

The New York Times

ERIC BRADLOW, SHANE JENSEN, JUSTIN WOLFERS AND ADI WYNER | Keeping Score FEB. 10, 2008

Report Backing Clemens Chooses Its Facts Carefully

Last week, [Roger Clemens](#) made the rounds on Capitol Hill to rebut charges by Brian McNamee, his former trainer, that he used steroids and human growth hormone late in his career. In addition, Clemens's agents from Hendricks Sports Management have provided a report loaded with numbers — 45 pages, 18,000 words and 38 charts — to support his position. You can find the report at the Web site www.rogerclemensreport.com.

But the value of evidence is not measured by the weight of a report; when examined carefully, the Clemens report does not make a convincing case for his innocence.

The report hinges on a critical question: Was Clemens's late-career success highly unusual? If so, an unusual late-career improvement lends credence to the [Mitchell report's](#) assertion that he used performance-enhancing drugs at various times from 1998 onward. The Clemens report tries to dispel this issue by comparing him with [Nolan Ryan](#), who retired in 1993 at 46. In this comparison, Clemens does not look atypical — both enjoyed great success well into their 40s. Similar conclusions can be drawn when comparing Clemens with two contemporaries, [Randy Johnson](#) and [Curt Schilling](#).

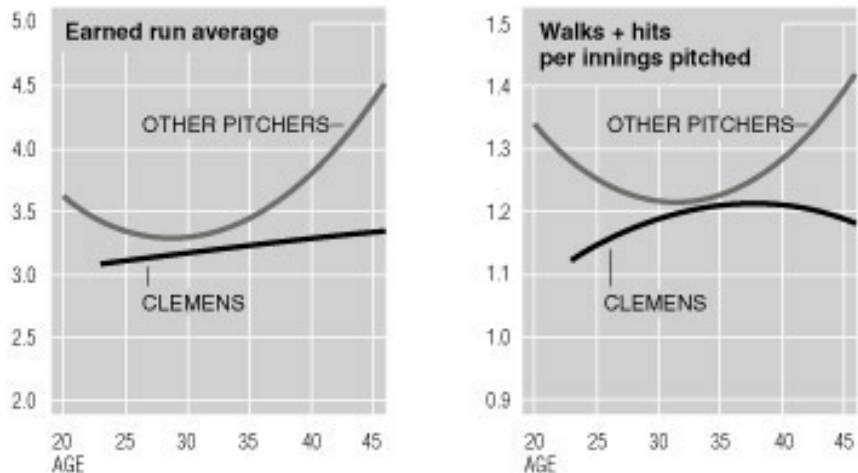
Yet such comparisons tell an incomplete story. By comparing Clemens only to those who were successful in the second act of their careers, rather than to all pitchers who had a similarly successful first act, the report artificially minimizes the chances that Clemens's numbers will seem unusual. Statisticians call this problem selection bias.

There is no doubt that Clemens was a great pitcher, but the question is whether he was much better past 36 or 37 (when he is suspected of having taken performance-enhancing drugs) than would have been expected based on his early career.

A better approach to this problem in-

Comparing Careers

A statistical comparison between Roger Clemens and the 31 other pitchers since 1968 who started at least 10 games in at least 15 seasons and pitched at least 3,000 innings. Note: A smooth curve is used to represent the data.



Sources: Eric Bradlow, Shane Jensen, Justin Wolfers and Adi Wyner

volves comparing the career trajectories of all highly durable starting pitchers. We have analyzed the progress of Clemens as well as all 31 other pitchers since 1968 who started at least 10 games in at least 15 seasons, and pitched at least 3,000 innings. For two common pitching statistics, earned run average and walks-plus-hits per innings pitched, we fitted a smooth curve to all the data from these 31 pitchers and compared it with those for Clemens's career.

Relative to this larger comparison group, Clemens's second act is unusual. The other pitchers in this durable group usually improve steadily early in their careers, peaking at around age 30. Then a slow decline sets in as they reach their mid-30s.

Clemens follows a far different path. The arc of Clemens's career is upside down: his performance declines as he enters his late 20s and improves into his

mid-30s and 40s.

The report correctly observes that he is not the only pitcher to excel at a comparatively old age, but it fails to note that he has taken an unusual path to that late-career success.

Another key shortcoming of the Clemens report is that it focuses almost exclusively on his E.R.A. But a pitcher's E.R.A. is affected by factors, like defense, that have nothing to do with his pitching. It is also affected by other factors, like the order of events — a triple, for instance, can be hit with the bases empty, or the bases loaded. So a pitcher's E.R.A. tends to bounce around a lot, and these ups and downs can help obscure patterns in career numbers.

Because E.R.A. can be so unreliable, analysts prefer to look at basic building blocks of talent like strikeout, walk, hit and home run rates. Clemens's walks-plus-hits rate, for instance, follows an

even more unusual trajectory late in his career, one that raises some suspicion.

Other measures suggest Clemens performed similarly to his contemporaries. But these comparisons do not provide evidence of his innocence; they simply fail to provide evidence of his guilt.

Our reading is that the available data on Clemens's career strongly hint that some unusual factors may have been at play in producing his excellent late-career statistics.

In any analysis of his career statistics, it is impossible to say whether this unusual factor was performance-enhancing drugs.

The Clemens report argues that his longevity "was due to his ability to adjust his style of pitching as he got older, incorporating his very effective split-finger fastball to offset the decrease in the speed of his regular fastball caused by aging." While this may be true, it is also just speculation: there is not a single number in the report quantifying the evolution of Clemens's pitch selection.

Statistics provide powerful tools for understanding the world around us, but the value of any analysis invariably comes down to choosing a useful statistic and an appropriate comparison group. Statisticians-for-hire have a tendency to choose comparison groups that support their clients. A careful analysis, and a better informed public, are the best defense against such smoke and mirrors.

Eric Bradlow, Shane Jensen, Justin Wolfers and Adi Wyner, professors at the University of Pennsylvania's Wharton School, researched and wrote this article. ■