, and complete text for 1713–1714. All 90 focal headings are included in the document register. Eleven further documents in volumes MOOC8/2 and MOOC8/3, which cover 1705–1720, carry no usable heading date and are excluded from the annual series; a few may belong to 1713 or 1714. Keyword flags for minor children, widowhood, guardianship, and enslaved asset rows help retrieve cases; they are not independently validated counts of orphans or families.

Table A5: Sensitivity of the probate activity comparison

| Baseline | Annual mean | 1713 / mean | 1713–1714 / expected |
|---|---|---|---|
| 1700–1712 | 7.85 | 6.88 | 5.74 |
| 1708–1712 | 10.80 | 5.00 | 4.17 |
| 1709–1712 | 10.75 | 5.02 | 4.19 |

The 1713 count is 54 and the 1714 count is 36. The combined ratio divides 90 by twice the baseline annual mean. These are ratios of dated documents, not relative mortality risks.

Nine explicitly identified groups of repeat estate papers contain 14 additional documents beyond one per group. They include the Basson–Clement, Verweij–Visser, Versch(u)ijr, Diemer, Cruse–

Corssenaer, Elbertsz, Besuijdenhout, and Pretorius–Elberts papers, and the Hijlon–Roomond documents. This is a set of demonstrated repetitions, not a complete entity-resolution exercise. Subtracting 14 from 90 would not yield a verified count of deaths or even of unique deceased estates. One paper may concern a couple; another may concern a living declarant or remarriage; dates within papers differ from heading dates.

The retained examples distinguish these possibilities. The De Jager document is headed 10 March 1714 but contains a declaration of 8 December 1713 by a living man. The Oberholster document of 15 January 1714 concerns remarriage and protection of existing property claims, without dating his first wife's death. The Pretorius–Elberts papers follow an estate inventoried in 1713 into division in January 1714. The Cruse–Corssenaer papers extend from July and August 1713 to an April 1714 annotation. These demonstrate why a high count in 1714 can include delayed and continuing administration; they do not establish that every 1714 paper concerns a 1713 epidemic death.

The machine-readable family-case register supplies the archival references, explicit relationships, transactions, and limits for ten selected cases or document groups. The complete focal text is retained alongside the 90-row audit. These cases were selected for explicit evidence of family arrangements or a useful source limitation. Their selection cannot support percentages of bereaved families, average orphan counts, treatment effects, or claims about typical care. Public transcriptions can be consulted through the GLOBALISE rendering of TEPC MOOC8 volumes 1–5. The local text has been compared with that public rendering; this is not a new paleographical validation of the estate manuscripts.

# C Journal provenance and worked passages

## C.1 Source structure and retrieval

The supplied workbook has one row per calendar date from 1700 to 1720, including leap days, with 7,670 distinct identifiers and dates. Four rows in 1702 lack usable text. The 7,666 usable entries contain 857,732 whitespace-delimited words after markup removal. Matching calendar days defines the supplied file's structure, rather than independently establishing archival completeness. The dagregister transcription was produced and supplied by the Tracing History Trust and shared by Helena Liebenberg. The analysis includes the full supplied corpus, with selected interpretively important passages additionally anchored to original images.

The full-corpus pipeline strips markup, normalises whitespace and preserves each identifier and calendar date. It then applies overlapping, case-insensitive retrieval expressions to every usable text. The smallpox search includes words containing *pok*, *pock* or *pocq*, and the stem *variol*. Other expressions cover historical variants of illness, death, hospitals, medicines, surgeons, healing, prescriptions, recovery, quarantine, and selected substances. The exact expressions are saved in the retrieval protocol and script. Refinements made during inspection include the historical *genesinge* and compound *opperchirurgijn*; overly broad matches to *geneegen* and *receptie* were removed. These

searches were developed for this investigation, not preregistered or assessed on a held-out test set.

Every broad smallpox candidate was inspected in context. In all, 18 entries match; 17 refer to smallpox, while the remaining entry contains *capok* in a criminal proceeding. Of the retained entries, 16 belong to 1713 and one to February 1714. The author verified every decision by hand. The ledger records the evidence and decision for every candidate, including the 3 May *gepokt* reference, which a search for *kinderpokjes* alone would miss. This is complete review of the retrieved candidate set, not proof that no implicit or differently transcribed reference remains elsewhere. An entry counts once even if it contains several references; retrospective discussion, prior exposure and consequences count alongside current illness. The series therefore measures explicit references, not daily incidence.

The expanded medical search retrieves 94 entries across the 21 years. Retrieved passages were screened in context, with full-entry follow-up for the examples cited in Section 8. The register preserves each passage and the interpretive notes. Relevant controls include medicine dispatches in 1704, treatment charges in 1708, prescription registers in 1710–1711, and hospital assistance in 1711. The 29 April 1713 reference to quarantine concerns plague and a ship in European waters; the 6 May reference to healing is the clerk's attribution about Khoesan knowledge. Neither records the selective import of smallpox medicine. The conclusion is that no corroborating passage was located for that specific account, with the journal's recording practices and the existence of separate medical records explicitly considered.

Annual frequencies in Table A6 are deterministic counts in the supplied text. Their denominators exclude missing texts. No sampling interval is attached: the uncertainty concerns transcription, retrieval, and interpretation rather than a random sample of these entries. Illness and death matches have not been promoted into exhaustive semantic labels. Death vocabulary can concern criminal proceedings or earlier events, and phrases for a dead calm (*dood stil*, *doodelijke stilte*) are excluded from the death count; *krankbesoeker* can designate an office. Changes in entry length and document content likewise affect the opportunity for a match. The 557 entries containing both illness and maritime vocabulary are co-occurrences, not a measure of the share of sickness affecting sailors. These distinctions permit full-corpus measurement without converting lexical matches into semantic estimates.

The 11-entry image concordance covers 17 February 1710, 13 April 1711, 6, 7, and 19 May, 13 June, 26 August, 28 November, and 17 December 1713, and 13 and 15 February 1714. The originals are NA, VOC archive 1.04.02, inventories 10730 (1710), 10731 (1711), 10732, and 10733. Saved viewer metadata and image manifests connect scan numbers to image identifiers. Each image/text comparison was made with AI assistance and then checked by hand by the author against the scans. The portion of the 15 February entry naming the first two successors was not located in the image sequence used here; scan 56 shows the nephew/heir and son/heir appointments, and the two brothers rest on the supplied transcription and Leibbrandt's abridgement. The 1710 prescription instructions are visible in inventory 10730, scan 62, with the date on scans 58 and 61. The 1711 resolution begins on inventory 10731, scan 162; scan 165 records the 16 enslaved hospital workers

Table A6: Full-corpus coverage and annual entry counts

| Year | Texts | Words | Illness | Death | Hospital | Smallpox |
|---|---|---|---|---|---|---|
| 1700 | 365 | 42,366 | 22 | 23 | 8 | 0 |
| 1701 | 365 | 24,838 | 8 | 12 | 2 | 0 |
| 1702 | 361 | 29,860 | 16 | 19 | 5 | 0 |
| 1703 | 365 | 28,001 | 16 | 18 | 4 | 0 |
| 1704 | 366 | 49,727 | 29 | 41 | 13 | 0 |
| 1705 | 365 | 55,944 | 49 | 54 | 18 | 0 |
| 1706 | 365 | 61,752 | 34 | 55 | 16 | 0 |
| 1707 | 365 | 69,511 | 40 | 72 | 19 | 0 |
| 1708 | 366 | 60,160 | 34 | 52 | 15 | 0 |
| 1709 | 365 | 65,502 | 45 | 59 | 19 | 0 |
| 1710 | 365 | 58,493 | 45 | 52 | 12 | 0 |
| 1711 | 365 | 56,964 | 44 | 45 | 16 | 0 |
| 1712 | 366 | 23,872 | 25 | 31 | 8 | 0 |
| 1713 | 365 | 41,650 | 39 | 48 | 9 | 16 |
| 1714 | 365 | 37,713 | 33 | 47 | 14 | 1 |
| 1715 | 365 | 27,036 | 28 | 37 | 15 | 0 |
| 1716 | 366 | 35,941 | 22 | 28 | 22 | 0 |
| 1717 | 365 | 25,332 | 29 | 36 | 11 | 0 |
| 1718 | 365 | 20,565 | 27 | 30 | 10 | 0 |
| 1719 | 365 | 18,747 | 21 | 36 | 6 | 0 |
| 1720 | 366 | 23,758 | 38 | 43 | 8 | 0 |

Illness, death, and hospital columns count entries matching the saved lexical expressions; categories overlap. Smallpox counts references after contextual inspection of all candidates. Words are whitespace-delimited after markup removal. Source: Tracing History Trust transcription. These counts describe the supplied corpus and do not estimate the completeness of its archival coverage.

and scan 166 the prescription-book requirement. The remaining institutional examples in the medical register rely on the Tracing History Trust transcription and are identified as such.

### C.2 Deaths counted "at this place"

The supplied text for 13 June 1713 reads:

> Heden kon men berekenen datter seder primo April 160 zoo mannen, vrouwen als kinderen door de kinderpokjes waren wegh gerukt allen op deese plaats.

A close translation is: "Today it could be calculated that since the first of April 160 men, women, and children had been carried off by smallpox, all at this place." The image contains the count, the start date, and the local qualification (inv. 10732, scans 135–136). The reading establishes a reported cumulative death count. It does not identify an entire-colony population or an infection denominator. Leibbrandt's abridgement places the count under 11 June (Leibbrandt, 1896, p. 256); the original dated journal sequence supports 13 June. This difference matters when specifying the observation window.

### C.3 The clerk's account of healing

The entry of 6 May 1713 describes Khoesan people as "dese ziekte niet kennende als noijt meer gesien en bij gevolge derselver geneesingh onkundigh" (inv. 10732, scans 114–115). The passage can be translated as "not knowing this disease, having never seen it before, and consequently being unfamiliar with its healing". The spelling and phrasing are reproduced from the supplied transcription; the substantive wording is visible in the image. The preceding sentence of the same entry concerns 13 white bodies awaiting burial, "alle hier geboorne of die de pokjes noit te vooren gehad hebben" (all born here or never previously having had smallpox). The same distinction appears on 3 May, when illness increased among "de noijt gepokt hebbende Affricanen". In the journal's usage *Affricanen* and *inboorlingen* denote people born at the Cape, including settler children, rather than Africans or Khoesan in general; on 24 June the clerk applied the term to the inhabitants of Drakenstein. These passages are therefore not counted as references to Khoesan people.

This is the clerk's attribution, rather than testimony directly recorded from a healer or patient. Translating *geneesingh* as healing or cure leaves open the practices he had in mind. The article distinguishes that contemporary belief from medical efficacy and from the separate, investigable claim about importing and selectively distributing medicines. The broader corpus establishes that medicines and organised care existed at the Cape; it has not yielded corroboration of the specific 1713 import-and-withholding account.

### C.4 Four communities and the renewal of leadership

The entry of 13 February 1714 refers to "het overschot van vier craalen (versamelingen)" and states that, according to the people reporting, "van de 10 persoonen hunner geselschap een over gebleeven

was" (inv. 10733, scans 51–52). The first phrase means the survivors or remainder of four kraals or communities; the second says that one of ten people in their company had remained. The entry connects the deaths of the captains to a request for four successors. The fraction is therefore local testimony, and the event is also a political initiative by survivors.

The entry of 15 February, which refers back to the delegation "waar af onder den 13:en deser gesprooken is", names four successors, each as heir of a dead captain: "Scipio Africanus, broeder en erfgenaam van den overledene capitein Asdrubal; Hanibal, broeder en erfgenaam van Jazon; Hercules, broeders zoon en erfgenaam van den gewesen capt: Hartloop; Kolinga, zoon en erfgenaam van Koeinga den Oude". Two brothers, a brother's son and a son thus succeeded. Scan 56 shows the last two appointments; Leibbrandt (1896, p. 258) lists all four. That evidence supports the discussion of kinship and succession in Section 7. It does not justify a population-wide Khoesan death rate. Leibbrandt groups the episode under 15 February (Leibbrandt, 1896, p. 258); the original separates the delegation and appointments. The chronology here follows that distinction.

### C.5 Company stock and flows

On 14 May 1713 the journal recorded that the Company had lost about 95 enslaved people in two months (vc-1713-134). The 26 August entry reports 441 Company-owned enslaved people, 127 fewer than previously, with loss attributed chiefly to smallpox. It also states that recruitment from Ceylon and local births brought the year's loss to nearly 200 once additions were allowed for (inv. 10732, scans 176–177). The prior stock of 568 is arithmetic: $441 + 127$. Additions would be approximately 73 only under the journal's simplified stock-flow accounting and its approximate loss figure. Neither the precise timing of additions nor the full exposed population is supplied. The ratio of approximate losses to the prior stock is consequently not a cohort mortality probability, and no infection count permits case-fatality estimation. The separate flow ledger preserves the wording, arithmetic, and assumptions.

### C.6 Flight, harvesting, and marriage

Three further image checks anchor the discussion of flight, labour, and marriage. The 19 May 1713 entry (inv. 10732, scan 121) records the reported flight of some of the remaining Cape Khoesan across the mountains and the killing of all but one of them by another group fearing infection (*ten merendeele, behalven een die echappeerde*). Its conditional wording is preserved in the interpretation. The 28 November entry (scan 260) states that smallpox had not entirely ceased, with people still ill in Drakenstein, and links the use of scythes to the loss of "het meerendeel der Hottentots", the majority of the Khoesan workers farmers had employed. This supports ongoing illness and labour adjustment, not an estimated fall in output or proof of permanent technical adoption.

The 17 December entry (scan 275) records the 32nd couple married since 21 July and connects renewed marriage to unions broken by smallpox. The unit is a married couple, not a death or a demonstrated remarriage of two widowed people. There is no comparable marriage series in this

analysis from which to estimate excess marriages. Its historical importance is the clerk's explicit connection between marriage and recovery, which the estate cases illuminate at the level of property and kin arrangements.

# D Computational protocol

## D.1 Tasks and checks

The full-corpus analysis uses the deterministic retrieval and contextual inspection described in Appendix C. Scripts export the complete entry ledger, exact expressions, all matched passages, annual counts, the 18 smallpox decisions, and medical examples. Neither handwriting recognition nor a trained semantic classifier is required to reproduce the lexical results from the supplied transcription. The contextual decisions were made with AI assistance, checked by the author, and saved with the passages on which they rest.

A hand-coded sample shows why the smallpox series covers every candidate. Two research assistants were given the same random sample of 79 journal entries to code independently for references to human illness or death, to smallpox, and to epidemic treatment. One completed 79 entries and the other 77. Their illness-or-death codes agreed with an AI-assisted reference coding on 74 of 79 and 76 of 77 entries, respectively (94 and 99 per cent); the disagreements concern a homicide, an individual illness, and routine hospital admissions. Neither reading found a smallpox reference in the sample. A sample of this size therefore cannot measure how reliably a search or classifier recovers so rare an event, and the analysis reads all retrieved smallpox candidates in context instead.

## D.2 Model choice, historical expertise, and costs

Text retrieval, semantic classification, and the recognition of handwriting are different tasks. Gonzalez Garcia and Weilbach (2023) examine language models as historical research assistants; their work does not validate the present labels. Li et al. (2025) study compact models for industrial classification, showing why model size should be considered in relation to a task and its training evidence. Semnani et al. (2025) instead develop a specialised model for historical text recognition. Its relevance is the need to evaluate handwriting recognition on appropriate documents, not a reason to assume that an off-the-shelf text model can read early modern Dutch accurately.

The model-size question depends on the task. The lexical counts are reproducible without a model; the interpretive assistance has no benchmark establishing superiority over a smaller model or a trained historian. Such a comparison would require independent reference judgments with relevant positive examples, a test set separated from tuning, documented model versions, and comparable error analysis. The article's contribution rests on its historical findings and inspectable evidence rather than a model-performance claim.

Computational assistance also carries costs that accuracy alone does not settle. Proprietary services can limit reproducibility and make scholarship dependent on changing commercial access.

Data selection and model training can carry linguistic and social biases into a colonial record already shaped by power. The work of transcribers, archivists, and readers must remain visible rather than being redescribed as a saving achieved by automation. Luccioni, Jernite and Strubell (2024) further show that deployment energy varies with the task and system. No energy measurements exist for this project, so no carbon-saving or low-impact claim is made. The defensible use here is bounded assistance with auditable outputs and explicit uncertainty, with historical interpretation and publication responsibility remaining with the author.

## E Purpose and limits of comparison

The article does not rank epidemics numerically. The available quantities differ in their populations, observation windows, and denominators, and no ordering would make them commensurable. The comparisons instead serve specific historical questions.

Schneider and Davenport's (2026) Swedish and Icelandic evidence helps separate case fatality from population mortality and provides a discussion of how care might matter when illness is simultaneous. The Cape does not have equivalent case counts. Muurling, Riswick and Buzasi's (2023) Amsterdam study concerns 1870–1872, in a different institutional and preventive setting; it is not treated as an early eighteenth-century control. The Leiden study (van Besouw, Curtis and van Oosten, 2026) examines epidemic mortality and property redistribution in 1635, but does not establish smallpox as the cause of that crisis; it is therefore not used as a smallpox comparison.

Kelton (2015) adds a colonial and Indigenous comparison. The relevant questions concern Indigenous responses, colonial disruption and the inadequacy of explaining losses by prior exposure alone. This comparison is qualitative and does not transfer a Cherokee mortality estimate or a uniform colonial mechanism to the Khoesan. The Cape contribution is its combination of fiscal stocks, estate arrangements and locally attributed testimony, with the different limits of each made explicit.

## References

**Fourie, Johan, Erik Green, Christiaan Burger, Chris de Wit, Kate Ekama, Jan Greyling, Hans Heese, Jan-Hendrik Pretorius, Robert Ross, and Dieter von Fintel.** 2024. "'Waar is beter dorp in Zuid Africa dan Stellenbos?': What the Stellenbosch–Drakenstein Tax Censuses Reveal." *South African Historical Journal*, 76(4): 420–446.

**Gonzalez Garcia, Giselle, and Christian Weilbach.** 2023. "If the Sources Could Talk: Evaluating Large Language Models for Research Assistance in History." *arXiv preprint 2310.10808.*

**Kelton, Paul.** 2015. *Cherokee Medicine, Colonial Germs: An Indigenous Nation's Fight against Smallpox, 1518–1824.* Norman:University of Oklahoma Press.

**Leibbrandt, H. C. V.** 1896. *Précis of the Archives of the Cape of Good Hope: Journal, 1699–1732.* Cape Town:W. A. Richards & Sons, Government Printers.

**Li, Lujun, Lama Sleem, Niccolò Gentile, Geoffrey Nichil, and Radu State.** 2025. "Small Language Models in the Real World: Insights from Industrial Text Classification." 971–982. Association for Computational Linguistics.

**Luccioni, Sasha, Yacine Jernite, and Emma Strubell.** 2024. "Power Hungry Processing: Watts Driving the Cost of AI Deployment?" 85–99. Association for Computing Machinery.

**Muurling, Sanne, Tim Riswick, and Katalin Buzasi.** 2023. "The Last Nationwide Smallpox Epidemic in the Netherlands: Infectious Disease and Social Inequalities in Amsterdam, 1870–1872." *Social Science History*, 47(2): 189–216.

**Schneider, Eric B., and Romola Jane Davenport.** 2026. "What Is the Case Fatality Rate of Smallpox?" *Population Studies*.

**Semnani, Sina J., Han Zhang, Xinyan He, Merve Tekgürler, and Monica S. Lam.** 2025. "CHURRO: Making History Readable with an Open-Weight Large Vision-Language Model for High-Accuracy, Low-Cost Historical Text Recognition." 34765–34812. Association for Computational Linguistics.

**van Besouw, Bram, Daniel R. Curtis, and Roos van Oosten.** 2026. "Epidemic Mortality and Redistribution of Housing Wealth in Leiden, 1632–1668." *Explorations in Economic History*, 101: 101768.