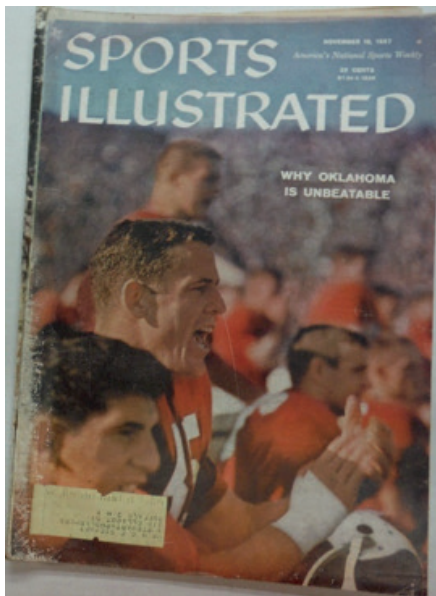


# The curse of Sports Illustrated

BY CHRIS POTTER

“Not a supernatural curse, but a basic statistical concept of blinding simplicity.”  
What is ‘regression to the mean’?



Oklahoma won a record 47 consecutive games to make the cover. That week they lost to Notre Dame.

**“If the editorial convocation here at ENT Towers were to call upon the staff to disrobe in the misguided hope of a boost in circulation, you can be reassured they would get pretty short shrift.”**

I am reliably informed that our former North American colonies publish a periodical known as *Sports Illustrated* (note, incidentally, the characteristically incorrect use of the plural noun - ‘sports’ is of course an adjective, as in ‘sports day’). My own personal awareness of this publication comes solely from annual tabloid coverage of the rather anachronous swimwear issue, where the finest female athletes of their generation are invited to disport themselves in various states of undress as though emancipation, suffragism and the Women’s Movement had never happened. Now don’t get me wrong, I appreciate the female form in all its glory as much as your average debauched roué, but if the editorial convocation here at ENT Towers were to call upon the staff to disrobe in the misguided hope of a boost in circulation, you can be reassured they would get pretty short shrift. One has to draw the line somewhere, you know.

Anyway, appearance on the cover of the magazine has long been seen as a poisoned chalice, the victim subsequently falling foul of some sort of jinx and experiencing a catastrophic slump of form. Indeed, a recent study claimed no fewer than 913 (37.2%) of those who have featured on the cover, out of the 2,456 covers to date, had experienced “significant misfortune” after publication [1]. Famous victims include Pete Rose, whose 44-game hitting streak ended, and Jon Peters, a high school pitcher with a 51-0 record, losing the only game of his school career the week after appearing.

Of course we are dealing here not with some supernatural curse, but a basic statistical concept of such blinding simplicity that it can appear obvious even to our densest core surgical trainee and his overworked synapse, but has consequences so profound it has fooled the finest minds of the ivory towers. For instance, Prof Horace Secrist wrote *The Triumph of Mediocrity in Business* in 1933, demonstrating beyond any reasonable doubt that the most successful US companies tended to become marginally less successful over a 10-year period,

whereas the least successful appeared to improve their performance. He claimed to have stumbled upon a profound universal economic truth to loud approval from his colleagues of ‘the dismal science’. Then a humble statistician pointed out he had spent a decade and 468 pages, 140 tables and 103 charts demonstrating the facile concept of ‘regression to the mean’.

In simple terms, any single test is an imperfect measure of a variable. Thus if we select a population on the basis of an extreme value of a variable, they are individually likely to be extreme in relation to their own personal mean. If we then retest that selected population, their values will tend to be less extreme. For example, imagine I have assembled all the trainees in our region in their underwear and made them perform standing jumps (the programme director has warned me about this once already). If I then select only those who jump two standard deviations beyond the mean, I am far more likely to select mediocre jumpers (who by definition are common) who happen to have fluked a big one, rather than freaks who average three deviations above the mean (and are by definition vanishingly rare) who have underperformed. Hence if I get that population to jump one more time (before the postgraduate dean arrives with the authorities) they are far more likely to underperform compared to their previous jump. They have thus regressed to the mean.

This has massive implications for medical research, which mainly seems to consist of using imperfect tests to select populations on the basis of extremely abnormal results. Now if we know the test: test reliability coefficient for a biochemical variable and set a threshold selection criterion of 3SDs from the mean, we can calculate how much regression to expect on retest for a variety of indices. For a highly reliable test such as serum sodium concentration, we might only expect a mean drop of 2.5% on retesting, but for lactate dehydrogenase we find a staggering 26% drop purely from regression to the mean. This agrees remarkably well with quantitative



Francis Galton, statistician extraordinaire. His autobiography includes a recipe for the perfect cheese sandwich.

estimates of the placebo effect in early studies, and may therefore go a long way to explain this mysterious phenomenon [2], and emphasise the importance of an adequate control group.

We owe the earliest description of this concept to that egg-headed cousin of Darwin, Francis Galton. An obsessive bean-counter, he charted the heights of parents and their offspring and found that the children of extremely tall or short parents tended to be less extreme in their heights. Importantly, he also found the converse was true, and thus regression also works backwards in time. He somehow found time to invent a pioneering audiometry device, the Galton Whistle, and satisfied himself statistically that the ladies of London were the most beautiful in the Kingdom, whereas the less said about Aberdeen the better. Rather endearingly he embarked on writing a somewhat ill-judged erotic novel at the age of 88. Somewhat less endearingly, he also founded the Eugenics movement.

Regression may also explain many of my experiences in higher surgical training. On the rare occasion I carried out a procedure with impeccable skill and precision, my boss would shower me with praise and clap me warmly on the back. I would then of course regress to my usual level of shambolic ineptitude for the next case. However, were I to make a series of uncharacteristically witless blunders, my trainer would berate me at length and thrash me soundly, followed by regression to my usual level of tolerable adequacy. Thus one may draw the conclusion that verbal and ritual humiliation is a highly effective technique, whereas moderate praise simply leads to complacency and

incompetence. Spare the rod and spoil the child.

One piece of sage clinical advice passed on by the wizened patriarchs of the speciality in olden days was always to endeavour to return a distressed patient's phone call by sundown. On a purely statistical basis, this makes great sense. To be desperate enough to pick up the phone and fight through to one's secretary, it is highly likely the patient is at least a couple of standard deviations worse than their usual diseased state. Thus a few words of advice and a well-timed brief clinic visit are all that may be needed to gain credit for the inevitable dramatic regression to the mean. If one is tempted to ignore the patient, they will no doubt attribute their subsequent improvement to something or somebody else.

Sport is perhaps unique in the field of human endeavour for its high levels of measurable performance variance. For every Bradman grinding away at a career average six standard deviations (yes SIX) above the mean, there are millions of lesser mortals who occasionally pop up to Olympus for a day. In May 1911, a journeyman Nottinghamshire quick bowler named Ted Alletson was sent out to bat at number nine against Sussex at Hove. The game seemed lost, and the captain told Ted (whose highest score in his previous 71 games was 81) to have a swing. What followed was the most devastating 90 minutes of batting in the history of the game as he smashed 189 off the bemused bowling attack, including 139 in the last 37 minutes. Five balls were lost, one firmly wedged in the woodwork of the South Stand, and the pavilion clock smashed beyond repair. Sadly, he never

**“If we select a population on the basis of an extreme value of a variable, they are individually likely to be extreme in relation to their own personal mean.”**

passed a hundred again, and ended his career with a batting average of 18. Few in sport or indeed life have ever regressed so far from such dizzy heights.

#### References

1. Langville AN, Meyer CD. *Who's # 1?* Princeton University Press; 2012.
2. McDonald CJ, Mazzuca SA, McCabe GP. How much of the placebo 'effect' is really statistical regression? *Stat Med* 1983;**2**:417-27.



**Chris Potter, MA, FRCS (Eng), FRCS (ORL),**

Consultant ENT Surgeon,  
Torbay Hospital, Torquay, UK.

**E: potler@doctors.org.uk**

**Declaration of competing interests:**  
None declared.