

The story of the fellow who did

Once upon a time, there was a fellow who wanted to be an engineer but ended up doing clinical trials. Eventually he came to head the coordinating center for the XYZ Alzheimer's Disease Prevention Trial (XYZADPT). The trial was funded by the NIH and started before the NIH policy on data sharing was announced and, hence, used consents not mentioning data sharing.

The trial involved six clinics that, together, enrolled 2,500+ persons into the trial over a four year period, starting in early 2000. Followup ended early 2007.

Subsequently, the funding agency started prodding the fellow to prepare and deposit a de-identified dataset for use outside the study research group. The fellow was reluctant to comply with the request due to concerns regarding risks of violating patients' rights to privacy seeing as consents did not mention data sharing, but eventually relented being careful to follow the yellow brick road of HIPAA de-identification and careful to get approval of his IRB before sending the dataset off to the repository specified by the funding agency for repose of datasets.

About a year later, the fellow was at the check-out counter of his "favorite" supermarket, waiting for the woman ahead of him to sort her coupons. It was during the wait that his gaze landed on the collection of rags masquerading as magazines at the check-out counter. The *National Enquirer* featured a full page cover of a famous black singer with the banner headline DEVASTATED. The singer was one of the fellow's idols years back so he peeked inside to see what the devastation was about. The fellow was saddened to learn that the singer was suffering from AD and flabbergasted to see reference to the XYZADPT and the singer's participation in it.

After furtively glancing about to make certain no one was looking, he slipped the copy onto the check-out counter.

About a week hence the fellow was summoned to his IRB about the article and a few weeks later he learned that he, his institution, and the XYZADPT officers and their respective institutions were named in a breach of confidentiality lawsuit brought by relatives of the singer.

"And now the rest of the story" in the words of Paul Harvey.

About a year before the story ran, a professor at Anything Goes University (AGU) requested the XYZADPT dataset from the repository. The professor used the dataset to address his questions and then "filed" it. Some months later he dug it out and handed it to a graduate student of his just starting work on an AD dissertation.

The student's wife was a journalism student and a close friend of a free lance reporter who, at the time, was working on a story concerning a famous black singer living with AD. The reporter, during background work, had learned that the singer had been enrolled in an NIH AD prevention trial in the early 2000s. Knowing that her girl friend's husband was working on an AD dissertation, the reporter asked if her husband might know of a such a trial.

A few days later, the student's wife passed her friend's question onto him. His curiosity was peaked seeing as, just a few days back, he had been given the XYZADPT dataset by his professor. He asked why the reporter wanted to know, whereupon his spouse told him that she was working on a story about a famous black singer who had been diagnosed with AD while enrolled in a trial to prevent AD.

One thing led to another and before long the student was on his laptop in the XYZADPT dataset.

His wife told him that the reporter knew the singer had been studied in a Boston clinic. The dataset and related data dictionary did identify clinics by location, but study publications did so the student matched published enrollment figures with counts by sorting on clinic. Matching the numbers to those published allowed the student to identify clinics by name. With the Boston clinic identified, the student limited his search to black males with AD in that clinic. There was only one such person.

The fellow who did

The student's wife also volunteered that her reporter friend knew that the singer was married and that he had dropped out of school after the 8th grade. The graduate student checked those fields and was satisfied he had found the singer in his dataset.

The next time the graduate student's wife saw her free lance reporter friend, she told her what her husband had found and the rest is history.

When the smoke finally settled and the music stopped, about three years after the *National Enquirer* article was published, the fellow's institution had agreed to pay an undisclosed amount for damages, including court costs incurred by the Boston site because data were deposited without approval of the Boston IRB. The settlement required the fellow's institution to admit negligence in allowing data to be deposited without the express approval of the Boston IRB.

A side bar: While deposit was at the NIH's urging, the NIH remained silent when the shooting started. The government has to agree to be sued. It declined the invitation.

The would-be engineer turned trialist was barred from serving as an investigator on any research involving human beings by his IRB. The legal representatives of the fellow's IRB argued that the fellow had been derelict in his duties in protecting the confidentiality of information in his keeping because he failed to recognize the risk of probabilistic identification.

The fellow retired and faded away without as much as a Timex Watch from the institution as a parting memento.

Moral: Don't be duped into believing de-identification will save you if lawyers come calling.