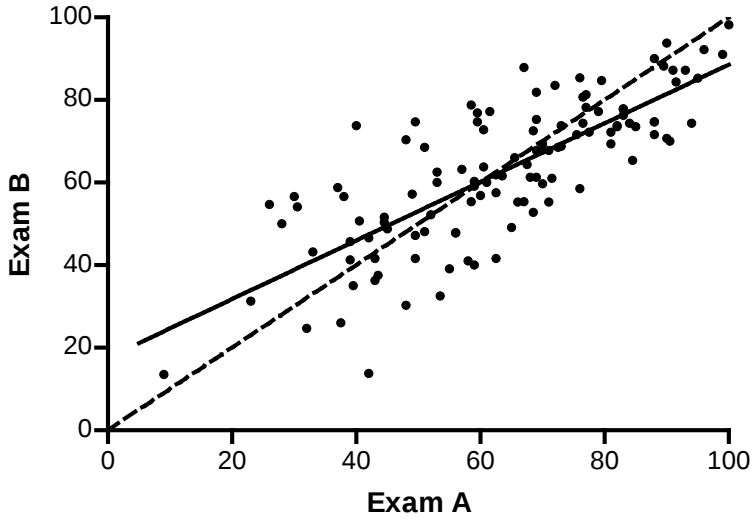
Assessment - What, why and how?

D.A. Saint, D. Horton and A. Elliott, School of Medical Sciences, University of Adelaide, Adelaide, SA 5005, Australia.

Introduction. Assessment is often used to "score" or "rank" students (purely summative)- here we examined how effective and efficient our own assessment methods were in achieving this. Summative assessment is often associated with learning that has short lasting outcomes. In response, teachers incorporate varying amounts of formative assessment into the curriculum, to better drive learning. Here, we examined one implementation of formative assessment and whether it did indeed improve student learning outcomes.

Methods. We examined the grade history of 130 students in 2nd year physiology including nine summative assessments, plus an exam component consisting of 5 sections. We examined the performance of students in each summative assessment and measured the degree of correlation between grade scores in each of these different assessment tasks. We also compared final course scores with those derived from a hypothetical scenario in which scores were derived from the final report (50%) and one exam section (50%) only. In a separate cohort of students (n=106), the assessment included a short in-class exam (paper A) early in the semester and which counted for 12.5% of the overall mark. The intention was to provide the students with realistic feedback on their own learning progress, hopefully to improve their final exam scores.

Results. Moderate correlations (r=0.4-0.6) were observed between final report and tutorials, exam and final report, and the exam and tutorials, respectively. The score from the hypothetical scenario showed a very strong correlation with final course score (r=0.98). Bland-Altman analysis revealed a mean bias (final score – scenario A) of 4.3 marks with a trend towards decreasing bias in higher achieving students. For the formative assessment cohort, correlation analysis showed that students who performed poorly in the formative assessment task (Exam A) tended to improve their performance in the final exam (Exam B). In contrast, higher performing students in Exam A showed a slight reduction in performance during the final exam. (Figure 1; dashed line is slope of 1, solid line is regression of the data, r=0.78; $P < 0.001$).



Discussion. Our data suggest that our current system of multiple assessments yields little additional benefit in ranking students compared to a scenario in which two smaller assessments is applied. Inclusion of an early formative type of assessment appeared to marginally improve student learning in lower performing students.