Supplementary Methods and Results

Children's Articulation Practice in SpeechLP Across Two Months · November 3, 2026

These methods and results accompany the main manuscript. Tables S1–S8 and Figures S1–S4 report the complete subgroup and sensitivity analyses. Aggregate exports retain the original Table 1–15 identifiers; the crosswalk in S8 maps them to the current labels.

These findings are preliminary and descriptive. Acceptance rates reflect the app's automated speech recognition decisions, which have not been validated against SLP ratings, and may also reflect changes in app scoring over time. They should not be interpreted as clinical articulation accuracy, within-platform growth, or evidence of treatment efficacy.

A primary goal in publishing these findings is to seek university research partners to collaborate on future research.

S1. Bootstrap implementation

Nominal 95% intervals for mean paired changes use NumPy default_rng(20261014) and 10,000 resamples of whole profiles, with profiles indexed by ascending SHA-256 profile key. Percentile endpoints use numpy.percentile at 2.5 and 97.5 with linear interpolation. Each subgroup starts a fresh generator and uses the same profile draws for all paired periods. Each between-group contrast starts a fresh generator and draws the complete higher-frequency resample array before the lower-frequency array, preserving each group's observed size. Resampling whole profiles preserves the matched month pairing and dependence between sounds contributed by the same profile; for the pair-weighted mean, each resampled profile retains its number of sounds. Changing profile order or random-number consumption can change finite-bootstrap endpoints by about a tenth of a point.

S2. Carried-over interval for 10+ first-month sessions

The 10+ session subgroup in Table S2 retains the interval generated in the earlier three-group analysis (1–5, 6–9 and 10+ sessions), whose resampling stream was shared across three session groups: +3.2 (-1.4 to 7.7). A same-procedure recomputation for this group alone gives -1.3 to 7.6; the difference is Monte Carlo variation. All other first-month group and between-group intervals use the procedure in S1.

S3. Analysis sources and computational reconstruction

Accuracy aggregates come from the October 14, 2026 sound-month analysis, with descriptive practice summaries restricted to the same 101-profile membership. A separate AI-assisted computational reconstruction from the source records reproduced cohort membership, monthly card, session and day summaries, sound means and direction counts. Separate reconstructions of the first-month frequency comparisons and the Day 1 sensitivity reproduced membership, grouping, point estimates, group-gain intervals and between-group gain intervals. The carried-over 10+ session interval was reproduced using the earlier shared stream (S2). Source records and analysis scripts are retained privately and are not publicly available.

S4. Cohort selection and analysis scope

The frozen eligible-card frame contained 60,780 records from 949 recorded profiles. Requiring full 60-day observation opportunity retained 672 and excluded 277. Among those 672 profiles, 559 had no eligible Month 2 practice and 12 returned without a matching sound target. The remaining 101 profiles, from 101 distinct family accounts, contributed 17,675 eligible cards in the two windows, including 11,902 cards on 660 matched profile-and-sound pairs. The five-card sensitivity retained 39 profiles and 171 pairs; the Day 1 sensitivity retained 95 profiles and 432 pairs. These overlapping samples are subsets of the primary cohort.

The frame follows the record-eligibility filters described in Methods and is not an enumeration of all registered app users. Selection counts were reconstructed offline from the frozen source extract during journal preparation. No inferential results were recomputed. The study size was determined by all qualifying records in the fixed window; no prospective power calculation or target sample size was specified. Eligibility and matching determine the analyzed sample rather than imputation of absent outcomes.

S5. Regular users and minimum-card sensitivity

The 47 regular users are part of the 101-profile cohort. Their profile-weighted mean rose by 8.1 points, and 207 of their 440 sound comparisons increased versus 127 that decreased and 106 unchanged. Because this subgroup was selected on full-window practice history, its larger gain does not show that greater use caused greater improvement.

With a five-card minimum, 88 of 171 comparisons increased (51.5%), 61 decreased (35.7%) and 22 were unchanged (12.9%). The profile-weighted mean change was +6.3 points (95% CI 1.4 to 11.3). The pair-weighted mean was +3.6 points with an interval including zero (-0.1 to 8.4). The direction pattern persists in this smaller sample, while the uncertainty depends on the weighting and selection rule.

Table S1. Primary result, regular-user subgroup and five-card sensitivity. Accuracy means weight profiles equally.

Analysis	Profiles	Pairs	Month 1 → 2 (%)	Change, pp (95% CI)	Higher pairs	Lower pairs	Same pairs
Primary cohort	101	660	69.4 → 74.9	+5.5 (0.4 to 10.4)	278 (42.1%)	177 (26.8%)	205 (31.1%)
Regular-user subset	47	440	69.0 → 77.1	+8.1 (3.5 to 12.9)	207 (47.0%)	127 (28.9%)	106 (24.1%)
Five-card sensitivity	39	171	72.3 → 78.6	+6.3 (1.4 to 11.3)	88 (51.5%)	61 (35.7%)	22 (12.9%)
Five-card regular subset	32	156	70.6 → 77.3	+6.7 (1.5 to 12.2)	78 (50.0%)	57 (36.5%)	21 (13.5%)

S6. First-month frequency comparisons

First-month sessions

Profiles with 1–4 first-month sessions gained 3.3 percentage points on average; those with 5+ gained 7.5 points. The lighter group increased from 70.0% to 73.2% (47 profiles); the more frequent group increased from 68.9% to 76.4% (54 profiles). The difference in gains was +4.2 points (95% CI −5.7 to 14.6), so the larger observed gain does not establish a difference between groups.

At the ten-session cutoff, the 10+ group gained 3.2 points (24 profiles), while the 1–9 group gained 6.2 (77 profiles). The difference was −3.1 points (95% CI −10.7 to 4.6). Higher accuracy in the 10+ group did not mean greater improvement.

Table S2. First-month sessions and paired accuracy change among the same 101 returning profiles. The five- and ten-count partitions overlap.

First-month practice	Matched profiles	Month 1 (%)	Month 2 (%)	Gain, pp (95% CI)
1–4 sessions	47	70.0	73.2	+3.3 (-5.4 to 11.9)
5+ sessions	54	68.9	76.4	+7.5 (2.2 to 13.0)
1–9 sessions	77	67.3	73.5	+6.2 (-0.4 to 12.6)
10+ sessions	24	76.2	79.4	+3.2 (-1.4 to 7.7)

Note. The 10+ session interval is carried over from an earlier shared random-number stream; a fresh-stream recomputation gives −1.3 to 7.6 (Supplementary Methods S2).

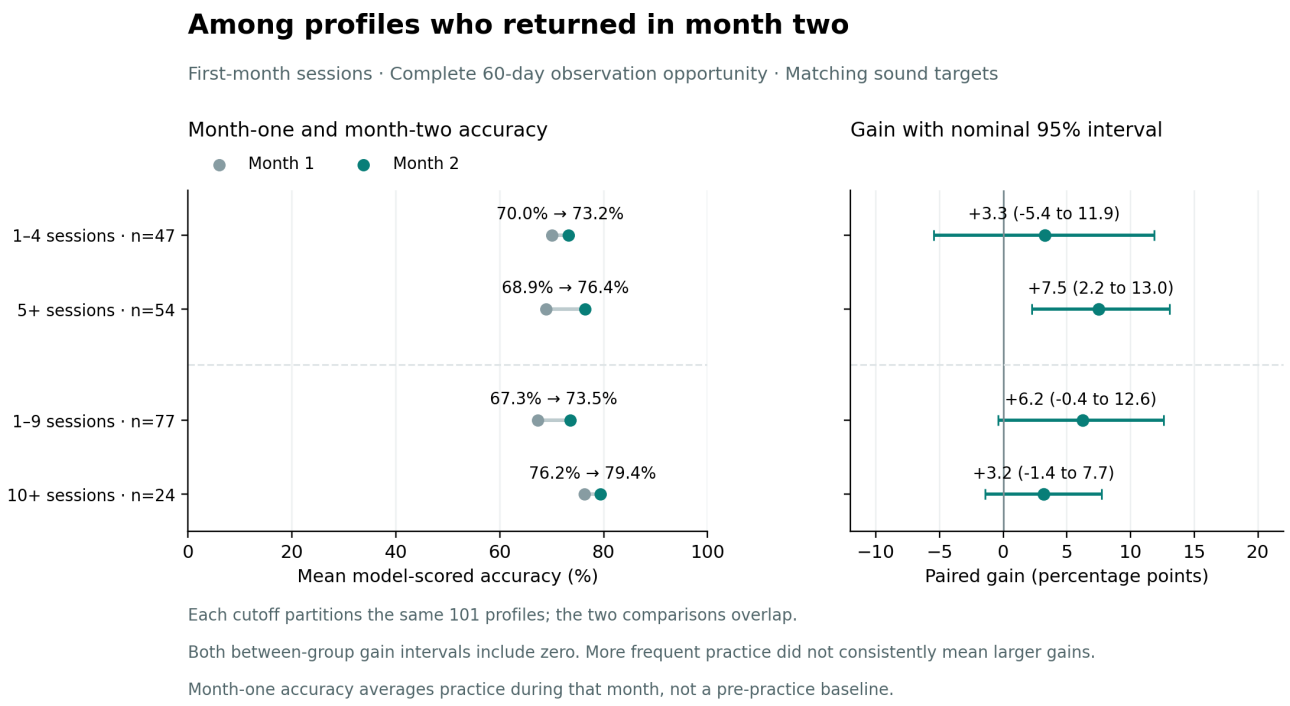


Figure S1. Month 1 and Month 2 means and paired gains by first-month sessions. Partitions overlap; direct gain contrasts appear in Table S4.

First-month practice days

Profiles practicing on 1–4 days increased from 68.9% to 73.4%, a gain of +4.5 points (58 profiles); those practicing on 5+ days increased from 70.1% to 77.0%, +6.9 points (43 profiles). The difference was +2.4 points (95% CI −7.0 to 12.0). At the ten-day cutoff, 10+ days yielded +3.7 points (10 profiles), versus +5.7 for 1–9 days (91 profiles), a difference of −2.0 points (95% CI −11.6 to 7.7). All practice-day comparisons count the first observed practice date.

Table S3. First-month practice days and paired accuracy change among the same 101 returning profiles. The five- and ten-count partitions overlap.

First-month practice	Matched profiles	Month 1 (%)	Month 2 (%)	Gain, pp (95% CI)
1–4 practice days	58	68.9	73.4	+4.5 (-3.3 to 12.0)
5+ practice days	43	70.1	77.0	+6.9 (1.7 to 12.8)
1–9 practice days	91	68.9	74.6	+5.7 (0.1 to 11.2)
10+ practice days	10	74.1	77.8	+3.7 (-4.0 to 11.6)

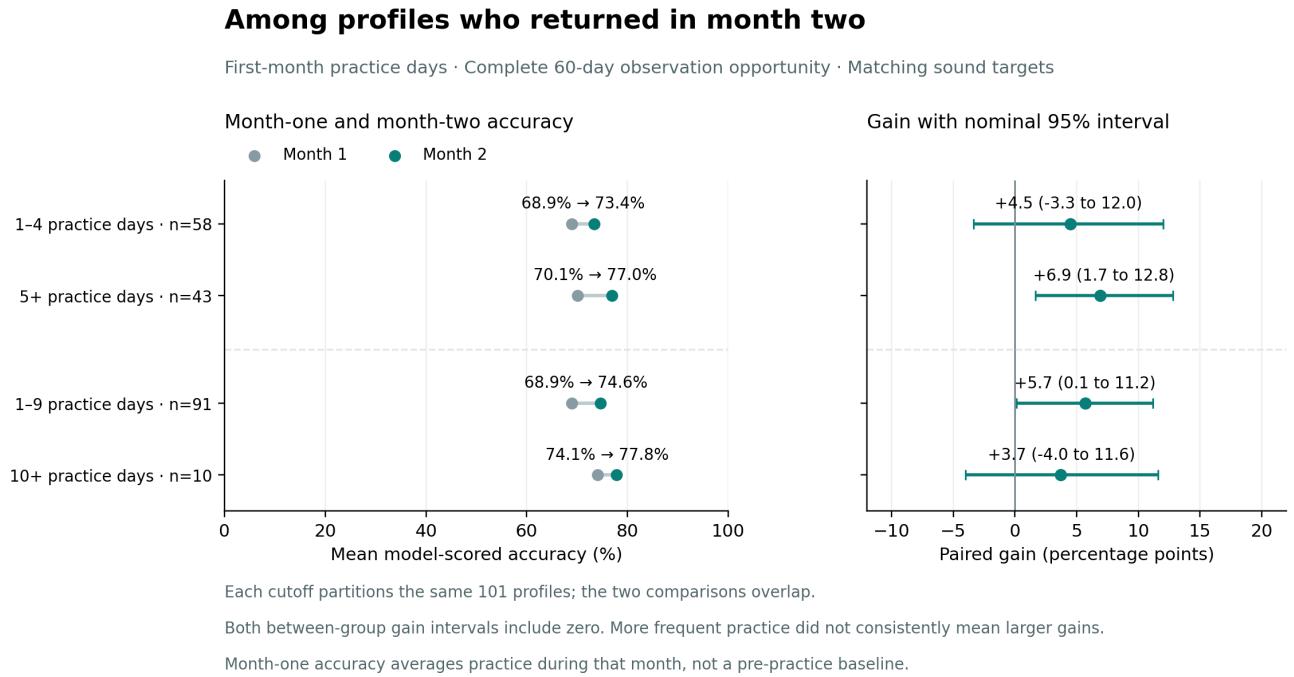


Figure S2. Month 1 and Month 2 means and paired gains by first-month practice days. Partitions overlap; direct gain contrasts appear in Table S4.

Table S4. Direct differences in gains for all four post hoc practice comparisons. Positive values favor the more frequent group; all nominal 95% intervals include zero.

Comparison	Difference in gain, pp (95% CI)
5+ sessions minus 1–4 sessions	+4.2 (-5.7 to 14.6)
10+ sessions minus 1–9 sessions	-3.1 (-10.7 to 4.6)
5+ practice days minus 1–4 practice days	+2.4 (-7.0 to 12.0)
10+ practice days minus 1–9 practice days	-2.0 (-11.6 to 7.7)

Every session-frequency and practice-day group had a higher mean in Month 2. The 5+ groups had larger observed gains than the 1–4 groups, while the 10+ groups had smaller gains than the 1–9 groups, so more practice was not consistently associated with greater improvement. Gains use unrounded means.

S7. Day 1 sensitivity results

In the 95-profile, 432-pair sensitivity sample, mean accepted-card accuracy was 70.1% on Day 1, 71.4% in Month 1 and 76.4% in Month 2. The Day 1-to-Month 2 gain was +6.3 (0.1 to 12.3) percentage points; the Month 1-to-Month 2 gain on the same sound set was +5.0 (-1.0 to 10.7) points. Parentheses show nominal 95% intervals. The baseline definition changes the estimated gain and its uncertainty.

Table S5. Day 1 sensitivity: overall means among the same 95 matched profiles and 432 sound pairs. Both gains use this narrower sample, with nominal 95% intervals.

Profiles	Sound pairs	Day 1 (%)	Month 1 (%)	Month 2 (%)	Day 1 → Month 2 gain, pp (95% CI)	Month 1 → Month 2 gain, pp (95% CI)
95	432	70.1	71.4	76.4	+6.3 (0.1 to 12.3)	+5.0 (-1.0 to 10.7)

The 5+ first-month session group gained +10.1 points from Day 1 to Month 2 (49 profiles), versus +2.3 points in the 1–4 group (46 profiles). The difference in gains was +7.8 points (95% CI −3.8 to 20.2). At the ten-session cutoff, gains were +7.8 for 10+ sessions (23 profiles) and +5.8 for 1–9 (72 profiles), with a difference of +2.0 points (95% CI −8.8 to 13.3). Both between-group intervals include zero.

Day 1-to-Month 2 gains were +5.7 points for 5+ first-month practice days (39 profiles) and +6.7 for 1–4 days (56 profiles), a difference of −1.0 points (95% CI −11.9 to 10.1). At the ten-day cutoff, the 10+ group gained +20.9 points (nine profiles), versus +4.8 for 1–9 days (86 profiles), a difference of +16.2 points (95% CI −0.1 to 32.2). The nine-profile group began at a lower Day 1 mean, 56.3%, and had much more room for its score to rise. Its large observed gain does not establish greater improvement from more practice.

Table S6. Day 1 sensitivity by first-month sessions. Each cutoff partitions the same 95 profiles; partitions overlap. Each row uses the same sound targets in all three periods. The 10+ sessions Day 1 to Month 2 lower bound is +0.04 before rounding.

First-month practice	Profiles	Sound pairs	Day 1 (%)	Month 1 (%)	Month 2 (%)	Day 1 → Month 2 gain, pp (95% CI)	Month 1 → Month 2 gain, pp (95% CI)
1–4 sessions	46	147	72.0	71.1	74.2	+2.3 (-8.0 to 11.8)	+3.1 (-6.8 to 12.4)
5+ sessions	49	285	68.4	71.7	78.4	+10.1 (3.3 to 17.2)	+6.7 (0.6 to 13.7)
1–9 sessions	72	288	70.1	69.6	75.9	+5.8 (-1.9 to 13.3)	+6.3 (-1.3 to 13.7)
10+ sessions	23	144	70.1	77.2	77.8	+7.8 (0.0 to 16.2)	+0.6 (-5.4 to 6.3)

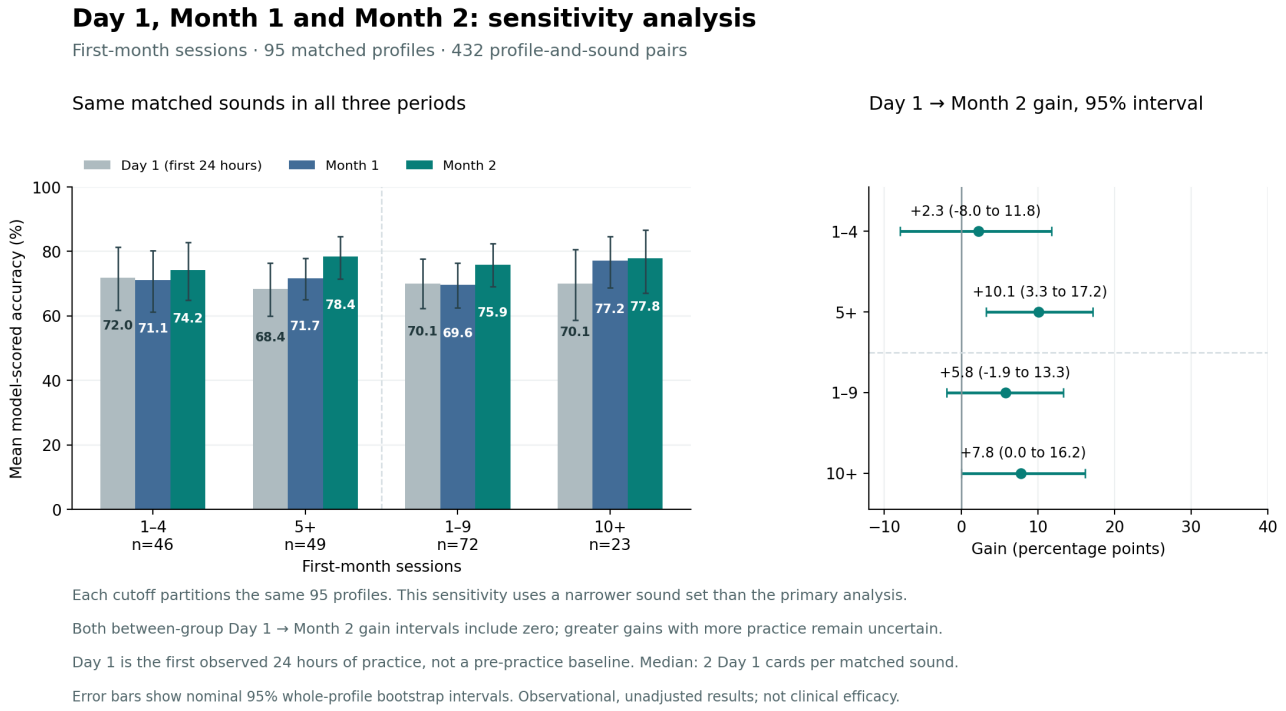


Figure S3. Day 1, Month 1 and Month 2 means and Day 1-to-Month 2 gains by first-month sessions. Direct gain contrasts appear in Table S8.

Table S7. Day 1 sensitivity by first-month practice days. Each cutoff partitions the same 95 profiles; partitions overlap. Each row uses the same sound targets in all three periods.

First-month practice	Profiles	Sound pairs	Day 1 (%)	Month 1 (%)	Month 2 (%)	Day 1 → Month 2 gain, pp (95% CI)	Month 1 → Month 2 gain, pp (95% CI)
1–4 practice days	56	204	68.4	68.8	75.1	+6.7 (-2.7 to 15.9)	+6.3 (-2.7 to 15.0)
5+ practice days	39	228	72.5	75.2	78.2	+5.7 (-0.2 to 12.1)	+3.0 (-3.1 to 9.1)
1–9 practice days	86	361	71.6	71.4	76.3	+4.8 (-1.8 to 11.0)	+4.9 (-1.5 to 11.1)
10+ practice days	9	71	56.3	72.1	77.2	+20.9 (5.6 to 35.8)	+5.1 (-4.0 to 14.0)

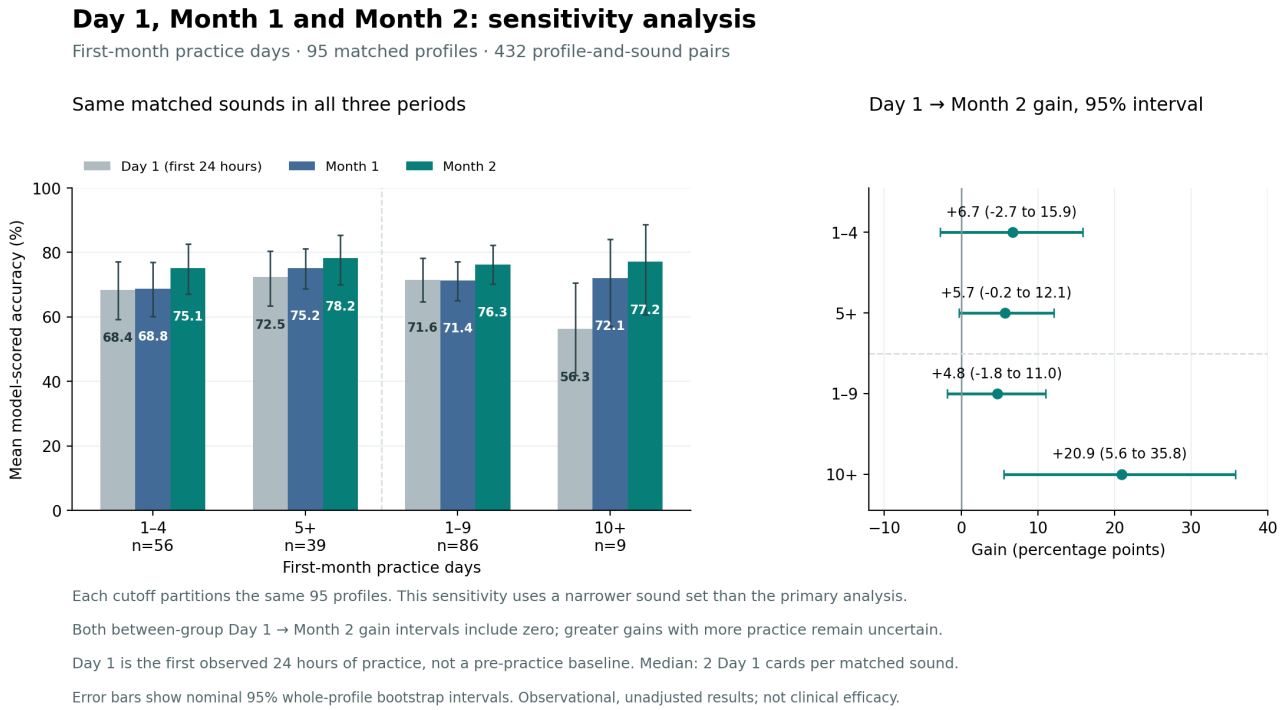


Figure S4. Day 1, Month 1 and Month 2 means and Day 1-to-Month 2 gains by first-month practice days. Direct gain contrasts appear in Table S8.

Table S8. Direct differences in Day 1-to-Month 2 gains for all four post hoc frequency comparisons. All nominal 95% intervals include zero.

First-month comparison	Difference in Day 1 → Month 2 gain, pp (95% CI)
5+ minus 1–4 sessions	+7.8 (-3.8 to 20.2)
10+ minus 1–9 sessions	+2.0 (-8.8 to 13.3)
5+ minus 1–4 practice days	-1.0 (-11.9 to 10.1)
10+ minus 1–9 practice days	+16.2 (-0.1 to 32.2)

S8. Table and figure crosswalk for aggregate exports

Exports retain the original Table 1–15 identifiers. Main Tables 1–7 keep their numbers; original Tables 8–15 map to Supplementary Tables S1–S8. Original Figures 1–4 become main Figures 2–5; original Figures 5–8 become Supplementary Figures S1–S4. Main Figure 1 is new. Plot assets and their embedded labels are unchanged.