Table 1: Table 1. Statistical Analysis Framework

Complete analytical approach from laboratory development to field validation

| Phase | Step | Research Question | Model | Method | Result | n | Performance*1* |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Laboratory Development | Discovery | Can immune genes predict infection costs? | DISC-1 | Linear regression (PC1, PC2 → weight loss) | Significant but modest prediction | n = 136 | R² = 0.106\*\*\* |
| Laboratory Development | Optimization | Can machine learning improve prediction? | DISC-2 | Random forest (19 genes → weight loss) | Substantial improvement achieved | n = 136 | R² = 0.476\*\*\* |
| Laboratory Development | Validation | Is the model reliable? | DISC-3 | Train-test cross-validation | Strong predictive accuracy | n = 95→41 | r = 0.79\*\*\* |
| Field Translation | Detection | Does the model work in wild populations? | FIELD-1 | Predicted vs. observed infection status | Successfully detects infection | n = 305 | +1.15%\*\*\* |
| Field Translation | Discrimination | Can it distinguish parasite species? | FIELD-2 | Predicted loss by species identity | Species-specific responses | n = 169 | E.f: +2.06%\*\*, E.r: +1.25%\*\* |
| Field Translation | Scaling | Does it correlate with infection severity? | FIELD-3 | Predicted loss vs. parasite load | Scales with infection intensity | n = 185 | r = 0.233\* |
| Biological Validation | Physiological relevance | Does it capture real health impacts? | PROOF-1 | Predicted loss vs. body condition | Correlates with actual body weight | n = 336 | ρ = -0.115\* |
| Biological Validation | Specificity | Is the response Eimeria-specific? | PROOF-2 | Predicted loss vs. parasite community | Specific to Eimeria infections only | n = 305 | p < 0.001\*\*\* |
| *1*\*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001 | | | | | | | |