

# Data-driven analyses revealed the comorbidity landscape of tuberous sclerosis complex

Kun-Hsing Yu, MD, PhD, Oren Miron, MA, Nathan Palmer, PhD, Dario R. Lemos, PhD, Kathe Fox, PhD, S.C. Kou, PhD, Mustafa Sahin, MD, PhD, and Isaac S. Kohane, MD, PhD

## Correspondence

Dr. Kohane  
Isaac\_Kohane@  
hms.harvard.edu

*Neurology*® 2018;91:974-976. doi:10.1212/WNL.0000000000006546

Tuberous sclerosis complex (TSC) is a genetic disorder affecting multiple organ systems observed in approximately 1 in 10,000 live births.<sup>1,2</sup> Patients with TSC have a higher risk of developing benign tumors in the brain, kidneys, heart, liver, lungs, and skin, and a variety of neuropsychiatric disorders.<sup>3</sup> However, due to the rarity of TSC, its comorbidity landscape has not been fully investigated. To characterize the comorbidities of TSC systematically, we employed a nationwide health insurance claims dataset to analyze the comorbidity landscape of TSC and determined the phenotypical differences between patients with TSC and participants without TSC matched on age and sex.

## Methods

Using unidentifiable member claims data from Aetna, which covers 45 million patients in the United States, we identified patients with TSC by the ICD-9 code 759.5 during the time frame from January 1, 2008, to February 29, 2016. To identify the comorbidity difference between participants with and without TSC with greater statistical power, for each patient with TSC, we selected 500 unique participants without TSC with the same age (in years) and sex. We mapped all ICD-9 codes associated with the participants in the insurance claims dataset to 1,865 phenome-wide association studies (PheWAS) codes,<sup>4</sup> and employed the standard Fisher exact test for each PheWAS code to identify the phenotype differences between participants with and without TSC.<sup>4</sup> We further examined the affected proportion of each PheWAS code in participants with and without TSC and used the insurance enrollment status to account for the number of at-risk patients in each age group. Since TSC is often diagnosed and treated by pediatric neurologists, we employed the Wilcoxon signed rank test to identify the difference in the affected proportion of each comorbidity between TSC and non-TSC groups between the ages of 0 and 20, and used the odds ratio (OR) to determine the effect size for patients aged 0–20. We controlled the false discovery rate at 0.05 using the Benjamini-Hochberg procedure. All analyses were performed using R version 3.3 with packages *icd* and *NMF*. This study was approved by the Harvard Medical School institutional review board. Consent to use the data was included in the approval.

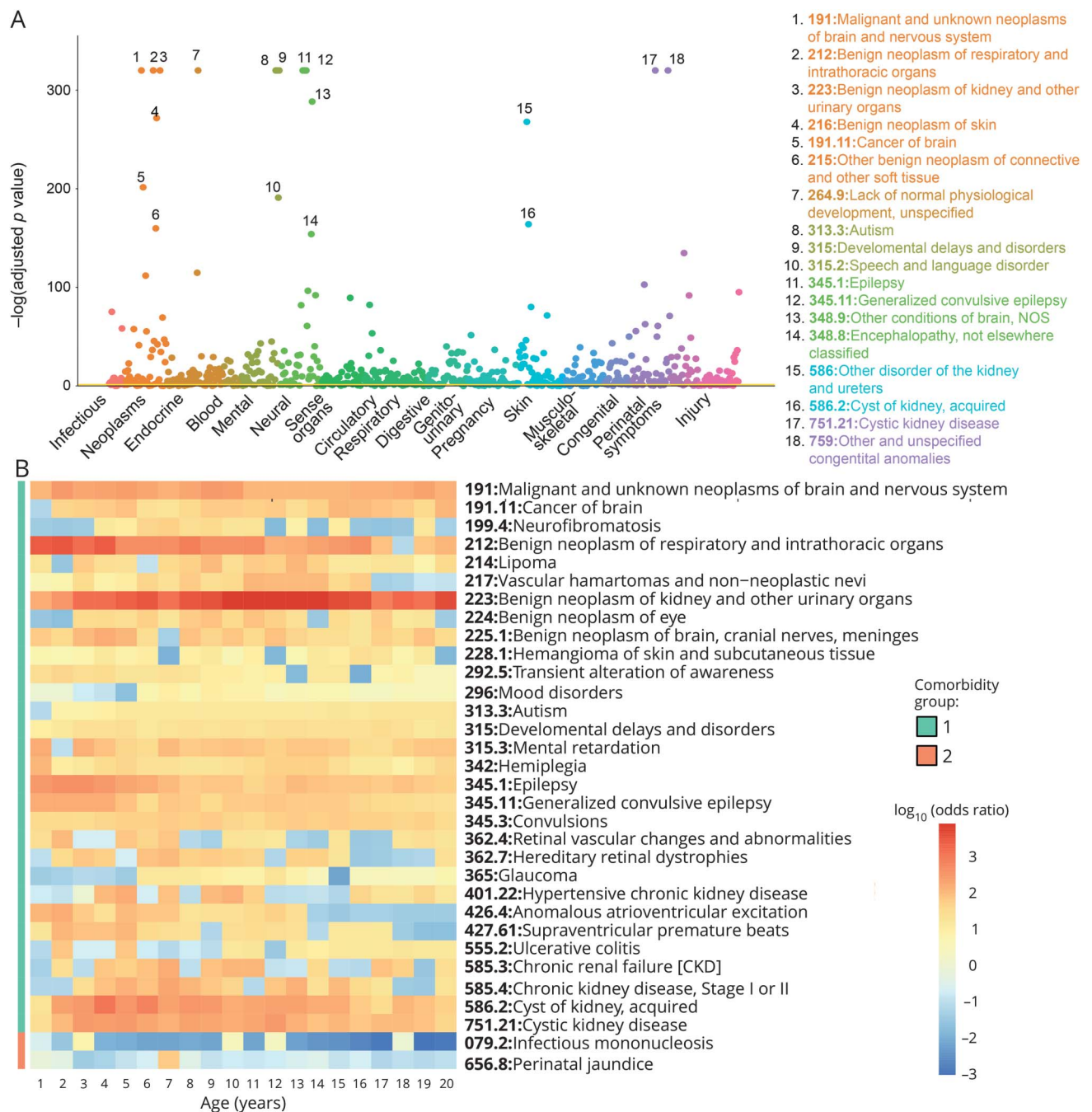
## Results

We identified 3,131 patients with TSC in the dataset. The mean age of patients with TSC in this cohort is 34.56 years, and the SD is 23.10 years. A total of 47.24% of the patients are male. The mean length of time a patient with TSC is covered is  $4.43 \pm 2.56$  years, while that for participants without TSC is  $3.99 \pm 2.60$  years. A global comorbidity landscape is shown in the figure. Consistent with the literature, patients with TSC under 20 years old ( $n = 1,296$ ) were more likely to develop benign neoplasm of the kidney (OR 2,265.3, 95% confidence interval [CI] 1,593.9, 3,219.6) (absolute odds in TSC; non-TSC 0.140;  $6.17 \times 10^{-5}$ ), malignant neoplasms of the brain (OR 177.2, 95% CI 141.9, 221.4) (absolute odds in TSC; non-TSC 0.0927;  $5.23 \times 10^{-4}$ ), epilepsy (OR 103.4, 95% CI 90.2, 118.5) (absolute odds in TSC; non-TSC 0.303;

From Harvard Medical School (K.-H.Y., O.M., N.P., I.S.K.), Boston; Harvard University (K.-H.Y., S.C.K.), Cambridge; Brigham and Women's Hospital (D.R.L.), Boston, MA; Aetna Inc. (K.F.), Hartford, CT; and Boston Children's Hospital (M.S., I.S.K.), MA.

Go to [Neurology.org/N](http://Neurology.org/N) for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

**Figure** The comorbidity landscape of patients with tuberous sclerosis complex (TSC)



(A) A Manhattan plot shows the significance level of TSC comorbidities. Common TSC comorbidities, including neoplasms of the brain, kidney, and respiratory organs, autism, epilepsy, developmental delays, and speech and language disorder, have the most significant adjusted *p* values. A total of 605 phenome-wide association studies (PheWAS) codes showed significant associations with TSC after correcting for multiple testing. (B) Distinct comorbidity patterns were observed in the TSC and non-TSC cohorts between age 0 and 20. Selected comorbidities with the most significant odds ratios (ORs) are shown. Patients with TSC are more likely to develop benign neoplasm of the kidney (OR 2,265.3), malignant neoplasms of the brain (OR 177.2), epilepsy (OR 103.4), and autism (OR 15.7), shown as comorbidity group 1, but less likely to develop infectious mononucleosis (OR 0.410) and perinatal jaundice (OR 0.720), shown as comorbidity group 2. PheWAS codes and descriptions were shown.

2.93\*10<sup>-3</sup>), and autism (OR 15.7, 95% CI 13.4, 18.4) (absolute odds in TSC; non-TSC 0.168; 0.0107), compared with the age- and sex-matched group. Based on our analysis, we further identified many less-reported comorbidities of TSC, including ulcerative colitis (OR 6.94, 95% CI 3.09, 15.6)

(absolute odds in TSC; non-TSC 4.65\*10<sup>-3</sup>; 6.70\*10<sup>-4</sup>) and mood disorders (OR 3.37, 95% CI 2.49, 4.58) (absolute odds in TSC; non-TSC 0.0343; 0.0102). In contrast, patients with TSC have significantly less infectious mononucleosis (OR 0.410, 95% CI 0.184, 0.914) (absolute odds in TSC; non-TSC

4.65\*10<sup>-3</sup>; 0.0113) and perinatal jaundice (OR 0.720, 95% CI 0.540, 0.961) (absolute odds in TSC; non-TSC 0.0385; 0.0534) (figure). All of the reported differences are statistically significant after correcting for multiple testing.

## Discussion

TSC is a relatively rare disease and it is difficult for pediatric neurologists to assemble data from thousands of patients, even in tertiary medical centers.<sup>5</sup> However, with a nationwide health insurance database, we identified more than 3,000 patients with TSC, and successfully quantified their phenotypic differences comparing to an age- and sex-matched non-TSC group in a systematic fashion. Patients with TSC are more likely to develop ulcerative colitis and psychiatric conditions; this should guide practitioners to look out for psychiatric and medical conditions in their patients. Additional studies using complementary datasets can confirm the identified comorbidities. Our approach could facilitate the generation of hypotheses for rare diseases,<sup>6</sup> and we can follow-up with the evolving phenotype of the patients during the whole period of health insurance coverage.

## Author contributions

K.-H. Yu: study concept and design, acquisition of data, analysis and interpretation of data, drafting the manuscript. O. Miron: critical revision of the manuscript for intellectual content. N. Palmer: acquisition of data, critical revision of the manuscript for intellectual content. D.R. Lemos: critical revision of the manuscript for intellectual content. K. Fox: acquisition of data, critical revision of the manuscript for intellectual content. S.C. Kou: critical revision of the manuscript for intellectual content. M. Sahin: critical revision of the manuscript for intellectual content. Isaac Kohane: acquisition

of data, critical revision of the manuscript for intellectual content, study supervision.

## Study funding

Study funded by the Harvard Data Science Fellowship. The computational infrastructure was supported by the NVIDIA Grant Program.

## Disclosure

K. Yu, O. Miron, N. Palmer, and D.R. Lemos report no disclosures relevant to the manuscript. K. Fox is employed by Aetna. S. Kou reports no disclosures relevant to the manuscript. M. Sahin received research funding from Roche, Novartis, Pfizer, and LAM Therapeutics, and served on the Scientific Advisory Board of Sage Therapeutics and PTEN Research Foundation and on the Professional Advisory Board of the Tuberous Sclerosis Alliance. I. Kohane reports no disclosures relevant to the manuscript. Go to [Neurology.org/N](http://Neurology.org/N) for full disclosures.

## Publication history

Received by *Neurology* March 1, 2018. Accepted in final form August 27, 2018.

## References

1. Hallett L, Foster T, Liu Z, Blieden M, Valentim J. Burden of disease and unmet needs in tuberous sclerosis complex with neurological manifestations: systematic review. *Curr Med Res Opin* 2011;27:1571–1583.
2. Osborne JP, Fryer A, Webb D. Epidemiology of tuberous sclerosis. *Ann NY Acad Sci* 1991;615:125–127.
3. Curatolo P, Moavero R, de Vries PJ. Neurological and neuropsychiatric aspects of tuberous sclerosis complex. *Lancet Neurol* 2015;14:733–745.
4. Denny JC, Ritchie MD, Basford MA, et al. PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. *Bioinformatics* 2010;26:1205–1210.
5. Dabora SL, Jozwiak S, Franz DN, et al. Mutational analysis in a cohort of 224 tuberous sclerosis patients indicates increased severity of TSC2, compared with TSC1, disease in multiple organs. *Am J Hum Genet* 2001;68:64–80.
6. Khoury MJ, Ioannidis JP. Medicine: big data meets public health. *Science* 2014;346:1054–1055.

---

## Visit the *Neurology*<sup>®</sup> Website at [Neurology.org/N](http://Neurology.org/N)

- More article-based content on home pages
- Streamlined menus and navigation
- Enhanced blog sections for specialty areas
- Same experience on desktop, tablet, and mobile devices
- Audio summaries of current issues
- Improved article reading experience; links more evident (pdf, analytics, social media)
- *Neurology*<sup>®</sup> *Clinical Practice* initiative “Practice Current” global surveys will be accessible across sites

**f** Find *Neurology*<sup>®</sup> on Facebook: <http://tinyurl.com/neurologyfan>

**t** Follow *Neurology*<sup>®</sup> on Twitter: <https://twitter.com/GreenJournal>