Hippocampal Acetylcholine Signalling and Running Speed

Six-Week Laidlaw Scholarship Placement Report, Barry Lab, UCL

Kai Gohil, MSci Neuroscience, UCL. Laidlaw Scholar.
Supervisor: Professor Caswell Barry, UCL Cell & Developmental Biology

September 2026

Abstract

Hippocampal acetylcholine (ACh) has been proposed to encode a state-transition prediction error (STPE), gating plasticity in the way dopamine does for reward prediction error (de Cothi, Shipley & Barry, 2026). The theory predicts a positive ACh–speed relationship, because unresolved transition error should accumulate faster the faster an animal moves. Using GRAB-ACh fibre photometry from a single mouse running a VR track across four sessions, I compared five candidate functions via AIC/BIC, diagnosed and split a data-quality artefact affecting two sessions, tested acceleration and speed×acceleration interactions, and, following a supervisor suggestion, cross-correlated ACh against speed for a temporal offset. The clearest result is that ACh changes roughly one second before the corresponding change in speed. The lag ranges from −1.0 to −1.2s and holds across all six data segments tested. Correcting for it improves the best model's fit in four of six segments when the sample filter is held constant (e.g. Day 1: R^2=0.506→0.595); the two noisiest pre-shift segments show no reliable change. A fixed sensor or processing-pipeline delay is still on the table as an alternative explanation. Acceleration is a weaker predictor and shows no comparable stable temporal relationship.

1. Introduction

This was a six-week Laidlaw Scholarship placement in Professor Caswell Barry's lab at UCL, looking at the relationship between hippocampal acetylcholine signalling and running speed in mice.

The starting point is a theoretical proposal by de Cothi, Shipley and Barry: hippocampal ACh encodes a state-transition prediction error (STPE), the gap between an animal's predicted next state and what it actually observes, computed via the hippocampal successor representation and gating plasticity through theta-timed spike-timing-dependent plasticity. It is the proposed structural-learning counterpart to dopamine's role in reward prediction error. Because transition error should accumulate faster the faster an animal moves, the theory predicts a positive ACh–speed relationship: something to be tested directly rather than regressed away as a confound.

The brief from Professor Barry was to find a simple, accurate function relating running speed to the hippocampal ACh signal, justified through proper model comparison (AIC/BIC) rather than a single R^2 value, and to extend that analysis as time allowed.

2. Methods

Data were recorded via GRAB-ACh fibre photometry (470 nm signal, 405 nm isosbestic control) from a single head-fixed mouse running a virtual-reality track across four recording sessions. Position was sampled via the VR system. $\Delta F/F$ was computed by regressing the signal channel against a slow-timescale fit of the isosbestic control channel and z-scoring the result (Figure 1); isosbestic regression R^2 served as a session-level data-quality indicator throughout. Running speed came from numerical differentiation of VR position, with teleport/lap-reset artefacts and extreme outliers removed. Acceleration was a second differentiation of speed.

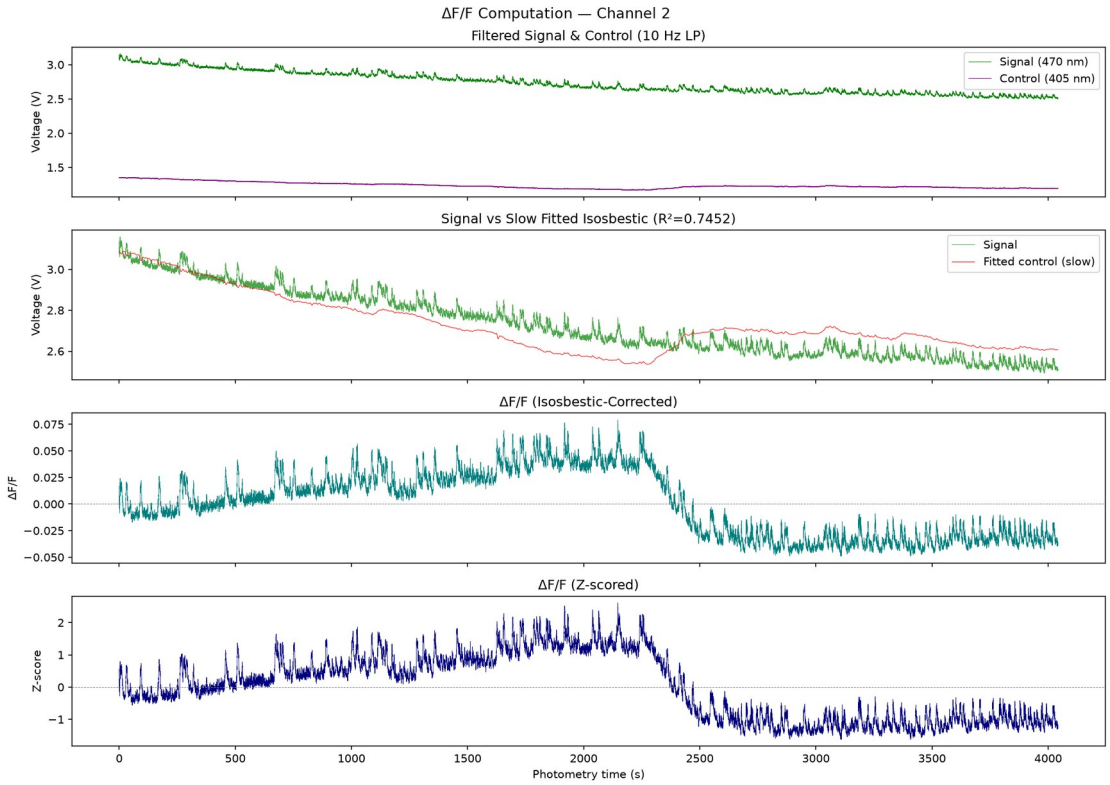

Figure 1. ΔF/F computation pipeline shown on the Day 2 pre-shift session (the artefact case discussed in §3.2). The bimodal step visible in the third panel and the poor isosbestic fit (R^2=0.7452) are what prompted the session-splitting fix. A clean session's diagnostic would show a higher isosbestic R^2 and no mid-session step.

Five candidate functions (linear, logarithmic, log-quadratic, power-law, saturating exponential) were fit to raw, unbinned, unsmoothed ΔF/F and speed via least-squares regression, and compared using AIC and BIC computed directly from residual sum of squares. I did this from first principles rather than using a packaged statistics function, deliberately. Data stayed unbinned throughout. Binning or smoothing would average away the noise AIC/BIC needs to distinguish real improvement from overfitting, including when acceleration and speed×acceleration terms were added, and when speed and ΔF/F were cross-correlated (Section 3.4).

Analysis code extends the lab's existing pipeline (basicPlotting.py, and fp_utils.py by Dr Sarah Shipley for TDMS parsing and sync-pulse alignment between the photometry and VR clocks). The new model-comparison, session-splitting, acceleration and cross-correlation scripts were written by me on top of that base.

3. Results

3.1 The baseline speed–ACh relationship

On Day 1, the first and most extensively analysed session (isosbestic R^2=0.978), all five functions rank consistently by AIC and BIC (Table 2). Day 4's isosbestic fit is marginally higher at 0.985, so "cleanest" by that specific metric belongs to Day 4; Day 1 got the deepest methodological development. The sample size for the fitting stage is n=12,154, which can be verified from Table 2's own values: the linear-model BIC−AIC gap of 14.81 equals k(ln n − 2) with k=2, giving ln n = 9.406 and n = 12,154 exactly. An earlier-stage n=17,060 in the working log reflects a pre-filtering intermediate step that does not survive to the model-fitting sample. Log-quadratic is best-supported (R^2=0.453), followed closely by saturating exponential (R^2=0.450). Log and unconstrained power-law are essentially indistinguishable by R^2 (≈0.44 each). Power-law's unconstrained fit converges to parameters (a≈11,469, b≈4.24×10^{-5}, c≈−11,470) that mathematically collapse to a disguised

log model (AIC ≈ −5294.87, tied with log's −5296.93); it offers no distinct shape. The constrained fit shown in Table 2 pins to its bounds and ranks worse. Linear is clearly worst (R^2=0.302).

Model	Form	AIC	BIC	R^2
Log-quadratic	a·ln(x)²+b·ln(x)+c	−5555.31	−5533.10	0.453
Saturating exp.	a(1−e^(−bx))+c	−5491.50	−5469.28	0.450
Log	a·ln(x)+b	−5296.93	−5282.12	0.441
Power-law (constr., b≥0.05)	a·x^b+c	−5215.64	−5193.42	0.438
Linear	a·x+b	−2590.01	−2575.20	0.302

Table 2. Five candidate functions fit to raw Day 1 data (n=12,154), ranked by AIC/BIC.

Log-quadratic's fitted form is $a \cdot \ln(x)^2 + b \cdot \ln(x) + c$ with $a < 0$ (concave-down). The turning point sits at $\ln(x_{turn}) = -b/(2a) \approx 6.23$, i.e. $x_{turn} \approx 508$ pos-units/s, roughly double Day 1's maximum recorded speed (~250, the 99.5th-percentile clip). Within the observed range the fitted curve is still slowly rising rather than turning over. Day 4's fit gives coefficients a≈−0.258, b≈2.125, c≈−2.688, and a turning point at ≈61 pos-units/s, which may fall within Day 4's observed speed range. Whether Day 4 shows a real within-range downturn is a specific, checkable open question rather than something I have assumed here (Section 5). The specific fitted (a, b, c) triple for Day 1's log-quadratic was not preserved in the working log to the precision needed to quote here, and is added to the followup list.

Even the best-supported model explains under half the variance in raw ΔF/F. This is a real limit of speed as a predictor: at any given speed, ΔF/F samples vary several times more than the range spanned by any fitted curve.

3.2 Data-quality checks: a real artefact and a false alarm

Two issues in the initial four-session comparison needed resolving before I could trust cross-day conclusions. Both are reported here because resolving them is as methodologically relevant as the results themselves.

Days 2 and 3 showed markedly weaker isosbestic regression fits (R^2≈0.72–0.75, versus ≈0.98 on Days 1 and 4) and a bimodal ΔF/F distribution. I traced this to a discrete mid-session step consistent with a fibre/probe-coupling shift (Figure 1 above shows this artefact on Day 2). Each session was split into "before" and "after" segments at its step, excluding the transition window, and refit independently. The fix recovered substantial signal: Day 2's post-shift half improved from R^2=0.017 (unsplit) to R^2=0.330; Day 3's from R^2=0.024 to R^2=0.411. The pre-shift halves improved by a smaller, inconsistent amount, an open question noted in Section 5.

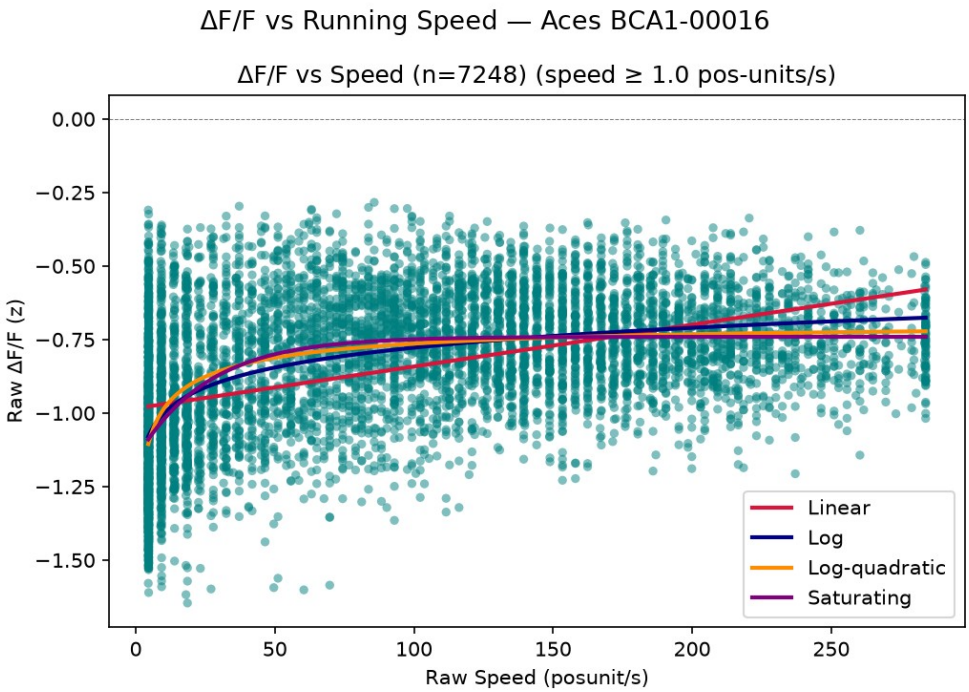

Figure 2. Four of the five candidate functions (power-law omitted for legibility; see Table 2 for its comparison values) fit to Day 2 post-shift raw data (n=7,248).

Separately, Day 4's log-quadratic fit briefly appeared unstable: one run showed a positive curvature coefficient, opposite in sign to every other session, with implausible predicted values. I initially concluded this was numerical ill-conditioning specific to Day 4. Tracing it further, the actual cause was an unlabelled console print. The array inspected under "log-quadratic" was power-law's output from the next block, not logquad's. Log-quadratic's true Day 4 fit had been correct from the first run. I record this retraction because catching my own mistaken diagnosis, and the lesson it left (trace a suspicious number back to the exact variable that produced it, not the nearest printed line), is as much a real part of this project as the results it sits alongside.

3.3 Acceleration as a candidate predictor

Following Professor Barry's instruction to extend the analysis beyond speed alone, I tested acceleration, a second numerical derivative of position, both as an independent predictor and in combination with speed. Acceleration alone showed a real, symmetric, quadratic relationship with ΔF/F, consistent with ACh scaling with the magnitude of speed change rather than its direction. A cubic term added no explanatory power, confirming the predicted symmetry. But acceleration explained substantially less variance than speed alone ($R^2 \approx 0.13$ versus ≈0.45–0.51).

Combining speed and acceleration additively, using each predictor's own best-established shape (log-quadratic for speed, quadratic for acceleration), improved on speed alone by a negligible amount ($R^2 = 0.5064 - 0.5068$ in the strongest test). I also tested two speed×acceleration interaction terms, prompted by a mid-placement AI-assisted (LLM) review of the working log that correctly flagged the additive-only models as unable to rule out interaction. Neither improved the fit. The full picture replicated across all six data segments (the two full sessions plus the four split halves). Speed is the dominant predictor of ACh in this dataset. Acceleration adds essentially nothing on top.

3.4 Cross-correlation: ACh changes before running speed

Following a supervision meeting, Professor Barry suggested computing a cross-correlogram between speed and ΔF/F directly, to test for a temporal offset rather than assuming the two are best compared at the same moment. I tested this for both speed and acceleration.

For speed, this is the clearest and most consistent finding of the placement. Cross-correlating raw, unsmoothed speed against raw ΔF/F recovers a negative peak lag (ACh changes before the corresponding change in speed) of −1.0 to −1.2 seconds, in the same direction and similar magnitude across all six data segments (the two full sessions, plus Days 2 and 3 each split into before/after halves), with no exceptions. The consistency is itself evidence the lag reflects something real: an artefactual estimate would be expected to vary more between clean and noisy segments than this does. A fixed sensor or processing-pipeline delay (sensor kinetics, isosbestic correction) is also on the table as an explanation for an apparent lead. It has been raised with the lab and is not yet ruled out. A synthetic sign-flip test and a sync-alignment check against the rig's known timing are the two concrete next steps for distinguishing a real anticipatory signal from a fixed processing delay (Section 5).

Table 3 uses the same speed>0 sample filter throughout for both the unshifted and lag-corrected fits. This avoids a filter mismatch present in some earlier working comparisons.

Segment	Peak lag	Logquad R^2 (unshifted, same filter)	Logquad R^2 (lag-corrected)	Gain
Day 1 (n=15,234)	−1.10s	0.506	0.595	+0.089
Day 2, before (n=10,613)	−1.00s	0.151	0.147	−0.004
Day 2, after (n=8,758)	−1.00s	0.386	0.503	+0.117
Day 3, before (n=8,033)	−1.00s	0.052	0.050	−0.002
Day 3, after (n=7,914)	−1.00s	0.484	0.525	+0.041
Day 4 (n=17,708)	−1.20s	0.379	0.467	+0.088

Table 3. Peak lag and log-quadratic R^2, before and after lag correction, same sample filter throughout.

Under this matched comparison, four of six segments show a clear improvement (+0.04 to +0.12). The two noisiest, pre-shift segments (Day 2 and Day 3 "before") show a small negative change (−0.002 to −0.004), i.e. no reliable improvement. This is a materially more honest picture than "every segment improves": lag correction helps where the underlying signal is already reasonably strong, and its effect on the two weakest segments cannot be distinguished from noise. Log-quadratic remains the best-supported model shape after correction in every segment, now explaining 47–60% of variance in the four higher-quality segments, up from 38–51% before correction (same filter, same four segments).

For acceleration, no comparable relationship emerged. Peak lags across the same six segments ranged from +1.5s to −0.1s with no consistent direction. Lag-correcting acceleration's own fit did not improve, and generally slightly worsened, its already weak R^2. The most defensible interpretation is that acceleration's relationship with ΔF/F is too weak for a reliable lag to be recoverable by this method, consistent with Section 3.3's finding that acceleration is the weaker predictor.

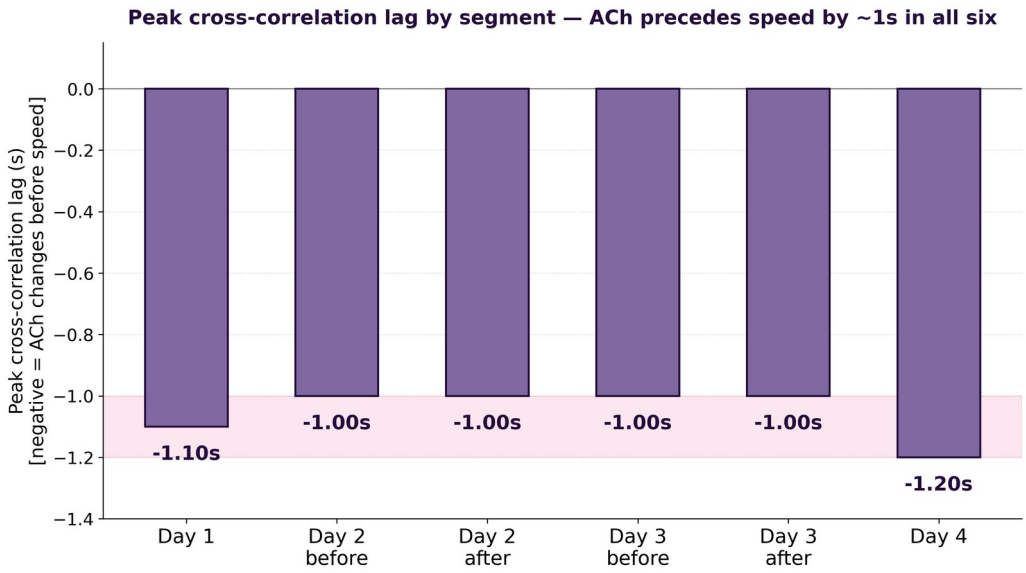

Figure 3. Bar chart summarising the peak cross-correlation lag (negative = ACh precedes speed) for each of the six data segments; not a raw correlogram trace.

4. Discussion

The lag finding is the most scientifically interesting result of the placement, for two reasons beyond its statistical strength. First, it was a direct answer to a hypothesis Professor Barry proposed in a supervision meeting, and it held up under a matched-filter re-analysis (Section 3.4), not just under the earlier unmatched comparison. Second, it sits in tension with a naive reading of the STPE framework: if ACh reports a prediction error computed from an already-experienced transition, a same-time or lagging relationship with speed might seem more natural than an anticipatory one. A one-second lead instead raises the possibility that the signal reflects something closer to an upcoming motor or behavioural state, rather than, or in addition to, a purely retrospective error. This has to be held alongside the unresolved sensor-delay explanation in Section 3.4. Until a sync check rules that out, the lead is consistent with either interpretation, and I do not think the data can currently distinguish them.

A smaller observation: before lag-correction, the saturating exponential model was a close second to log-quadratic in every session; after correction, its ranking drops in most sessions. I do not have a settled explanation, and flag it as an open question. It may bear on whether ACh's relationship with speed has a real ceiling (Section 3.1) or keeps rising slowly outside the observed range.

5. Limitations and Future Work

In rough priority order: (1) rerun all six segments through one identical pipeline, printing n, R^2, AIC and BIC for both the unshifted and lag-shifted fit of every model under one consistent sample filter, so every number in this report and the accompanying poster traces to a single, directly comparable table rather than being assembled from separately-run analyses as it is now; (2) run a synthetic sign-flip test and check the photometry-to-VR sync alignment against the rig's documented timing, to rule out a fixed processing delay as the source of the lag; (3) check Day 4's log-quadratic turning point (≈61 pos-units/s) against its actual observed speed range, to test whether Day 4 shows a genuine within-range downturn unlike Day 1; (4) extend lag-correction to the combined speed-and-acceleration models rather than treating them separately; (5) preserve Day 1's fitted (a, b, c) triple for log-quadratic explicitly, alongside the summary statistics currently recorded. This project used data from a single mouse across four sessions. The consistency of the lag finding across sessions would not by itself establish that it generalises across animals; confirming that would need multi-animal data.

6. Conclusion

This placement produced a defensible, AIC/BIC-justified account of how hippocampal ACh relates to running speed, ruled out acceleration as a stronger predictor, and, prompted by supervisor feedback, found a consistent one-second lag between ACh and running speed that improves the fitted models in four of six segments once sample filtering is held constant. Two methodological lessons (Section 3.2) and one methodological correction to the lag comparison itself (Section 3.4) were caught during the write-up and are reported here alongside the results they informed, rather than smoothed over.

7. Reflection on Methods Lessons

7.1 The lesson I did not expect to learn

Coming into this placement I thought the hard bit would be understanding the theory. It was not. The theory (that hippocampal acetylcholine encodes a prediction error about state transitions, and should therefore track how fast an animal is moving) took an afternoon to read and a supervision meeting to argue about. What actually took the six weeks was learning that a number appearing on a screen is not the same thing as knowing what it means, and that most of what a scientist does is close the gap between those two things.

7.2 Trust the array, not the print statement

The single lesson I keep coming back to happened halfway through the placement. Day 4's log-quadratic model kept returning a positive curvature coefficient: the wrong sign, giving physically impossible predicted values. I spent the better part of a session convinced this was numerical ill-conditioning specific to Day 4, and wrote up a whole retraction of the day's results on that basis. It wasn't. The array being printed on my screen under the label "log-quadratic" was actually the next model's output. A missing print label, three lines up. The fit itself had been correct from the first run. I learned, in a way I don't think I could have learned from a textbook, that a results table computed directly from a function's return value can be right at the same time as the debugging trail explaining it is wrong, and that the discipline is to trace every suspicious number back to the exact variable that produced it, not to the nearest printed line.

7.3 Statistics is a way of asking questions honestly

The brief for the placement asked for AIC/BIC-based model comparison rather than R^2 alone, and I didn't fully understand why until I started using it. R^2 will always improve when you add parameters. AIC and BIC penalise complexity, so a more flexible model has to actually explain more variance to win. Computing them from first principles rather than a stats library forced me to understand what each term meant, and to catch my own mistake later when I compared unshifted and lag-corrected fits under two different sample filters, a mistake the R^2-only version would have hidden. The corrected comparison told a materially more honest story, and it was better for the write-up, not worse, that I had to explain the correction rather than paper over it.

7.4 How much scientists actually depend on each other

The data I worked with had been collected by another scholar; the analysis pipeline I built on top of was written by Dr Sarah Shipley; the questions that shaped my six weeks came from a supervision meeting with Professor Barry where he suggested cross-correlating for a temporal offset rather than assuming same-time comparison. The cross-correlation ended up being the most interesting finding of the placement. I don't think I would have got there on my own inside a six-week window. Research in a working lab is a chain of people handing on tools, data, and questions, and the discipline is knowing which link you are and treating the others with the seriousness they deserve.

7.5 What I would carry into the next project

Three specific habits. First, label everything you print. Second, when you catch yourself with a clean-looking result, ask what filter or sample you used, then verify it matches the comparison you are drawing. Third, when a supervisor makes a suggestion in a meeting you weren't expecting, take it seriously enough to test properly. The cross-correlation suggestion took two hours of extra work and turned out to be the placement's headline finding. I don't think any of these are specifically neuroscience lessons; they are lessons about doing careful work with quantitative data, and I expect to keep using them.

References

de Cothi, W., Shipley, S., & Barry, C. (2026). Acetylcholine: a candidate substrate for hippocampal predictive learning? Nature Reviews Neuroscience. https://doi.org/10.1038/s41583-026-01058-w

Acknowledgements

With thanks to Professor Caswell Barry and Dr Sarah Shipley for supervision and guidance throughout this placement, and to the Laidlaw Scholarship programme for funding it.