To whom this concerns:

I have learned that the Geisel School of Medicine Ethics Committee has opened numerous student integrity investigations on the basis of semi-random log data. I write because I fear that the Committee may have fallen victim to the statistical fallacy known as the Birthday Paradox. The Birthday Paradox is well described in a 2012 Scientific American article that may be accessed at: https://www.scientificamerican.com/article/bring-science-home-probability-birthday-paradox/. Suffice it to say that if you host a party of 23 random people and wonder about the probability that at least two of them will have the same birthday, the probability is NOT 23/365 (6.3%) as many people assume but in fact is over 50%. Using analogous reasoning, let us imagine that you have students take a test with 100 questions and that you have a rogue ipad, computer, or other device automatically conducting 100 downloads of information pertinent to the test during the test-taking period. The probability of the log data showing a synchronous answering of a question concurrent with exactly the relevant information being automatically downloaded from the school’s e-curriculum platform is 1 – (.99^4950) = (1 - 0.0000000000000000002478638881) = .999999999999999998632472229 or, if you do a bit of rounding = 1 (i.e., certainty). I recommend that before the Committee invests more confidence than warranted in the use of their integrity-testing software, that they recruit a statistician to check their statistical assumptions about the rarity of synchrony between the student's answer and the content of an automated upload.

Because each student used their respective devices differently the night before the exam, some students had repeated automatic downloads that had nothing to do with the test while other equally-innocent students' rogue devices happened to engage in spontaneous downloading of information that coincided with information pertinent to one or two questions on the test. I therefore recommend that the Ethics Committee discount the likelihood of a student cheating if only one or two questions on any single exam were answered synchronously with the relevant information being downloaded spontaneously by a secondary device. If five or more questions on the test were answered synchronously with the information being downloaded on the Ipad, that would be more likely to be the result of a student deliberately downloading relevant information from Canvas, but still not deterministic.

Accusing a student of cheating on an exam is such a serious charge that the Committee should seek corroboration by different methods before reaching a judgment. Other corroborative evidence might consist of the student being chronically poorly prepared to answer questions in class and the student's past exam performance having a Jekyll & Hyde quality of being embarrassingly poor some times and amazingly superior other times. It is also incumbent on the School to warn students that other devices that they have used to download information relevant to exams should be disconnected from power sources, including batteries, and not just left in sleep mode during the exam. In the absence of ancillary corroborative evidence of cheating, the current accusations based solely on the currently available log data strike me as being irresponsible and likely to lead to grave injustice being done.

To tell innocent students that digital forensic evidence "proves" that they cheated and that they can either "take responsibility" and throw themselves on the mercy of the administration or risk getting expelled reminds me of the police use of polygraphs to extort false confessions from suspects who were later demonstrated to be completely innocent. The boilerplate apologies that many students expressed
before the Committee looked more like the product of coercive administration coaching than the authentic recriminations of miscreants caught in the act of cheating.

Sincerely yours,

William J. McCarthy, Ph.D.
Adjunct Professor, Health Policy & Management
UCLA Fielding School of Public Health
wmccarth@ucla.edu; 310-456-3829