For the Members of the Public Health Committee

Testimony in support of Raised SB 130, February 24, 2016

Susan Israel, MD

I would like to thank the members of the Public Health Committee for raising SB 130 which would enable patients to keep their health care services from inclusion in the All-Payer Claims Database. This would ensure that the APCD is accountable directly to consumers for them to consider its risk/benefit ratio and for them to be full participants in their healthcare decisions.

Rhode Island allows patients to opt-out of its APCD and does not take names and complete addresses. Connecticut takes both, along with Social Security numbers and family information. At the least, the claims should not be fully identified to the computer systems of the company managing the data, as they include every diagnosis we have had, the name of every physician we've seen, and every test, procedure and drug taken, with the dates. (1)

The data which will be released to researchers and state agencies will be de-identified, but it will still be possible to re-identify it. There is a .22% (2) rate for the data going to outside researchers which could amount to over 7000 people in CT. For the data going to Access Health Analytics, the rate is 87% (3) as it could be merged with the identified enrollment data and voter registration lists. The Dept. of Public Health could merge it with the identified hospital discharge summaries, the Tumor Registry or with its infectious disease, STI's or newborn DNA data bases. (Please see detailed explanations of the percentages later.)

In order to use the APCD for research and to merge it with the electronic medical record, a master patient index number may be assigned to each CT citizen. But that number may be re-identifiable over time as it would be possible to recognize people from all that medical and demographic information collected about them in one place. (4)

Other concerns are that researchers can re-disclose the data to third parties, making it more susceptible to hackers and that the APCD will accept only a signature as the proof that the data have been destroyed after use. The APCD carries cyber liability insurance to protect from breaches such as those recently to Anthem, Excellus BCBS, and hospitals in California and from the 40% (5) of leaks caused by employee error.

Since the price transparency provided by the APCD could be achieved by requiring providers to post their prices directly to consumers, the APCD ought to address the ongoing privacy concerns and allow patients to control who sees their medical data.

The following are further details and explanations to the oral testimony above.

I would like to add that the 2nd Court of Appeals, in their comments on the Patriot Act, questioned whether it was a violation of the 4th Amendment for the government to take metadata, including medical records, without consent.

The medical insurance claims data of the APCD inherently have many inaccuracies which is a problem for using it for research. One way this occurs is that to order testing various diagnoses are used that may not precisely reflect the patient's actual problem. This is one of the reasons there is thought to merge the APCD with the entire medical record. But this would create more problems for privacy as it will be difficult to keep a Master Patient Index number anonymous when it is paired with so much medical and demographic information.

The data given to the researchers will be fully HIPAA compliant "de-identified" data (18 identifiers removed). However, Health and Human Services has shown that even so, with the demographics of only the year of birth, three-digit ZIP for populations greater than 20,000, marital status, ethnicity, and gender, the data would have a 0.22% (2) rate of patient re-identification, that is 2,200 people per million or about 7700 people in CT. Research has shown a .04% (6) rate of re-identification just using the year of birth, 3-Digit ZIP code and gender. Both of these rates would be much higher if the accompanying medical histories were merged along with those demographics.

The Limited Data Sets, with only 16 of the 18 identifiers removed, that will be given to Access Health Analytics employees, will include additionally the full date of birth and the full ZIP code (along with the gender) which increase the likelihood of its re-identification rate to 87% (3). This could be done by merging it with other data bases such as that of voter registration lists and their own identified enrollment data.

I would like to stress that it is of concern that the DPH has so much identified medical information on CT citizens which it could merge with the APCD to re-identify it. Additionally, the identified Controlled Substance Registry data is available to the DPH, law enforcement and providers. So if a psychiatrist prescribes an anti-anxiety medication (benzodiazepine), it will be listed there and thus remove behavioral health privacy from the patient. It seems likely that behavioral health/HIV/substance abuse information will also be sent to the APCD without patient consent.

Public figures with name recognition and their families particularly will have to worry that their medical information is not leaked or hacked from all of the many state data bases or that it is ever used to pressure them.

Thank you again for your consideration of all these reasons.

References:

1. Health-Care Industry Spending More on Security But Not Ready for Cyberattack. Health IT Law & Industry Report: "... FBI Federal investigators also warned Nov. 9 (2015) that cybercriminals are increasingly using sophisticated techniques to gain access to health-care organizations' IT systems. Hackers are "doing their homework" on senior personnel before launching phishing attacks or other

campaigns aimed at accessing the troves of personal data health-care companies store. Donald Good, deputy director of the FBI's Cyber Division, said at the cybersecurity summit. These more targeted attacks can be harder to detect and are resulting in larger and larger breaches of data, he said. …."

2. Kwok P. Lafky D. Harder Than You Think: A case Study of Re-identification Risk of HIPAA-Compliant Records. 8. Anderson N. "Anonymized" data really isn't—and here's why not.http://www.amstat.org/meetings/jsm/2011/onlineprogram/AbstractDetails.cfm?abstractid=302255

3. Sweeney L. Simple demographics often identify people uniquely. Carnegie Melon University. Data Privacy Working Paper 3; 2000.

http://dataprivacylab.org/projects/identifiability/paper1.pdf & TheDataMAP. Matching known patients to known health records in Washington State 2012-2013. & TheDataMap. http://www.thedatamap.org/risks.html & TheDataMap. All the places your data may go; 2012-2013. www.thedatamap.org/.

4. Anderson N. "Anonymized" data really isn't—and here's why not. Law & Disorder/Civilization & Discontents. Ars Technica; Sep 8, 2009. www.arstechnica.com/tech-policy/2009/09/your-secrets-live-online-in-databases-of-ruin/

5. Ponemon Institute, www.ponemon.org

6. National Committee on Vital and Health Statistics Ad Hoc Work Groups for Secondary Uses ofHealthData.HearingProceedings;Aug.23,2007.www.ncvhs.hhs.gov/transcripts-minutes/transcript-of-the-august-2-2007-ad-hoc-workgroup-for-secondary-uses-of-health-data-hearing/.

Beck M. Doctors Could be Penalized for Ordering Prostate Tests. The Wall Street Journal. November 19, 2015.

.