Journal of Population Ageing
https://doi.org/10.1007/s12062-026-09571-7

RESEARCH

Variation in multimorbidity between and within age groups

Joel E. Cohen[1,2,3] · T. J. Lin[4] · Alice Eggleston[5] · Mirta Milanes[5] · Amro Hassan[5] · Andrea Cassells[4] · Jonathan N. Tobin[4,6]

Received: 12 January 2026 / Accepted: 8 June 2026

Abstract

The chronic conditions associated with age are an increasing challenge to individual and public health in many countries. In general, older age groups have a greater average multiplicity of chronic conditions (greater "multimorbidity"). But the variation (not only the average) of multimorbidity within groups of increasing age seems not to have been studied empirically. The variation in multimorbidity in groups of increasing age could, in principle, increase or decrease. We use the enhanced Charlson Comorbidity Index (eCCI) as a quantitative index of a person's concurrent chronic medical conditions (multimorbidities) in 39 categories, weighted by severity. We demonstrate systematic relationships between the mean eCCI and the variance of eCCI of people by age group in de-identified electronic health records of 238,156 adults served by four United States health systems (HSs), two in Chicago, two in New York City. The relation of variance to mean eCCI *across* age groups is described by a generalization of Taylor's power law. *Within* almost every age group, the frequency distribution of eCCI per individual is well approximated by a negative binomial distribution. The means and the variances of the fitted negative binomial distributions *within* age groups account approximately for the relations of mean and variance of eCCI *across* age groups. The clear findings that variance of eCCI within an age group rises with the mean eCCI, and that the variance exceeds the mean, imply that increasing average multimorbidity brings more variable or heterogeneous multimorbidity. Both clinical care and health policy need to take account of the variation or heterogeneity of multimorbidity among older adults.

Keywords Charlson Comorbidity Index · Comorbidity · Electronic health records (EHR) · Federally Qualified Health Centers (FQHC) · Fluctuation scaling · Negative binomial distribution · Practice-based research network (PBRN) · Taylor's law · Variance function

Extended author information available on the last page of the article

Published online: 01 July 2026

Introduction

Multimorbidity is defined as the presence of multiple chronic conditions in one individual. Multimorbidity focuses on the entire health of an individual. By contrast, comorbidity commonly centers on one chronic disease and its possibly associated, or "comorbid," chronic conditions. Here we describe and model mathematically the variation in multimorbidity between and within age groups of some low-income and minority urban Americans. Our findings invite parallel investigations of other populations to test the generality of the new patterns we reveal.

Chronologically older individuals are more likely to experience more multimorbidity, given greater exposure time at risk of developing new diagnoses (Beard et al., 2016, p. 2146). But the average and the variation of multimorbidity between and within age groups are little understood quantitatively. For example, the quantitative relationship between the mean and the variance of a quantitative index of age-specific multimorbidity as adult age increases appears not to have been examined (i.e., Divo et al., 2014; Chen et al., 2022, 2023; Chowdhury et al., 2023; Bensken et al., 2024). Here we quantify and describe mathematically systematic patterns in a quantitative index of multimorbidity at two levels of analysis: between groups of increasing adult age, and within age groups. We relate within-group and between-group averages and variation in multimorbidity.

We are motivated to examine the quantitative relationships between the mean and the variance of multimorbidity at different adult ages by the widespread observation of a power-law relation (now often called Taylor's law) between the mean and the variance in ecology (Taylor 1961, 2019; Cohen et al. 2013a), demography (Cohen et al. 2013b; Xu and Cohen 2021; Benassi et al. 2023a, b; among others), earth sciences (Calif & Schmitt, 2015; Tippett & Cohen, 2016; among others), epidemiology (Keeling & Grenfell, 1999; Keeling et al., 2001; Cohen, Davis, and Samorodnitsky, 2022; Benassi, Naccarato, and Xu, 2023b; among others), and many other fields (Eisler, Bartos and Kertesz, 2008). Physicists refer to a power-law variance function as fluctuation scaling.

Here, Taylor's law would predict that, for each specific adult age group, the variance of an index of multimorbidity approximates a power of the mean of that index. The data here reject Taylor's law in its original form but support a generalization of Taylor's law known as a quadratic Taylor's law. *Within* groups of increasing adult age, we examine the distribution of this index of multimorbidity across individuals of the same age to see how well this microscopic level of variation can account for the variation of multimorbidity across ages. We find that, within each age group, the frequency distribution of this index of multimorbidity is well approximated by a negative binomial distribution.

Data, definitions, and methods

Data source

From the electronic health records (EHRs) of 238,156 adults in four health systems (HSs) of networks of Federally Qualified Health Centers (FQHCs) in low-income communities

in New York City and Chicago in the United States, we extracted each person's chronological age at the date of analyzing the EHR and an integer-valued index of multimorbidity called the "enhanced Charlson Comorbidity Index" (henceforth eCCI) (Charlson et al., 2025). These data were accessed for research purposes on 2024-12-28. We assessed each person's age and eCCI once only. This is a cross-sectional, not a longitudinal study.

The Health Resources and Services Administration (HRSA) Health Center Program Uniform Data System (UDS) provides a publicly available summary of the four populations served by the four HSs in this study (https://data.hrsa.gov/topics/healthcenters/uds/overview?grantNum=H80CS00597).

Age groups

The New York HSs 1 and 2 gave each adult's age in whole years plus a decimal fraction of a year based on the number of days from the date of birth to the date on which eCCI was computed. The Chicago HSs 3 and 4 rounded age to the nearest whole year.

We constructed two systems of age groups, one year wide and four years wide.

For the one-year groups, we rounded ages from NY to the nearest year, like the ages from Chicago. We then excluded from further analysis single years of age with fewer than 100 people (here, only some of the oldest adults) because age groups with few people were judged to give unreliable estimates of the mean and variance of multimorbidity. The numbers of people in each one-year group with 100 or more people, by age and HS, are plotted in Supplementary Information (hereafter SI) Fig. 1.

To give statistically credible information about the relative frequency of each eCCI in each age group, we devised 16 age groups containing four years of chronological age each: 19–22 years (that is, from 20 minus one up to and including 20 plus two), 23–26 (that is, from 24 minus one up to and including 24 plus two), and so on up to 79–82 (that is, from 80 minus one up to and including 80 plus two). For brevity in figures and tables, we identified each such age group by its year of chronological age that is divisible by 4. For example, the 19–22 group is called 20 and the 79–82 group is called 80. SI Fig. 2 shows the number of people in each age group in each HS.

Enhanced Charlson comorbidity index (eCCI)

The eCCI of a patient is the sum of the points assigned in Table A.1 of Charlson et al. (2025, pp. 6–7, Appendix A) to each chronic condition in the patient's EHR. For each age group separately, we analyzed the mean, variance, and probability distribution (frequency histogram) of eCCI.

The eCCI used here is designed to anticipate risk of hospitalization and disability (Charlson et al. 2025). The first three chronic conditions in the eCCI with lowest "Comorbidity Weight" equal to 1 each are myocardial infarction, congestive heart failure, and peripheral vascular disease or bypass. The last three chronic conditions in the eCCI with highest "Comorbidity Weight" equal to 6 each are metastatic solid tumor, HIV or AIDS, and any transplant: renal, heart, liver, bone marrow, or lung. Of the 39 conditions in the eCCI, only these three conditions have weight 6. The remaining conditions have weight 1 (a total of 13 conditions), 2 (11 conditions), or 3 (12 conditions).

Earlier versions of Charlson's Comorbidity Index were designed to anticipate the risk of death (Charlson et al. 1994, Charlson et al. 2022, Charlson and Wells 2022; see also Ly et al., 2025 and the on-line calculator at https://www.mdcalc.com/calc/3917/charlson-comorbidity-index-cci).

Two of the four HSs (HS 1 = NYC-a and HS 2 = NYC-b) are in New York City and two (HS 3 = Chicago-a and HS 4 = Chicago-b) are in Chicago, Illinois. In New York, the two health systems are Family Health Centers at NYU Langone and Community Healthcare Network. In Chicago, the two health systems are Erie Family Health Centers and Friend Health. The HSs and individuals are fully de-identified for this analysis because the data did not identify any individual or which HS was which in either city.

SI Table 1 summarizes statistics of age and eCCI in the people included in this study, by HS. The data on HS, chronological age, and eCCI for each individual, on which this study rests, are posted in the SI file HealthSystemAgeCCI20250323.csv.

Variance functions

The variance function of eCCI in a set of age groups shows the variance of eCCI in each age group on the vertical axis as a function of the mean eCCI at that age on the horizontal axis.

Taylor's law (TL) (Taylor, 1961) proposes that, with $a > 0$ and ordinarily $b > 0$,

$$\log_{10}(variance\ of\ eCCI\ in\ age\ group\ x) \approx a + b \log_{10} mean\ of\ eCCI\ in\ age\ group\ x \tag{1}$$

The variance-mean power law, $variance\ of\ eCCI\ in\ age\ group\ x \approx 10^{a}(mean\ of\ eCCI\ in\ age\ group\ x)^{b}$, is mathematically equivalent to (1).

The quadratic Taylor's Law (Taylor et al., 1978, p. 388, Eq. (14) (QTL) proposes that

$$\begin{aligned}\log_{10}(variance\ of\ eCCI\ in\ age\ group\ x) \approx \\ a + b \log_{10}(mean\ of\ eCCI\ in\ age\ group\ x) \\ +c[\log_{10}(mean\ of\ eCCI\ in\ age\ group\ x)]^2,\end{aligned} \tag{2}$$

where the constant c may be positive or negative. When $c = 0$, the QTL is just TL.

In each HS, to estimate the coefficients a, b, c, we fitted linear and quadratic regressions by ordinary least squares (OLS) using Matlab function regress (Matlab, 2024).

For each HS separately, we tested whether QTL described the data significantly better than TL by an F-test:

$$F = \frac{\frac{RSS_{TL} - RSS_{QTL}}{df_{TL} - df_{QTL}}}{\frac{RSS_{QTL}}{df_{QTL}}} \tag{3}$$

where RSS_{TL}, RSS_{QTL} are the residual sums of squares of TL and QTL, respectively, and df_{TL}, df_{QTL} are the corresponding degrees of freedom. Then, if *fcdf* is the cumulative distribution function of the F distribution with degrees of freedom df_{TL}, df_{QTL}, we calculate

$$P = 1 - fcdf\left(F, df_{TL}, df_{QTL}\right). \tag{4}$$

If P is larger than a critical value, e.g., 0.05, then QTL is not a significantly better description of the data than TL. If P is smaller than a critical value, then QTL is a statistically significant improvement over TL.

Negative binomial distribution

The negative binomial distribution (NBD) is a classical integer-valued nonnegative random variable that is "over-dispersed" in the sense that its variance always exceeds its mean (unlike the Poisson and binomial distributions, for example). The NBD can be parameterized in various ways. The probability density function of the NBD in Matlab (https://www.mathworks.com/help/stats/nbinpdf.html), which we used for these computations, is: for $0 < p < 1,\ r > 0$ (r may or may not be an integer),

$$Pr\left(eCCI = x\right) = \left(\frac{\Gamma\left(r+x\right)}{\Gamma\left(r\right)\Gamma\left(1+x\right)}\right) p^r (1-p)^x, \quad x = 0,\ 1,\ 2,\ 3,\ \ldots \tag{5}$$

If eCCI is negative binomially distributed as in (5), the population mean of eCCI is

$$E\left(eCCI\right) = \frac{r\left(1-p\right)}{p} \tag{6}$$

and the population variance of eCCI is

$$Var\left(eCCI\right) = \frac{r\left(1-p\right)}{p^2}. \tag{7}$$

If r is constant and only p varies, then

$$Var\left(eCCI\right) = E\left(eCCI\right) + \left(\frac{1}{r}\right)\left[E\left(eCCI\right)\right]^2 \tag{8}$$

and the variance is a convex increasing quadratic function (8) of the mean, not consistent with TL (1). In the limit as r approaches infinity, the variance function (8) asymptotically becomes $Var\left(eCCI\right) = E\left(eCCI\right)$, the variance function of the Poisson distribution, which is consistent with TL with $a = 0,\ b = 1$. In the limit as r approaches 0, the second term in (8) dominates the first and (8) asymptotically becomes TL (1) with $10^a = 1/r$ and $b = 2$. We estimated the parameters p and r of the NBD by maximum likelihood using nbinfit in Matlab (2024b).

Results

Mean and variance of eCCI by age group

In each HS, for each one-year age group with 100 or more individuals, the mean eCCI and the variance of the eCCI of the individuals of that age generally increase with age (Fig. 1), but not in HS 4 for ages above roughly 65. A Poisson distribution could not adequately describe the frequency distribution of eCCI by age group in any HS because in a Poisson distribution, variance equals mean. Here variance generally exceeds mean, except for some of the youngest age groups.

It seems plausible to speculate that the decline in HS 4 in the mean and the variance of eCCI of the individuals above roughly age 65 reflects the selective removal, by institutionalization or death, of individuals with higher multimorbidity. The cross-section of the active EHRs on a single date does not permit testing this speculation, as this cross-section combines the effects of age (an individual attribute), aging (the process of growing older) and cohort (unique generational exposures).

Variance function

Figure 2 compares the (age, eCCI) data with two models: Taylor's law (TL) and the quadratic Taylor's law (QTL).

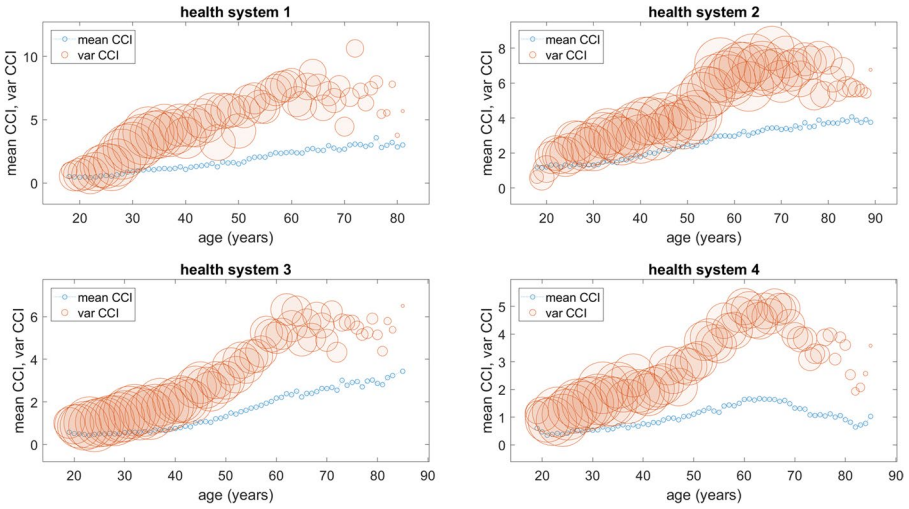

Fig. 1 The mean eCCI (blue circles) and the variance (red circles) of the eCCI of the individuals of each single year of age with 100 or more individuals, by HS. The area of each red circle is proportional to the number of individuals of the age corresponding to the abscissa of the center of the circle. The smallest circles at the extreme right of each panel show ages with 100 individuals. Source: MeanCCIvsAgeVarCCIvsAge4HSagemin100_20250131-122802.png

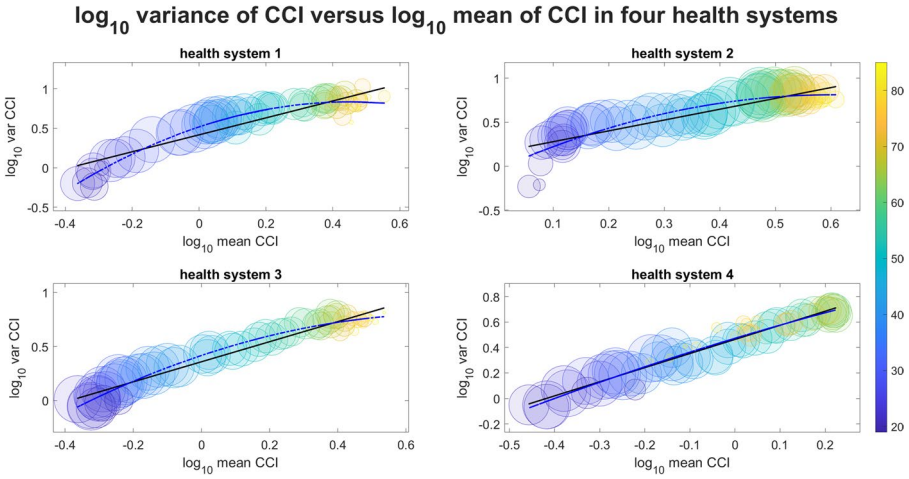

Fig. 2 The $\log_{10}$ variance of the eCCI of the individuals of each one-year age group with 100 or more individuals as a function of the $\log_{10}$ mean eCCI of each group (center of the circles) and two models: Taylor's law (1) (black solid straight line) and the quadratic Taylor's law (2) (blue dashed curved line). The area of each circle is proportional to the number of individuals in the corresponding age. The color of each circle represents age from youngest (deep purple or blue) to oldest (yellow). Source: logVarCCIvslogMeanCCITLQTLall4HSagemin100_20250131-140652.png

Taylor's law

SI Table 2 gives the estimated coefficients a, b of Taylor's law (1) and their nominal 95% confidence intervals (CIs). The mean of eCCI accounted for most of the variation in the variance of eCCI: the coefficient of determination R^2 varied from a low of 0.81 in HS 2 to the high of 0.95 in HS 3 and HS 4. In HS 1, b does not differ from 1, while b is less than 1 in HS 3 and significantly greater than 1 in HS 2 and HS 4. The coefficient a exceeds 0 in all four HSs, whereas in a family of Poisson distributions, $a \approx 0$.

Non-linearity of log variance as a function of log mean

In Fig. 2, on log-log coordinates, the variance functions of HS 1, HS 2, and HS 3 appear to be concave but that of HS 4 is very close to linear. In HS 4, the unusual declines with increasing age in both the mean and the variance of eCCI above age roughly 65 show that, at least in this instance, the mean eCCI predicts variation in eCCI better than an increasing function of age predicts variation in eCCI.

SI Table 3 gives the OLS estimates and 95% confidence intervals of the parameters of the quadratic Taylor's law (2). The mean of eCCI accounted for even more of the variation in the variance of eCCI than the linear Taylor's law (1): the coefficient of determination R^2 varied from a low of 0.85 in HS 2 to the high of 0.97 in HS 3. The confidence intervals of c, the coefficient of the quadratic term, lie below zero for HSs 1, 2, 3, confirming the visual impression that their variance functions are concave on log-log coordinates. By contrast, the coefficient of c for HS 4 does not differ significantly from 0, again confirming the visual impression that its variance function

is approximately linear on log-log coordinates. SI Table 4 gives the statistics of the TL and QTL regressions. The last two columns, Fcompare and P, strongly confirm that QTL describes the variance functions of all four HSs better than TL, and that the log-log variance functions are significantly concave for HSs 1, 2, and 3.

Frequency histogram of eCCI within each age group

The frequency histogram of eCCI within an age group in a HS shows how many people of that age and HS had eCCI equal to 0, how many had eCCI equal to 1, how many had eCCI equal to 2, and so on. Each HS has 16 four-year age groups, labeled 20, 24, …, 80. For all 64 of these age groups (16 age groups for each of 4 HSs), the variance of eCCI exceeded the mean eCCI. Figures 3, 4, 5 and 6 show the observed counts (black × marks) of people in HS 1–4, respectively, who have each eCCI and the expected counts (green line) from the fitted NBD. As is apparent from visual inspection, the agreement of the empirical counts with the NBD is generally excellent. Some deviations arise at the largest values of eCCI where fewer people are observed. Because the vertical axis (number of people) is on a logarithmic scale, the rare counts of zero people are omitted.

The estimated parameters of the NBD fitted to the empirical frequency distributions of eCCI account qualitatively for the approximation to Taylor's law with exponent near or less than $b = 1$ in these data. Using the estimated parameter values r, p (shown in SI Fig. 3), Fig. 7 plots the predicted population variance $Var\,(CCI) = r(1-p)/p^2$ as a function of the predicted mean $E\,(CCI) = r(1-p)/p$ for each age group in each HS (blue × markers) and a (black) curve, a power function corresponding to Taylor's law, fitted by least squares to the 16 values of $(r(1-p)/p, r(1-p)/p^2)$ for the 16 age groups in each HS. Only for HS 2 is $b > 1$ significantly, assuming the

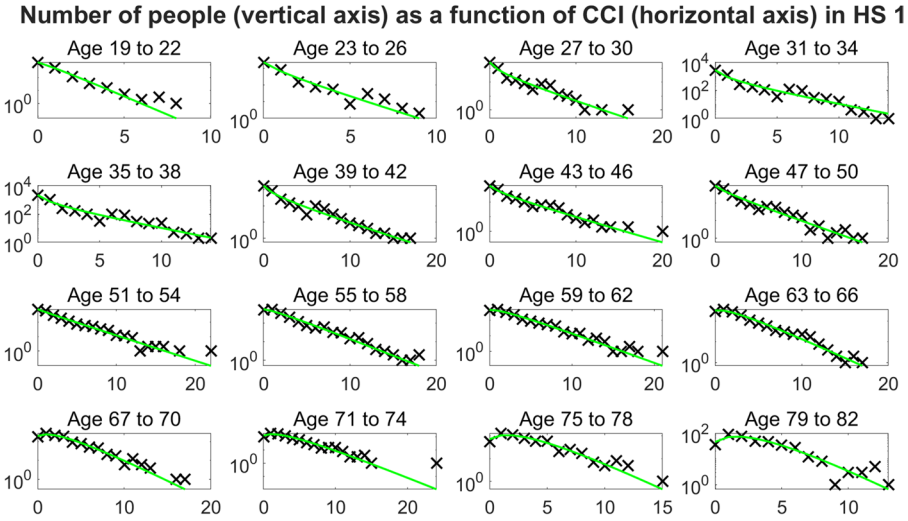

Fig. 3 For HS 1, the number of people in each 4-year age group (vertical coordinate of black ×) who have each value of eCCI (horizontal coordinate of black ×), and the fitted negative binomial distribution (green line). Source: PeopleVsCCIbyAgeHS1_20250115-160715.png

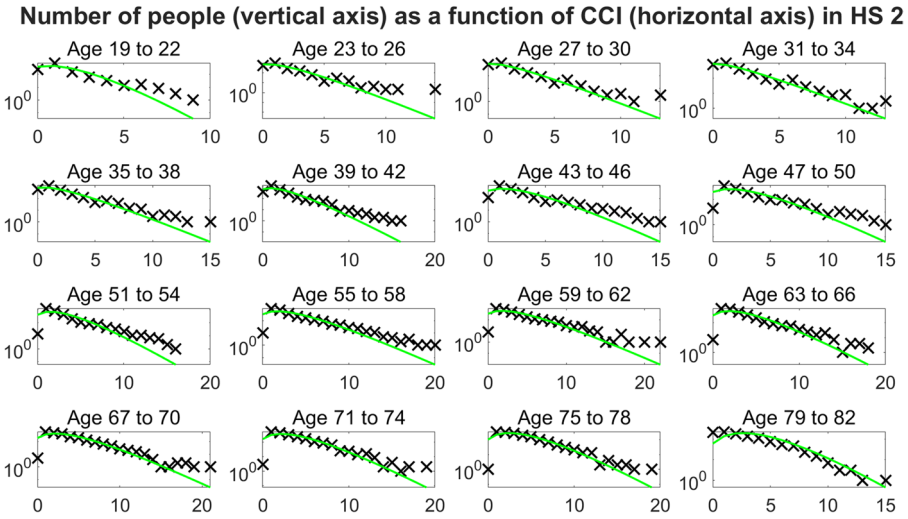

Fig. 4 For HS 2, the number of people in each 4-year age group (vertical coordinate of black ×) who have each value of eCCI (horizontal coordinate of black ×), and the fitted negative binomial distribution (green line). The sample variance differs most from the variance predicted by the negative binomial distribution in the age groups 36–54. The other HSs do not show similar deviations in these age groups. Source: PeopleVsCCIbyAgeHS2_20250115-160715.png

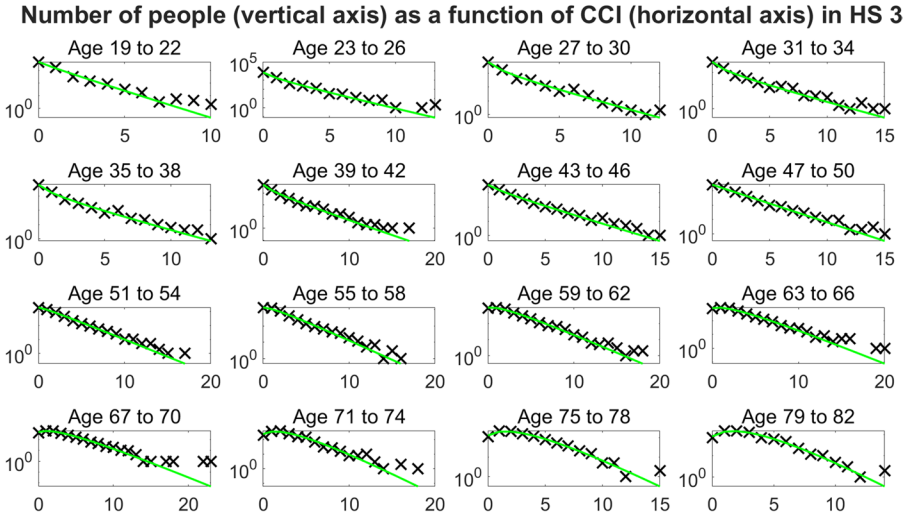

Fig. 5 For HS 3, the number of people in each 4-year age group (vertical coordinate of black ×) who have each value of eCCI (horizontal coordinate of black ×), and the fitted negative binomial distribution (green line). Source: PeopleVsCCIbyAgeHS3_20250115-160715.png

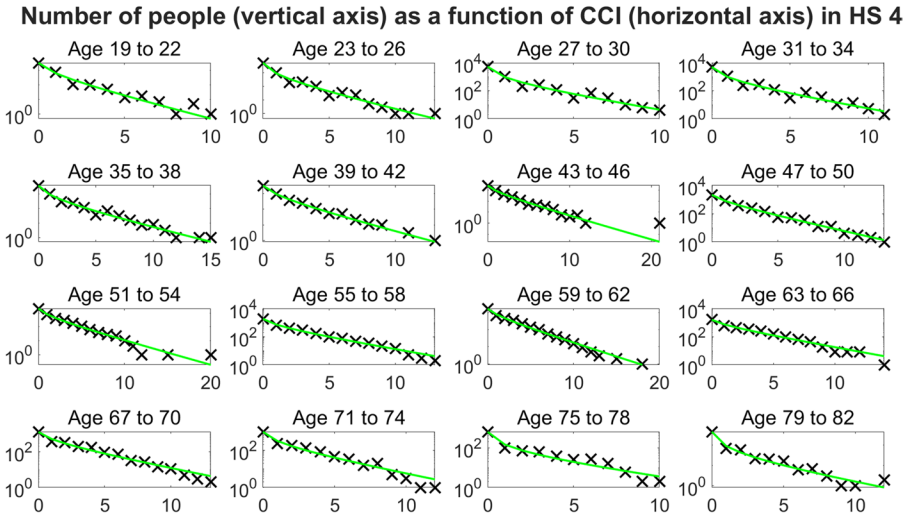

Fig. 6 For HS 4, the number of people in each 4-year age group (vertical coordinate of black ×) who have each value of eCCI (horizontal coordinate of black ×), and the fitted negative binomial distribution (green line). Source: PeopleVsCCIbyAgeHS4_20250115-160715.png

correctness of the calculated confidence interval. In HS 1 and 3, $b < 1$ significantly, and b is not significantly different from one in HS 4.

For HS 1 (New York) and HS 3 (Chicago) in Fig. 7, the age-specific predicted $Var\,(CCI)$ (blue × marks) increases as a concave function of the age-specific predicted mean $E\,(CCI)$ compared to the fitted power-law (black) curve. This observation is consistent with the curved path of the circles in Fig. 2's top left panel for HS 1 (New York). Likewise, in Fig. 2, New York's HS 2 gives an M-shaped trajectory of circles for $Var\,(CCI)$ that is roughly concordant with the blue × markers in Fig. 7, top right panel, compared to the fitted TL. Similarly, the panels for Chicago in the bottom row of Fig. 7, like the panels for Chicago in the bottom row of Fig. 2, roughly support TL.

Discussion

Summary and interpretation of main results

We show that in four HSs, two in New York City and two in Chicago, the variation (technically, the variance) in the enhanced Charlson Comorbidity Index (eCCI) is systematically related to the average (or mean) of eCCI over all ages, including above age 65, in approximate quantitative agreement with the quadratic Taylor's law (QTL). This pattern is consistent with the negative binomial distribution (NBD) of eCCI within each age group.

In all four HSs, in each age group, the variance of eCCI substantially exceeds the mean of eCCI. This observation rejects a Poisson model of variation in eCCI. Poisson variation, in which the variance of eCCI equals the mean of eCCI, would arise if

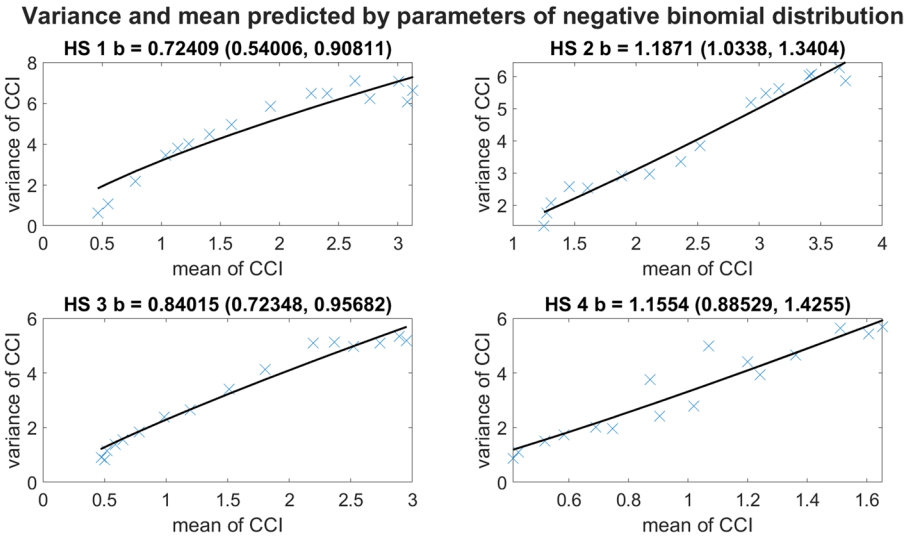

Fig. 7 The variance $Var\,(eCCI) = r(1-p)/p^2$ of a negative binomial distribution of individual eCCI as a function of the mean $E\,(CCI) = r(1-p)/p$ for each age group in each HS (blue × markers) and a (black) curve, corresponding to a fitted Taylor's law (1). Source: NegativeBinomialPredictedTLbyHS_20250117-115016.png

every cause of an increase by one unit in eCCI were independent of every other cause and had an equal probability of causing an increase in eCCI. The medical conditions that are counted in eCCI are known *not* to be independent and *not* to be of identical probability. Moreover, a majority of the chronic conditions listed in the definition of eCCI are given a weight greater than 1, so the eCCI is not incremented by independent units. The sample of people from each HS has more than Poisson variation in eCCI, suggesting that the factors that contribute to eCCI are positively associated, i.e., they tend to vary together. Moreover, these associations may change with age.

The variance and the mean of eCCI rise with age, at least up to age around 65, in all four HSs. Aging brings increasing average multimorbidity (known already) and increasing diversity of multimorbidity (described here). Increasing age brings some older people more, and some older people less, multimorbidity than would be expected from the increase in the mean of eCCI alone. Clinical practice guidelines should recommend protocols that take account of the increased diversity of multimorbidity among older patients. To illustrate the problem, Martin et al. (2024, p. 2) pointed out that "only 1 of 17 [clinical practice] guidelines approved by the Australian National Health and Medical Research Council (NHMRC) considered multimorbidity." They added that "Age alone is a limited guide to heterogeneity, as socioeconomic factors strongly impact age of onset of multimorbidity." Many clinical practice guidelines address multimorbidity indirectly through recommendations for subgroups of patients with a primary diagnosis who also have additional chronic conditions, e.g., hypertension and diabetes (American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines, 2025, Sect. 5.3 on comorbidities).

Future research

We found that the negative binomial distribution (NBD) describes well the frequency distribution (histogram) of eCCI within nearly every four-year age group in all four HSs. The NBD results from at least three different processes or models (Bailey, 1964; Grimmett & Stirzaker, 2001): heterogeneity, contagion or cumulative disadvantage, and sampling.

Heterogeneity: the NBD is a mixture of Poisson distributions with gamma-distributed means. In greater detail, if an age group contains subgroups, and each subgroup has a Poisson distribution of eCCI, and the mean eCCIs of the different subgroups vary according to a gamma distribution, then the whole age group has an NBD of eCCI.

Testing this mechanism empirically requires stratifying each age group into subgroups by socially observable features (for example, sex, marital status, education, socioeconomic status, tobacco use, drug use, alcohol use, exercise, sleep, diet, and relationships), environmental exposures, genotypes, biomarkers and other characteristics. Then it must be shown whether each subgroup has a Poisson distribution of eCCI and whether the mean eCCIs of the different subgroups vary according to a gamma distribution. The required data are cross-sectional, not longitudinal.

Contagion: the more chronic conditions a person has, at any age, the higher the risk that the person will acquire an additional chronic condition. Statistical contagion in multimorbidity could be viewed as cumulative disadvantage. Testing this mechanism empirically requires longitudinal data.

Sampling: when the parameter r is a positive whole number, suppose that, over time, each person experiences a succession of independent random trials. Each trial can either succeed or fail. If the trial fails, the person's eCCI increases by one. If the trial succeeds, the person's eCCI remains unchanged. Once a person has r successes, the sequence of trials stops. The number x of failures up to that point is the person's observed eCCI.

The limited data we have available for this study are insufficient to select among these possible mechanisms. In addition, we cannot say to what extent the overdispersion observed in the distribution of eCCI in each age group is due to the differential weighting of chronic conditions with scores of 1, 2, 3 or 6 and to what extent the overdispersion is due to heterogeneity in the relative frequencies (estimated probabilities) of the different conditions. In future work, we could simulate record-level data under assumptions supported by the data and see where the observed data fall along the sampling distribution (as in Talih & Fricker, 2002). The multiple possibilities suggest questions for future research.

All FQHCs studied here serve low-income and minority communities, and all are subject to adverse effects of social determinants of health, resulting in accelerated biological aging (Charlson et al., 2024). Log-variance of eCCI increases less than linearly, i.e., concavely, with log-mean of eCCI in HS 1, HS 2, and HS 3. A reviewer suggested that this concave relation could profitably be compared with the log-log variance functions of populations with high socioeconomic status and more rectangular survival curves. We suggest that future work comparing the allostatic load (Seeman et al., 2001; McEwen, 2012) of different populations could reveal possible associations among life stressors, survival and the slope of the log-log variance function.

For possible future clinical interventions, to identify risk groups and opportunities for patient education and more intensive monitoring and intervention, it would be desirable to stratify each age group by characteristics such as sex, marital status, educational attainment, behavior (smoking, drug use, alcohol use, exercise, sleep, diet, relationships), socio-economic status, environmental exposures, and others. The features associated with individuals with high and low eCCI (relative to their age group) may offer clues to possible preventive interventions or behavioral changes.

The 39 conditions included in the eCCI generate $2^{39} \approx 5.5 \times 10^{11}$ possible combinations of conditions. Identifying each patient's combination of conditions would make it possible to examine the marginal frequency of each condition and the associations, positive and negative, among conditions. The UK Biobank, EPIC Cosmos, TriNetX and other biobanks and electronic health records (EHR) repositories may offer opportunities to extend this type of analysis.

Limitations of the present analysis include the cross-sectional nature of the data. We compare individuals of different ages who are observed only once. We are unable to disentangle the mixture of effects of age, aging, period, and cohorts. Longitudinal electronic health records containing medical history for the oldest age groups would make it possible to calculate the eCCI retrospectively for those individuals, say, at each age 10, 20, 30, … years earlier, and to compare the trajectories (mean, variance, and relation of variance to mean) over increasing age of individuals at an advanced age who have low eCCI with the trajectories of individuals who have high eCCI at an advanced age, allowing a more granular observation of the association between age and CCI variability among extreme phenotypes, such as younger individuals with high multimorbidity and older individuals with low multimorbidity. Longitudinal records of medical history would also make it possible to evaluate the influence of selective or differential survival on the distribution of eCCI among the living.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12062-026-09571-7.

Acknowledgements This collaboration among The Rockefeller University Center for Clinical and Translational Science, The Rockefeller University Laboratory of Populations, Clinical Directors Network (CDN), and Weill Cornell Medicine, is part of a cluster randomized clinical trial funded by the Patient-Centered Outcomes Research Institute (PCORI). J.E.C. thanks Roseanne Benjamin for help during this work and Dr. Ruth Kevess-Cohen M.D. for comments based on her geriatric clinical practice, as well as Mary E. Charlson MD, Martin T. Wells, PhD, James P. Hollenberg, MD, Rosio Ramos, BA, Robin Andrews, DSW, Shelly S. Sital, MPH, and Ilana Mittleman, MPH.

Author contributions J.E.C. conceived the study, analyzed the data, and drafted the original text, figures, and tables. T.J.L. and A.H. developed software to extract and check the data. A.C., A.E. and M.M. administered the project. J.N.T. acquired the funding and supervised and administered the project. All authors reviewed the manuscript.

Funding This collaboration among The Rockefeller University Center for Clinical and Translational Science, The Rockefeller University Laboratory of Populations, Clinical Directors Network (CDN), and Weill Cornell Medicine, is part of a cluster randomized clinical trial funded by the Patient-Centered Outcomes Research Institute (PCORI).
The trial was funded by PCORI Contract Number IHS-2017C3-8923 to CDN with additional infrastructure support provided by AHRQ (N2-PBRN Grant #1 P30-HS-021667 to CDN); Grant UL1 TR001866 from the National Center for Advancing Translational Sciences (NCATS), National Institutes of Health (NIH)

Clinical and Translational Science Award (CTSA) program to The Rockefeller University; INSIGHT (PCORI Award #R-1306-03961 to Weill Cornell Medicine); and CAPriCORN (PCORI Award #CDRN-1306-04737 to Northwestern Medicine). PCORI scientific staff played no role in study design and conduct. The statements presented in this publication are solely the responsibility of the authors and do not necessarily represent the views of the PCORI, its Board of Governors or Methodology Committee.

Data Availability The data are provided in full in the supplementary file HealthSystemAgeCCI20250323.csv.

Declarations

Competing interests The authors declare no competing interests.

References

American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. (2025). Clinical practice guidelines: 2025 AHA/ACC/AANP/AAPA/ABC/ACCP/ACPM/AGS/AMA/ASPC/NMA/PCNA/SGIM guideline for the prevention, detection, evaluation and management of high blood pressure in adults. *Circulation*, *152,* e114–e218. https://www.ahajournals.org/doi/10.1161/CIR.0000000000001356

Bailey, N. T. J. (1964). The elements of stochastic processes with applications to the natural sciences. Wiley, New York. ISBN: 978-0-471-52368-0.

Beard, J. R., Officer, A., de Carvalho, I. A., Sadana, R., Pot, A. M., Michel, J. P., Lloyd-Sherlock, P., Epping-Jordan, J. A. E., & G M E E (Geeske). Peeters, Wahyu Retno Mahanani, Jotheeswaran Amuthavalli Thiyagarajan, Somnath Chatterji. (2016). The World report on ageing and health: a policy framework for healthy ageing. *Lancet, 387,* 2145–2154. https://doi.org/10.1016/S0140-6736(15)00516-4

Benassi, F., Naccarato, Alessia, & Xu, M. (2023a). Daily Covid-19 infected population densities in Italian provinces follow Taylor's law. *Mathematical Population Studies*. https://doi.org/10.1080/08898480.2022.2155415

Benassi, F., Naccarato, A., & Salvati, L. (2023b). Testing Taylor's Law in urban population dynamics worldwide with simultaneous equation models. Economies, 11, 56 https://doi.org/10.3390/economies11020056

Bensken, W. P., Navale, S. M., McGrath, B. M. (2024). Variation in multimorbidity by sociodemographics and social drivers of health among patients seen at community-based health centers. *Journal of Multimorbidity and Comorbidity, 14,* 1–12. https://journals.sagepub.com/doi/10.1177/26335565241236410

Calif, R., & Schmitt, F. G. (2015). Taylor law in wind energy data. *Resources*, *4*, 787–795. https://doi.org/10.3390/resources4040787

Charlson, M. E. Carrozzino, Danilo, Guidi, Jenny, and Patierno, Chiara. (2022). Charlson Comorbidity Index: A critical review of clinimetric properties. *Psychotherapy and Psychosomatics, 91,* 8–35. https://doi.org/10.1159/000521288

Charlson, M. E., Mittleman, I., Ramos, R., Cassells, A., Lin, T. J., Eggleston, A., Wells, M. T., Hollenberg, J., Pirraglia, P., Winston, G., & Tobin, J. N. (2025). Preventing tipping points in high comorbidity patients: A lifeline from health coaches - rationale, design and methods. *Contemporary Clinical Trials, 152*:107865. https://doi.org/10.1016/j.cct.2025.107865

Charlson, M. E., Wells, M. T., Hollenberg, J., Ramos, R., Martinez, G. M., Gerard, M. J., Cassells, A., Lin, T. J., Mittleman, I., Eggleston, A., & Tobin, J. N. (2024). Examining individual- versus population-level social determinants of health in a cluster randomized trial of health coaches for patients with multiple chronic conditions. *J Clin Transl Sci*, *8*(1), e191. https://doi.org/10.1017/cts.2024.598

Charlson, M. E., & Wells, M. T. (2022). Comorbidity: from a confounder in longitudinal clinical research to the main issue in population management. *Psychotherapy and Psychosomatics*, *91*, 145–151. https://doi.org/10.1159/000521952

Charlson, M., Szatrowski, T. P., Peterson, J., & Gold Jeffrey. (1994). Validation of a combined comorbidity index. *Journal of Clinical Epidemiology, 47*(11), 1245–1251. https://doi.org/10.1016/0895-4356(94)90129-5

Chen, H. L., Yu, X. H., Yin, Y. H. (2023). Multimorbidity patterns and the association with health status of the oldest-old in long-term care facilities in China: a two-step analysis. *BMC Geriatrics 23*, 851. https://doi.org/10.1186/s12877-023-04507-8

Chen, Y., Shi, L., Zheng, X., Yang, J., Xue, Y., Xiao, S., Xue, B., Zhang, J., Li, X., Lin, H., et al. (2022). Patterns and determinants of multimorbidity in older adults: study in health-ecological perspective. *International Journal of Environmental Research and Public Health*, *19*(24), 16756. https://doi.org/10.3390/ijerph192416756

Chowdhury, S., Rahman, D. C., Das, T. C., Sunna, J., Beyene, A., & Hossain (2023). Global and regional prevalence of multimorbidity in the adult population in community settings: a systematic review and meta-analysis. *eClinicalMedicine*. https://doi.org/10.1016/j.eclinm.2023.101860

Cohen, J. E., Davis, R. A. & Samorodnitsky, G. (2022). COVID-19 cases and deaths in the United States follow Taylor ' s law for heavy-tailed distributions with infinite variance. *Proceedings of the National Academy of Sciences, 119*(38), e2209234119, September 12. https://doi.org/10.1073/pnas.2209234119

Cohen, J. E., Xu, M., & Brunborg, H. (2013b). Taylor's law applies to spatial variation in a human population. *Genus*, *69*(1), 25–60. http://www.jstor.org/stable/genus.69.1.25

Cohen, J. E., Xu, M., & Schuster, W. S. F. (2013a). Stochastic multiplicative population growth predicts and interprets Taylor's power law of fluctuation scaling. *Proceedings of the Royal Society Series B, 280*(1757), 20122955. https://doi.org/10.1098/rspb.2012.2955

Divo, M. J., Martinez, C. H., & Mannino, D. M. (2014). Ageing and the epidemiology of multimorbidity. *European Respiratory Journal*, *44*(4), 1055–1068. https://doi.org/10.1183/09031936.00059814

Eisler, Z., Bartos, I., & Kertész János. (2008). Fluctuation scaling in complex systems: Taylor's law and beyond. *Advances in Physics* 57(1):89–142. arXiv:0708.2053v2 [physics.soc-ph]. https://doi.org/10.1080/00018730801893043

Grimmett, G. R., & Stirzaker, D. R. (2001). Probability and random processes, 3rd Edition. Oxford University Press, New York.

Keeling, M. J., & Grenfell, B. T. (1999). Stochastic dynamics and a power law for measles variability. *Phil Trans R Soc Lond B*, *354*, 769–776. https://doi.org/10.1098/rstb.1999.0429

Keeling, M. J., Rohani, P., & Grenfell, B. T. (2001). Seasonally forced disease dynamics explained as switching between attractors. *Physica D: Nonlinear Phenomena*, *148*, 317–335. https://doi.org/10.1016/S0167-2789(00)00187-1

Ly, K., Wakefield, D., & ZuWallack, R. (2025). The usefulness of Charlson Comorbidity Index (CCI) scoring in predicting all-cause mortality in outpatients with clinical diagnoses of COPD. *Journal of Multimorbidity and Comorbidity*, *15*, 1–7. https://doi.org/10.1177/26335565251315876

Martin, F. C., Quinn, T. J., Straus, S. E., Anand, S., van der Velde, N., & Harwood, R. H. (2024). New horizons in clinical practice guidelines for use with older people. *Age and Ageing, 53*(7):afae158. https://doi.org/10.1093/ageing/afae158

MATLAB Version (2024b). 24.2.0.2740171 Update 1, Natick, Massachusetts: The MathWorks Inc., 2024.

McEwen, B. S. (2012). Brain on stress: how the social environment gets under the skin. *Proc Natl Acad Sci U S A, 109*(Suppl 2):17180-5. Erratum: Proc Natl Acad Sci U S A. 2013; 110(4):1561. https://doi.org/10.1073/pnas.1121254109

Seeman, T. E., McEwen, B. S., Rowe, J. W., & Singer, B. H. (2001). Allostatic load as a marker of cumulative biological risk: MacArthur studies of successful aging. *Proc Natl Acad Sci U S A*, *98*(8), 4770–4775. https://doi.org/10.1073/pnas.081072698

Talih, M., & Fricker, R. D. (2002). Effects of neighbourhood demographic shifts on findings of environmental injustice: a New York City case-study. *J R Statist Soc A*, *165*(2), 375–397.

Taylor, L. R. (1961). Aggregation, variance and the mean. *Nature*, *189*(4766), 732. https://doi.org/10.1038/189732a0

Taylor, L. R., Woiwod, I. P., & Perry, J. N. (1978). The density-dependence of spatial behaviour and the rarity of randomness. *Journal of Animal Ecology*, *47*(2), 383. https://doi.org/10.2307/3790. 406, Junehttp://www.jstor.org/stable/3790

Taylor, R. A. J. (2019). *Taylor's power law: order and pattern in nature*. Academic. https://doi.org/10.1016/C2016-0-00333-9

Tippett, M. K., & Cohen, J. E. (2016). Tornado outbreak variability follows Taylor's power law of fluctuation scaling and increases dramatically with severity. *Nature Communications*, *7*, 10668. https://doi.org/10.1038/ncomms10668

Xu, M., & Cohen, J. E. (2021). Spatial and temporal autocorrelations affect Taylor s law for US county populations: descriptive and predictive models. *Plos One*, *16*(1), e0245062. https://doi.org/10.1371/journal.pone.0245062

Authors and Affiliations

Joel E. Cohen[1,2,3] · T. J. Lin[4] · Alice Eggleston[5] · Mirta Milanes[5] · Amro Hassan[5] · Andrea Cassells[4] · Jonathan N. Tobin[4,6]

✉ Joel E. Cohen
cohen@rockefeller.edu

✉ Jonathan N. Tobin
jntobin@cdnetwork.org

T. J. Lin
tlin@cdnetwork.org

Alice Eggleston
aeggleston@alliancechicago.org

Mirta Milanes
mmilanes@alliancechicago.org

Amro Hassan
ahassan@alliancechicago.org

Andrea Cassells
acass@cdnetwork.org

1 Laboratory of Populations, Rockefeller University, New York, USA

2 Department of Statistics, University of Chicago, Chicago, IL, USA

3 Department of Statistics, Columbia University, New York, USA

4 Clinical Directors Network (CDN), New York, USA

5 Alliance Chicago, Chicago, USA

6 Center for Clinical and Translational Science, The Rockefeller University, New York, USA