MHJ

LUMORA JOURNALS

Multidisciplinary Healthcare Journal

Vol. 1, No. 1 (2025) · e-ISSN 3152-8775

Original research · DOI: not registered

Selected diagnosed conditions and self-rated health among United States adults: a cross-sectional survey analysis of NHANES 2017–2018

Aliya Durrani[1],*, Aliya Jasmine[2], Alpana Mishra[3], Alpo Vuorio[4]

Submission date	Acceptance date	Publication date
6 January 2025	20 January 2025	30 January 2025

[1] Peshawar Medical College, Peshawar, Pakistan

[2] Sri Ramachandra Institute of Higher Education and Research, Chennai, India

[3] Kalinga Institute of Medical Sciences, Bhubaneswar, Odisha, India

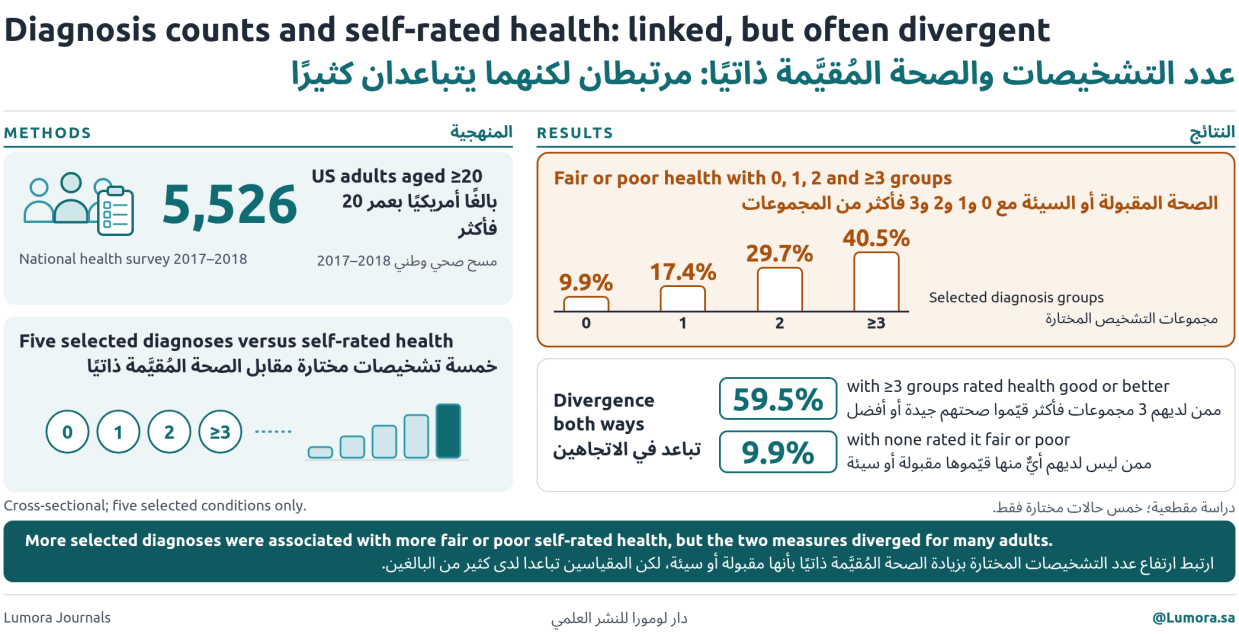
[4] University of Helsinki, Helsinki, Finland

* Correspondence: fatemalumora@gmail.com

English edition. An Arabic edition of this article is available: link to be added at publication.

Cite as: Durrani A, Jasmine A, Mishra A, Vuorio A. Selected diagnosed conditions and self-rated health among United States adults: a cross-sectional survey analysis of NHANES 2017–2018. Multidisciplinary Healthcare Journal. 2025;1(1):MHJ-2026-000321. © The authors. Published open access under the Creative Commons Attribution 4.0 International licence (CC BY 4.0), https://creativecommons.org/licenses/by/4.0/.

Graphical abstract



Graphical abstract: More selected diagnoses were associated with more fair or poor self-rated health, but the two measures diverged for many adults. Designed by the Lumora Journals production team using hand-coded HTML/SVG rendered in Google Chrome.

Abstract

Background: Diagnosis counts and general health ratings are often assumed to move together. We quantified how often they diverge, in both directions, using a restricted five-condition list that is not a comprehensive multimorbidity measure.

Methods: This exploratory secondary analysis included 5526 adults aged at least 20 with valid household-interview responses in NHANES 2017–2018. We counted ever-diagnosed hypertension, diabetes, asthma, arthritis and cancer/malignancy; retained five health-response levels; and estimated weighted proportions with stratified-cluster variance. Logistic models with 4747 complete cases gave odds ratios and covariate-standardized prevalences. Sensitivity analyses addressed borderline diabetes, current asthma, the component list, missing income, non-melanoma skin cancer and the degrees of freedom.

Results: In the highest count group, 59.5% reported good or better health; with no selected diagnoses, 9.9% reported fair or poor health. Fair or poor health was 9.9%, 17.4%, 29.7%, 40.5% across counts 0, 1, 2 and at least 3 (covariate-standardized 8.5% to 42.6%; prevalence ratio for at least 3 versus 0, 5.00 (3.36–7.45)). Adjusted odds ratios versus 0 were 2.42 (1.72–3.40), 5.78 (3.50–9.55), 10.20 (5.50–18.94), exceeding unadjusted values mainly because age was strongly related to the count but not to fair or poor health at a given count. Missing-income, cancer-item and coding sensitivities gave similar estimates.

Conclusions: Higher restricted diagnosis counts were associated with less favorable self-rated health, but the two measures diverged for many adults in both directions. These cross-sectional results support reporting the joint distribution; they do not establish causality, comprehensive multimorbidity, prognosis or an individual decision rule.

Keywords: self-rated health; multimorbidity; chronic disease; health surveys; NHANES; cross-sectional studies

Introduction

Self-rated health and a count of diagnosed conditions describe related but different aspects of health. A person may regard a treated chronic disease as compatible with good health, whereas another person with few recorded diagnoses may experience pain, fatigue, functional difficulty, distress or an undiagnosed disorder. Neither measure is a reference standard for the other; their joint distribution shows where a diagnosis-based summary and the respondent's own appraisal do not coincide.

General health ratings are associated with subsequent mortality across observational studies [1,2]. Jylhä proposed that a rating combines bodily information with the context in which a respondent interprets it [3], so an identical number of diagnoses need not have an identical meaning to different people. Survey data do not measure that reasoning directly, and explanations based on adaptation or symptom burden remain hypotheses.

The meaning of a condition count depends on which conditions are counted. Prevalence studies differ in their diagnosis lists and operational definitions [4], and measures of morbidity burden range from simple counts to weighted indices chosen according to purpose and available data [5,6]. A five-condition count is therefore not a comprehensive measure of chronic disease, and the same count can represent different combinations, severities and durations of illness.

The association between the number of conditions and self-rated health is well established. In 25,268 adults aged 39–79 years registered in English general practice, the odds of moderate or poor health rose with the number of chronic conditions (odds ratio 3.4 for three or more), and psychosocial measures did not explain the association [7]. Among Medicare recipients three years after hip replacement, self-rated health was associated with functional, mental-health and other health domains as well as medical comorbidity [8]. In United States national surveys, the odds of worse self-rated health increased with the number of co-occurring mental, substance-use and chronic medical conditions [9], and some combinations of treated conditions carried about sevenfold adjusted odds of poor health [10]. NHANES 2017–2018 itself has been used to examine thirteen chronic conditions and social determinants of multimorbidity [11].

NHANES, the National Health and Nutrition Examination Survey, provides a probability-based source of health information on the United States civilian noninstitutionalized population [12]. Studies of the association itself are reviewed above; what is less well described is how often the two measures disagree. This paper therefore focuses on discordance in a nationally representative sample. It uses a clearly delimited interview-only list of five condition groups, preserves all five health-response levels, reports discordance in both directions, and shows how alternative coding, missing covariates and the choice of effect measure change the estimates.

The primary question is how self-rated health varies among adults reporting zero, one, two, or at least three of five selected condition groups: ever-diagnosed hypertension, diabetes, asthma, arthritis and cancer or malignancy. Secondary questions concern fair or poor health among adults without these diagnoses, good or better health among adults with several, and the sensitivity of the pattern to coding and selection. The adjusted analysis is descriptive; it does not identify a causal effect or support individual clinical decisions.

Methods

Design and reporting

We conducted an exploratory cross-sectional secondary analysis of the public NHANES 2017–2018 household-interview files. The analysis plan was not registered or deposited before analysis; all analyses are reported, and additional analyses are identified as such. The study was not registered prospectively, and no confirmatory hypotheses are claimed. Two other manuscripts by the same authors use this survey cycle, one on examination-defined undiagnosed hypertension and one on sleep timing; they address different outcomes and measures and are disclosed to the editor.

Reporting follows the cross-sectional items of STROBE, Strengthening the Reporting of Observational Studies in Epidemiology [13,14], and the completed checklist is provided as a supplementary file. The analysis follows NCHS, National Center for Health Statistics, guidance on the quality of NHANES analyses [15].

Participants and interview measures

Five public files (demographics; hospital utilization and access to care; medical conditions; diabetes; and blood pressure and cholesterol) were linked by the public respondent sequence number [16–20]. Adults aged at least 20 years were eligible because the arthritis and cancer questions begin at age 20 [18]. Inclusion required a positive household-interview weight, a valid general-health response and a valid response for each of the five condition groups; no examination or laboratory item was required. Figure 1 shows the participant flow.

The general-health question offers excellent, very good, good, fair or poor [17]. We retained the five-level distribution; the primary binary outcome combined fair and poor, and excellent or very good was described as a narrower favorable category. Refused and don't-know responses were not reassigned. This coding is a communication choice, not a validated clinical threshold.

Hypertension was a report that a doctor or other health professional had ever said the respondent had high blood pressure; it reflects awareness of a past diagnosis, not examination-defined or uncontrolled hypertension [20]. A yes response to the diabetes question counted as diagnosed diabetes, whereas no and borderline responses did not [19]; a sensitivity analysis excluded borderline responses.

Asthma, arthritis and cancer or malignancy were ever-diagnosed items in the medical-conditions file [18]. Asthma could refer to a past condition, arthritis was not subtyped, and cancer was not restricted to active disease or a site. The count summed the five groups, with three to five combined as at least three; this grouping avoids sparse categories and does not imply equivalent clinical burden.

The count excludes many relevant conditions, including other cardiovascular diagnoses, mental-health disorders, chronic kidney disease, undiagnosed conditions and functional impairment, and applies no severity weights. A zero count means none of the five groups, not absence of illness, and a count of two or more is not comprehensive multimorbidity.

Covariates and missingness

The adjusted model included age (linear per ten years; ages of 80 or older are released as 80), sex (two public categories), the five released race and Hispanic-origin categories (reference non-Hispanic White),

education and family income-to-poverty ratio [16]. Race and Hispanic origin were survey descriptors, not biological causes; the residual other-race category combines heterogeneous groups.

Education was grouped as below high school, high school and above high school; income-to-poverty ratio was continuous, using the released value and its upper limit. Adjustment was descriptive and does not establish exchangeability or remove all confounding. Flexible terms, interactions and alternative combinations were not selected after inspecting significance.

Participants with valid outcome and condition answers contributed to the descriptive domain even if a covariate was missing; the adjusted model used complete cases. Missing education and income overlap. Survey weighting does not correct item-level missingness [21], so the fair-or-poor prevalence of retained and excluded participants is reported and missing income was examined in sensitivity analyses.

Survey estimation and association model

Household-interview weights were used because every variable came from the interview or demographics components [16–21]. The design retained all participants with a positive interview weight, with eligibility, valid responses and complete-case status defined as domains, so that domain variances reflect the original primary sampling units [22]. The masked stratum and primary-sampling-unit identifiers were used, with units nested within strata.

Weighted percentages were ratios of weighted totals, with Taylor-linearized standard errors. The design had thirty primary sampling units in fifteen strata, giving fifteen design degrees of freedom. Confidence intervals for proportions were logit-transformed and used a t critical value with the design degrees of freedom; they describe sampling uncertainty only.

Survey-weighted logistic regression modeled fair or poor health with zero selected groups as the reference: an unadjusted model and a covariate-adjusted complete-case model. Odds ratios used design-based sandwich standard errors and t-based intervals with fifteen design degrees of freedom. Because the adjusted model estimates twelve coefficients besides the intercept, intervals were also calculated with the residual degrees of freedom (15 − 12 = 3).

An odds ratio is not a prevalence ratio, and the two diverge when the outcome is common [23]. We therefore also report covariate-standardized prevalences and prevalence ratios alongside the observed weighted percentages.

R survey procedures were used for the primary design-based estimates [24]. A separate implementation in Python 3.10 or later (numpy, pandas, scipy) reconstructed the weighted ratios, linearization and logistic sandwich covariance from the public files; code reproducing every reported estimate is provided as Supplementary Code S1. Findings were not selected on hypothesis-test thresholds, the analyses are exploratory, and intervals are pointwise.

Sensitivity analyses and estimate reliability

Sensitivity analyses excluded borderline diabetes; replaced ever-diagnosed asthma with current asthma, requiring a valid current answer among those with a lifetime diagnosis; and omitted each component in turn from the at-least-two-group summary to show dependence on the list.

Four additional analyses were performed. First, the unadjusted model was refitted in the complete-case domain and covariates were added stepwise, to explain the difference between unadjusted and adjusted

odds ratios. Second, covariate-standardized prevalences of fair or poor health were obtained by assigning every complete-case respondent to each count group, averaging the weighted adjusted-model predictions, and forming prevalence ratios and differences versus zero groups; their intervals used a delete-one-primary-sampling-unit jackknife (30 replicates). Third, missing income-to-poverty ratio was handled among 5516 respondents with valid education by a missing-income category (income below 1, 1 to below 2, 2 to below 4, and 4 or more) and by multiple imputation (predictive mean matching with five donors, 20 imputations; imputation model including the count, age, age squared, sex, race and Hispanic origin, education, the outcome and the log weight; Rubin's rules with Barnard–Rubin degrees of freedom). Fourth, respondents whose only reported cancer type was non-melanoma skin cancer (codebook code 32) were excluded, alone or together with those reporting only skin cancer of unknown type (code 33) [18].

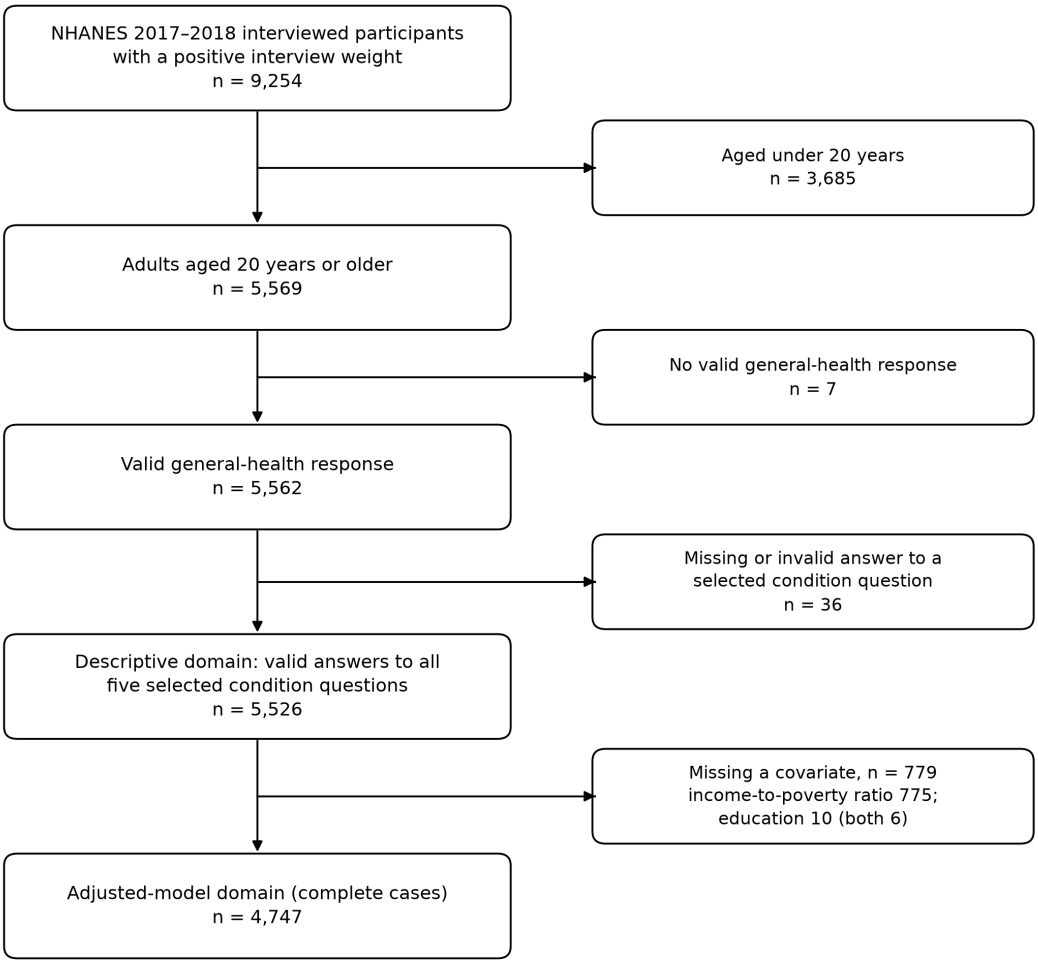
Reliability was judged using sample size, design degrees of freedom, effective sample size and interval width, following NCHS guidance [25,26]; cells with fewer than thirty events are flagged.

Results

Participants and selected diagnoses

Of 9,254 interviewed participants with positive interview weights, 5,569 were aged at least 20 years; 5,562 gave a valid general-health response and 5,526 gave valid answers to all five condition questions, forming the descriptive domain (Figure 1). The adjusted model included 4,747 participants; 779 lacked at least one covariate (income-to-poverty ratio 775, education ten, both six). The descriptive domain contained 1,404 fair-or-poor responses, a weighted prevalence of 18.7% with a 95% confidence interval (CI) of 16.7–20.9%.

Figure 1. Participant flow, NHANES 2017–2018



Counts are unweighted respondents. Covariate exclusions overlap: six respondents lacked both income-to-poverty ratio and education.

The zero-, one-, two- and at-least-three-group categories contained 2,163, 1,576, 1,077 and 710 respondents (weighted 43.2%, 29.0%, 17.7% and 10.1%). Mean released age rose from 39.7 to 65.0 years across the groups, and the highest group had a larger female share (Table 1).

Table 1. Participant characteristics by selected diagnosed-condition count, NHANES 2017–2018

Characteristic	0 groups	1 group	2 groups	3 or more groups
Respondents, n	2163	1576	1077	710
Weighted group share, %	43.2 (39.7–46.8)	29.0 (26.2–32.0)	17.7 (15.8–19.7)	10.1 (8.9–11.5)
Mean released age, years	39.7	48.7	59.4	65.0
Female, %	51.5 (48.1–54.8)	49.6 (45.1–54.1)	52.2 (48.5–55.8)	58.6 (52.4–64.5)
Below high school, %	11.5 (8.9–14.6)	9.3 (7.7–11.2)	12.8 (9.5–17.1)	13.2 (10.7–16.2)
Income-to-poverty ratio below 1, %	12.1 (10.2–14.3)	13.6 (11.2–16.4)	11.6 (8.8–15.2)	12.4 (8.0–18.6)
Valid education responses, n	2159	1574	1076	707

Characteristic	0 groups	1 group	2 groups	3 or more groups
Valid income responses, n	1848	1374	923	606

n, unweighted number of respondents; CI, confidence interval. Percentages use household-interview weights and Taylor-linearized clustered/stratified variance; logit intervals use 15 design degrees of freedom.

Selected groups: ever-diagnosed hypertension, diabetes, asthma, arthritis and cancer or malignancy. This list is not a comprehensive multimorbidity measure. Borderline diabetes does not count as diagnosed diabetes.

NHANES, National Health and Nutrition Examination Survey. Percentage entries show estimate (95% CI); means are in years, not percentages. Age 80 represents 80 years or older. Education and income rows use their available-response denominators, shown separately.

Weighted component prevalences were 32.4%, 11.4%, 15.0%, 28.0% and 11.2% for hypertension, diabetes, ever-asthma, arthritis and ever-cancer. Components overlap within individuals, and fair or poor health among respondents reporting each ranged from 27.0% to 38.5% (Table 2).

Table 2. Selected diagnosis-group frequencies and fair or poor health among respondents reporting each group

Ever-diagnosed group	Respondents with group, n	Weighted prevalence, % (95% CI)	Fair/poor health among group, % (95% CI)
Hypertension	2104	32.4 (29.5–35.3)	30.9 (26.6–35.6)
Diabetes	866	11.4 (10.3–12.7)	38.5 (33.1–44.3)
Asthma	829	15.0 (13.3–16.7)	27.5 (21.9–33.9)
Arthritis	1687	28.0 (25.0–31.2)	27.5 (22.7–32.8)
Cancer	582	11.2 (9.7–12.7)	27.0 (20.2–35.1)

n, unweighted number of respondents; CI, confidence interval. Estimation as described for Table 1.

Selected groups as defined in Table 1.

The diagnosis groups overlap and their prevalences must not be summed as a population share. Ever-cancer is not restricted to active malignancy. Component-specific health prevalences are not comparisons adjusted for severity or other diagnoses.

Health distribution and conditional association

Fair or poor health increased across the four count categories: 9.9% (95% CI 8.3–11.9%), 17.4% (15.4–19.6%), 29.7% (25.6–34.1%) and 40.5% (33.2–48.3%) (Table 3; Figures 2 and 3). The sequence is cross-sectional and does not establish a dose-response mechanism or an individual trajectory.

Table 3. Five-level self-rated health within selected diagnosed-condition count categories

Condition groups	Health response	Respondents, n	Weighted %	95% CI, %
0	Excellent	398	21.5	17.4–26.2
0	Very good	709	36.3	33.0–39.7
0	Good	761	32.3	28.5–36.3
0	Fair	268	9.1	7.7–10.8
0	Poor	27	0.8	0.4–1.4

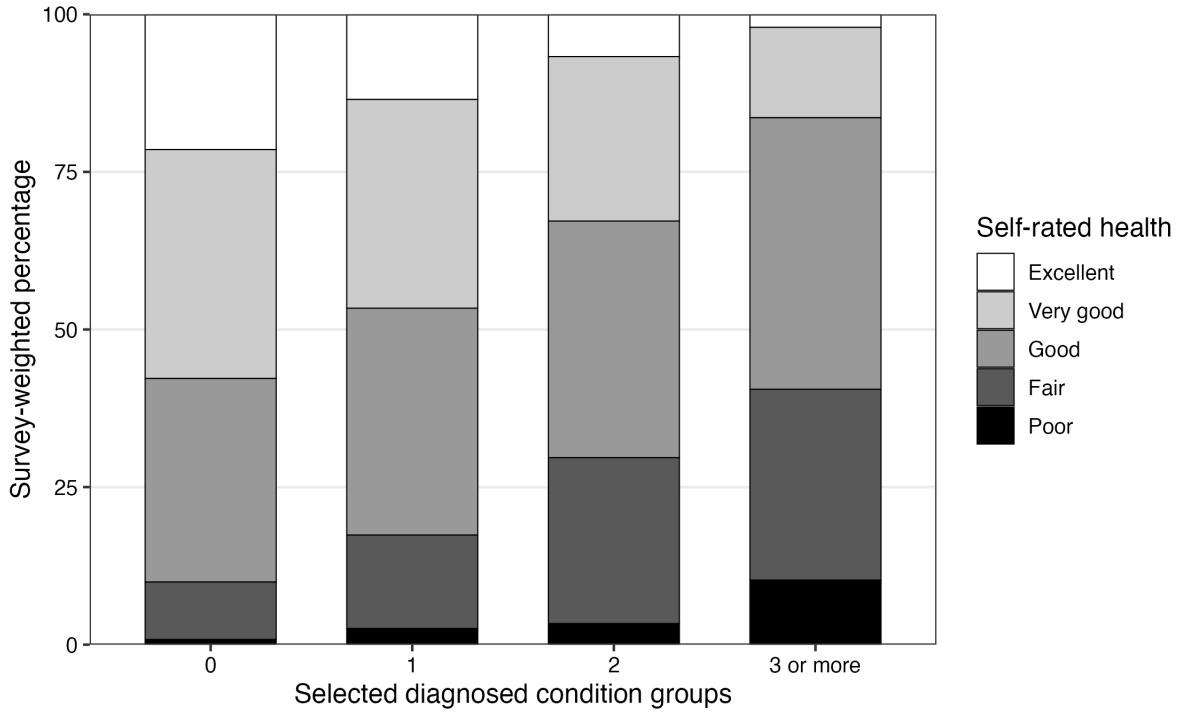
Condition groups	Health response	Respondents, n	Weighted %	95% CI, %
1	Excellent	159	13.5	9.8–18.3
1	Very good	451	33.1	30.1–36.3
1	Good	602	36.0	32.6–39.6
1	Fair	309	14.8	12.6–17.3
1	Poor	55	2.6	1.6–4.1
2	Excellent	57	6.7	4.4–10.0
2	Very good	227	26.1	22.1–30.5
2	Good	391	37.5	33.4–41.8
2	Fair	338	26.3	22.6–30.4
2	Poor	64	3.3	2.5–4.5
3 or more	Excellent	11	2.1	0.8–5.1
3 or more	Very good	70	14.3	11.2–18.1
3 or more	Good	286	43.1	37.5–48.8
3 or more	Fair	252	30.3	23.7–37.7
3 or more	Poor	91	10.2	7.1–14.5

n, unweighted number of respondents; CI, confidence interval. Estimation as described for Table 1.

Selected groups as defined in Table 1.

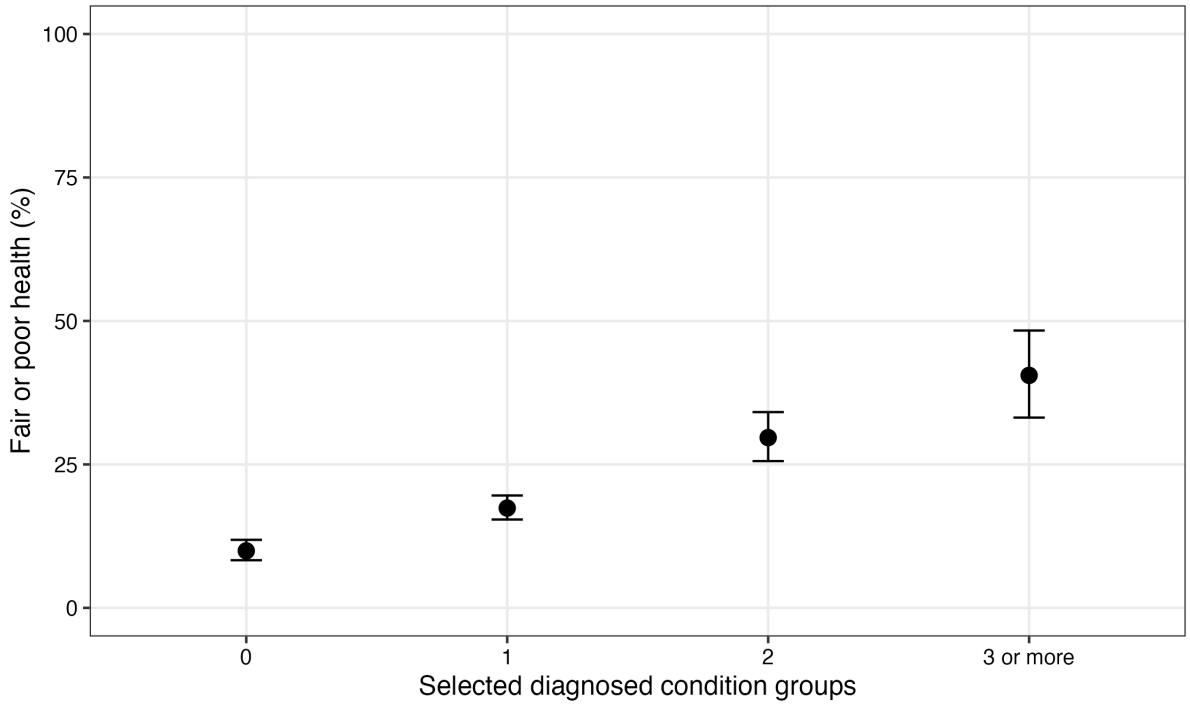
Within each count category, the five weighted percentages sum to 100 before rounding. The n column counts respondents in that response cell. Cells with fewer than 30 responses are zero-group poor health (27) and at-least-three-group excellent health (11); interpret these narrow cells cautiously.

Figure 2. Distribution of five-level self-rated health by selected diagnosed-condition count



Within-group survey-weighted percentages. NHANES, National Health and Nutrition Examination Survey. The count includes five ever-diagnosed groups, not all illness. No individual transition is shown.

Figure 3. Fair or poor self-rated health and design-based uncertainty by selected diagnosed-condition count



Points are weighted percentages; error bars are 95% confidence intervals from Taylor-linearized variance and a logit transformation with 15 design degrees of freedom. No fitted individual risk curve is shown.

Discordance was present in both directions. Among adults reporting none of the selected diagnoses, 9.9% (95% CI 8.3–11.9%) reported fair or poor health. Among those reporting at least three groups, 59.5% (51.7–66.8%) reported good or better health and 16.4% (12.6–21.1%) excellent or very good health. Two cells were small: 27 zero-group respondents reported poor health and eleven at-least-three-group respondents reported excellent health.

The unadjusted odds ratios for one, two and at least three groups versus none were 1.91 (1.49–2.44), 3.82 (2.87–5.10) and 6.17 (4.29–8.87); adjusted odds ratios were 2.42 (1.72–3.40), 5.78 (3.50–9.55) and 10.20 (5.50–18.94) (Table 4; Figure 4). The model cannot account for severity, unlisted conditions, access to care or how respondents interpret general health.

The adjusted odds ratios exceeded the unadjusted ones mainly because of age. In the complete-case domain the unadjusted odds ratios were 1.93 (1.51–2.47), 4.22 (3.01–5.92) and 6.76 (4.47–10.23); adding age alone gave 2.11, 5.18 and 8.84, and adding sex and race and Hispanic origin gave 2.23, 5.68 and 9.88. Age was strongly related to the count (at least three groups: 1.2%, 6.9% and 24.6% of adults aged 20–39, 40–59 and 60 years or older in the complete-case sample, n = 4,747) but not to fair or poor health at a given count (adjusted odds ratio per ten years 0.94, 95% CI 0.84–1.05), so adjustment removed negative confounding by age. Covariate-standardized prevalences of fair or poor health were 8.5% (6.3–10.7%), 17.3% (15.0–19.5%), 31.1% (25.6–36.6%) and 42.6% (32.8–52.4%), giving prevalence ratios versus zero groups of 2.03 (1.54–2.68), 3.66 (2.54–5.28) and 5.00 (3.36–7.45) and prevalence differences of 8.8, 22.6 and 34.1 percentage points (Table 4).

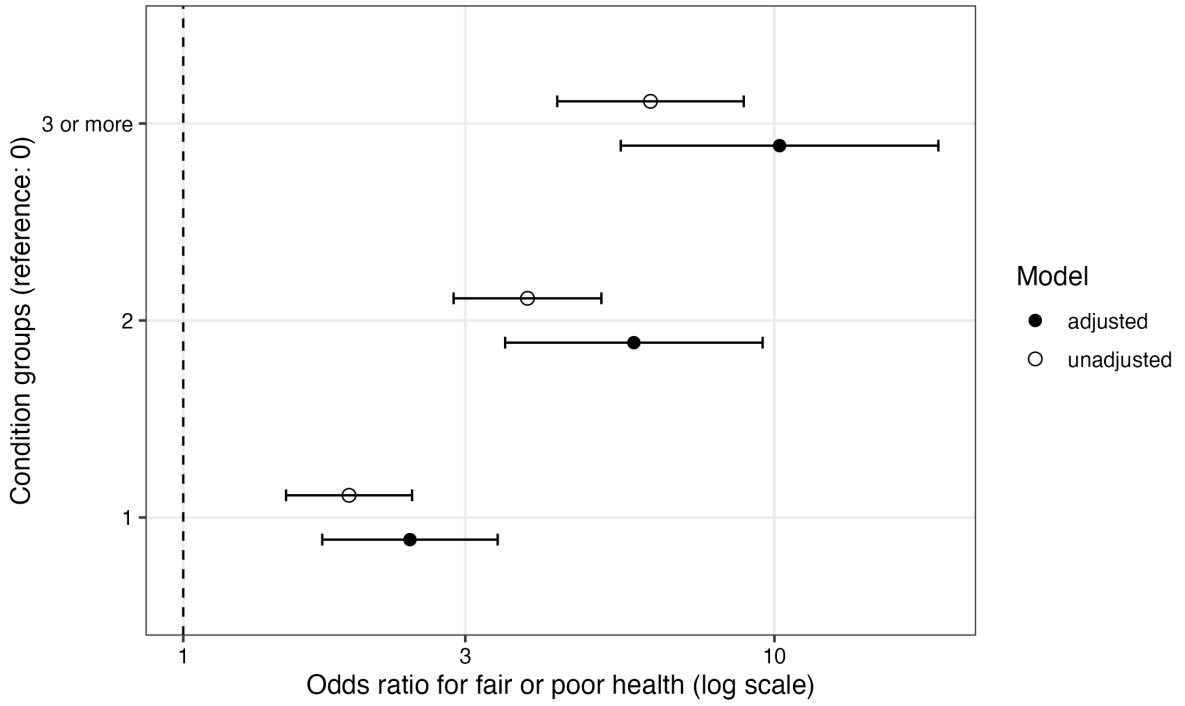
Table 4. Survey-weighted odds ratios for fair or poor self-rated health by selected diagnosed-condition count

Count contrast versus 0	Unadjusted OR (95% CI)	Adjusted OR (95% CI)	Standardized fair/poor, % (95% CI)	Standardized PR (95% CI)	Complete-case respondents, n
0 (reference)	1.00 (reference)	1.00 (reference)	8.5 (6.3–10.7)	1.00 (reference)	1846
1 versus 0	1.91 (1.49–2.44)	2.42 (1.72–3.40)	17.3 (15.0–19.5)	2.03 (1.54–2.68)	1374
2 versus 0	3.82 (2.87–5.10)	5.78 (3.50–9.55)	31.1 (25.6–36.6)	3.66 (2.54–5.28)	922
3 or more versus 0	6.17 (4.29–8.87)	10.20 (5.50–18.94)	42.6 (32.8–52.4)	5.00 (3.36–7.45)	605

OR, odds ratio; CI, confidence interval; PR, prevalence ratio; n, unweighted number of respondents. Reference: 0 selected groups. Standardized prevalences average adjusted-model predictions over the complete-case population with every respondent assigned to each count; their intervals use a delete-one-primary-sampling-unit jackknife (30 replicates).

Unadjusted model: 5526 respondents. Adjusted model: 4747 complete cases; age per 10 years, sex, five released race/Hispanic-origin categories, three education categories and income-to-poverty ratio. Taylor-linearized sandwich standard errors; OR intervals use 15 design degrees of freedom.

Figure 4. Unadjusted and adjusted survey-weighted associations with fair or poor self-rated health



OR, odds ratio; CI, confidence interval. Horizontal error bars are 95% CIs on a logarithmic axis; models as in Table 4. An OR is not a prevalence ratio.

Sensitivity and selection

Excluding 169 borderline-diabetes respondents gave fair-or-poor percentages of 9.7%, 17.3%, 30.1% and 40.5%, and using current asthma gave 10.0%, 18.4%, 29.8% and 42.3% (Table 5).

Table 5. Fair or poor health under alternative coding and complete-case selection

Analysis domain	Selected count	Respondents, n	Weighted % (95% CI)
Exclude borderline diabetes	0	2118	9.7 (7.9–11.8)
Exclude borderline diabetes	1	1517	17.3 (15.4–19.4)
Exclude borderline diabetes	2	1030	30.1 (26.1–34.3)
Exclude borderline diabetes	3 or more	692	40.5 (33.3–48.2)
Current asthma definition	0	2294	10.0 (8.3–12.0)
Current asthma definition	1	1521	18.4 (15.4–21.9)
Current asthma definition	2	1047	29.8 (25.9–34.1)
Current asthma definition	3 or more	646	42.3 (33.8–51.3)
Complete-case covariates	0	1846	9.2 (7.4–11.4)
Complete-case covariates	1	1374	16.4 (14.3–18.8)
Complete-case covariates	2	922	30.0 (25.7–34.7)
Complete-case covariates	3 or more	605	40.7 (33.0–48.9)

n, unweighted number of respondents; CI, confidence interval. Estimation as described for Table 1.

Selected groups as defined in Table 1.

Borderline sensitivity excludes 169 primary-domain respondents. Current-asthma sensitivity requires a valid current response for those with ever-asthma, while valid never-asthma is negative; the other four definitions are retained. Complete-case rows describe the adjusted-model population without covariate standardization.

The proportion with at least two selected groups ranged from 14.7% to 24.5% across the five leave-one-component calculations, and fair or poor health in that group from 32.2% to 38.5% (Table 6). Estimates therefore depend on the condition list and should not be compared directly with studies counting different conditions.

Table 6. Sensitivity of restricted multiple-condition summaries to omission of one diagnosis group

Omitted component	At least 2 remaining groups, % (95% CI)	Respondents with at least 2, n	Fair/poor among at least 2, % (95% CI)
Hypertension	14.7 (13.2–16.5)	980	36.1 (29.0–43.9)
Diabetes	24.0 (21.5–26.6)	1499	32.2 (27.0–37.9)
Asthma	23.6 (21.5–25.9)	1572	34.3 (29.7–39.1)
Arthritis	16.8 (15.1–18.8)	1115	38.5 (32.6–44.9)
Cancer	24.5 (22.0–27.2)	1615	35.2 (30.0–40.8)

n, unweighted number of respondents; CI, confidence interval. Estimation as described for Table 1.

Selected groups as defined in Table 1.

Each row omits the named group and counts the other four. The at-least-two-group prevalence uses the 5526 primary-domain respondents; the health prevalence uses the restricted multiple-group subgroup. No component was selected for omission on the basis of statistical significance.

Fair or poor health was 18.1% (16.1–20.4%) in the complete-case domain and 22.8% (18.0–28.5%) among the 779 participants excluded for missing covariates.

Table 7 collects the stepwise models and the additional sensitivity analyses. Handling missing income differently changed the adjusted odds ratios little. With a missing-income category (5516 respondents) they were 2.42 (1.69–3.45), 5.25 (3.30–8.37) and 9.32 (5.04–17.24); with multiple imputation of the 769 missing income values they were 2.43 (1.70–3.47), 5.29 (3.31–8.44) and 9.32 (5.10–17.04). Of the 582 respondents reporting cancer, 63 reported only non-melanoma skin cancer; excluding them gave fair-or-poor percentages of 9.9%, 17.4%, 30.3% and 41.1% and adjusted odds ratios of 2.39 (1.69–3.37), 5.87 (3.51–9.80) and 10.28 (5.61–18.85). Also excluding the 32 respondents whose only cancer was skin cancer of unknown type gave similar results. With the residual degrees of freedom (3), the adjusted intervals widened to 1.45–4.03, 2.74–12.23 and 4.05–25.69.

Table 7. Additional sensitivity analyses: odds ratios for fair or poor self-rated health by selected diagnosed-condition count

Analysis	Respondents, n	1 versus 0, OR (95% CI)	2 versus 0, OR (95% CI)	3 or more versus 0, OR (95% CI)
Unadjusted, descriptive domain (primary)	5526	1.91 (1.49–2.44)	3.82 (2.87–5.10)	6.17 (4.29–8.87)
Unadjusted, complete-case domain	4747	1.93 (1.51–2.47)	4.22 (3.01–5.92)	6.76 (4.47–10.23)

Analysis	Respondents, n	1 versus 0, OR (95% CI)	2 versus 0, OR (95% CI)	3 or more versus 0, OR (95% CI)
Age only, complete-case domain	4747	2.11 (1.57–2.84)	5.18 (3.24–8.29)	8.84 (5.13–15.24)
Age, sex, race and Hispanic origin	4747	2.23 (1.63–3.05)	5.68 (3.49–9.24)	9.88 (5.59–17.49)
Fully adjusted (primary)	4747	2.42 (1.72–3.40)	5.78 (3.50–9.55)	10.20 (5.50–18.94)
Fully adjusted, residual degrees of freedom (3)	4747	2.42 (1.45–4.03)	5.78 (2.74–12.23)	10.20 (4.05–25.69)
Missing-income category	5516	2.42 (1.69–3.45)	5.25 (3.30–8.37)	9.32 (5.04–17.24)
Multiple imputation of income (M = 20)	5516	2.43 (1.70–3.47)	5.29 (3.31–8.44)	9.32 (5.10–17.04)
Excluding non-melanoma-skin-cancer-only respondents	4693	2.39 (1.69–3.37)	5.87 (3.51–9.80)	10.28 (5.61–18.85)
Also excluding unknown-type-skin-cancer-only respondents	4667	2.39 (1.70–3.36)	5.93 (3.55–9.92)	9.96 (5.37–18.48)

OR, odds ratio; CI, confidence interval. Reference: 0 selected groups. Adjusted models include age per 10 years, sex, race and Hispanic origin, education and income-to-poverty ratio (categorical with a missing category in the missing-income row). Intervals use 15 design degrees of freedom except where stated; multiple-imputation intervals use Barnard–Rubin degrees of freedom. The cancer-item rows are fully adjusted complete-case models.

Discussion

Principal findings and contribution

Higher counts of five selected ever-diagnosed condition groups were associated with more fair or poor general health, yet the two measures diverged for many respondents: most adults with at least three groups rated their health good or better, and about one in ten with none rated it fair or poor. The joint distribution shows both an overall gradient and its imperfect correspondence for individuals; neither measure substitutes for the other.

The gradient agrees with studies in English general practice, Medicare recipients and United States national surveys [7–10] and with a broader thirteen-condition analysis of the same NHANES cycle [11]. The contribution of the present analysis is the discordance and the measurement choices behind it: a delimited interview-only list, the full five-level response distribution, standardized prevalences alongside odds ratios, and sensitivity to coding, missing covariates and the cancer item. The estimates are restricted-condition estimates, not a more complete national multimorbidity burden.

The pattern is compatible with the literature linking perceived health to morbidity and outcomes [1–3], but cross-sectional responses cannot show whether a diagnosis changed a health rating, whether poor perceived health led to investigation, or whether common factors influenced both.

Discordance and measurement choices

Good or better health with several lifetime diagnoses may reflect remote, treated or manageable conditions, the equal weighting of very different conditions, past rather than active cancer, or appraisal relative to age; these explanations were not measured. The finding shows that a restricted count and a general health rating are not equivalent constructs, not that the diagnoses are unimportant or that need is absent.

Fair or poor health with a zero count likewise should not be labeled unexplained illness: undiagnosed conditions, functional difficulty, mental-health symptoms and unlisted diseases lay outside the measure, and access to diagnosis can affect a reported count. A zero count must not become a screening rule for low need.

Item time frames matter. Lifetime and current asthma definitions serve different purposes, and mixed time frames remain across the other components. Excluding borderline diabetes avoids forcing a distinct response into either category but cannot resolve its heterogeneity. Removing respondents whose only cancer was non-melanoma skin cancer changed the estimates little.

The leave-one-component results show the consequences of the list rather than a search for an optimal score [4–6]. A comprehensive multimorbidity measure would need a wider list, current-status definitions, severity or function information, and separate validation.

Uncertainty, selection and generalizability

Interview weights and masked design variables address the probability sample, but the valid-response domain remains subject to item missingness and reporting error. The results describe the civilian noninstitutionalized population in 2017–2018 and cannot establish current prevalence, regional need or post-pandemic change.

Only fifteen design degrees of freedom were available, and intervals were much wider with the residual degrees of freedom of the adjusted model, although all lower limits remained above one. Small cells, such as excellent health in the highest count group, should not support detailed subgroup narratives [25,26].

The adjusted model has further limitations: linear age and income terms, top coding, a heterogeneous residual race category, and no severity, function or unlisted diagnoses. Fair or poor health was more common among participants excluded for missing covariates, although a missing-income category and multiple imputation gave similar odds ratios; imputation assumes that income is missing at random given the modeled variables.

Because fair or poor health was common in the upper groups, odds ratios overstate relative prevalence: the adjusted odds ratio of 10.20 for at least three groups corresponded to a standardized prevalence ratio of 5.00 [23]. The absolute and standardized prevalences are the more direct description.

Implications

The findings support reporting patient-reported general health alongside diagnosis-based summaries in descriptive research. Survey researchers should show the actual condition list, preserve the response categories, report retained and excluded populations, and distinguish observed or standardized prevalences from odds ratios. The results do not validate a clinical instrument or inform individual treatment.

Future work could examine diagnosis combinations, current-disease definitions, function or longitudinal change, each with its own question and analytic plan.

Conclusion

Among adults aged at least 20 in NHANES 2017–2018, higher counts of five selected ever-diagnosed condition groups were associated with more fair or poor self-rated health, but the two measures diverged for many adults: some with no selected diagnoses rated their health unfavorably, and most with at least three rated it good or better. These exploratory cross-sectional findings support transparent joint description, not a comprehensive multimorbidity estimate, a causal effect, an individual prognosis or a clinical decision rule.

Declarations

Institutional Review Board Statement: The NHANES protocols were approved by the NCHS Ethics Review Board; the 2017–2018 cycle was conducted under the continuation of protocol 2011-17 and under protocol 2018-01 from October 26, 2017 [27]. This approval applies to the source survey. The present analysis used the public-use data for statistical analysis under the NCHS data-user agreement, without re-identification or linkage to individually identifiable information [28]. This study was exempt from ethics review because the data are public (secondary analysis of public, de-identified data).

Data and Code Availability Statement: No new data were generated. The analysed public-use NHANES 2017–2018 files are available from the National Center for Health Statistics [12,16–20]. Python code that downloads these files and reproduces every count and estimate in this article is provided as Supplementary Code S1.

Informed Consent Statement: Participants in the source survey gave consent to NCHS for the household interview [29]. This secondary analysis used de-identified public-use data, with no contact with participants.

Funding: This research received no funding.

Conflicts of Interest: The authors declare no competing interests.

Author contributions: [Pending: CRediT author contributions and confirmation that all authors approved the final version — to be supplied by the authors]

References

1. Idler EL, Benyamini Y. Self-rated health and mortality: a review of twenty-seven community studies. J Health Soc Behav. 1997;38(1):21-37. doi:10.2307/2955359

2. DeSalvo KB, Bloser N, Reynolds K, He J, Muntner P. Mortality prediction with a single general self-rated health question. A meta-analysis. J Gen Intern Med. 2006;21(3):267-75. doi:10.1111/j.1525-1497.2005.00291.x

3. Jylhä M. What is self-rated health and why does it predict mortality? Towards a unified conceptual model. Soc Sci Med. 2009;69(3):307-16. doi:10.1016/j.socscimed.2009.05.013

4. Fortin M, Stewart M, Poitras ME, Almirall J, Maddocks H. A systematic review of prevalence studies on multimorbidity: toward a more uniform methodology. Ann Fam Med. 2012;10(2):142-51. doi:10.1370/afm.1337

5. Huntley AL, Johnson R, Purdy S, Valderas JM, Salisbury C. Measures of multimorbidity and morbidity burden for use in primary care and community settings: a systematic review and guide. Ann Fam Med. 2012;10(2):134-41. doi:10.1370/afm.1363

6. Nicholson K, Almirall J, Fortin M. The measurement of multimorbidity. Health Psychol. 2019;38(9):783-90. doi:10.1037/hea0000739

7. Mavaddat N, Valderas JM, van der Linde R, Khaw KT, Kinmonth AL. Association of self-rated health with multimorbidity, chronic disease and psychosocial factors in a large middle-aged and older cohort from general practice: a cross-sectional study. BMC Fam Pract. 2014;15:185. doi:10.1186/s12875-014-0185-6

8. Perruccio AV, Katz JN, Losina E. Health burden in chronic disease: multimorbidity is associated with self-rated health more than medical comorbidity alone. J Clin Epidemiol. 2012;65(1):100-6. doi:10.1016/j.jclinepi.2011.04.013

9. Walker ER, Druss BG. Cumulative burden of comorbid mental disorders, substance use disorders, chronic medical conditions, and poverty on health among adults in the U.S.A. Psychol Health Med. 2017;22(6):727-35. doi:10.1080/13548506.2016.1227855

10. Schiltz NK. Prevalence of multimorbidity combinations and their association with medical costs and poor health: a population-based study of U.S. adults. Front Public Health. 2022;10:953886. doi:10.3389/fpubh.2022.953886

11. Mossadeghi B, Caixeta R, Ondarsuhu D, Luciani S, Hambleton IR, Hennis AJM. Multimorbidity and social determinants of health in the US prior to the COVID-19 pandemic and implications for health outcomes: a cross-sectional analysis based on NHANES 2017-2018. BMC Public Health. 2023;23(1):887. doi:10.1186/s12889-023-15768-8

12. National Center for Health Statistics. NHANES 2017–2018 overview [Internet]. [cited 2026 Oct 9]. Available from: https://wwwn.cdc.gov/Nchs/Nhanes/continuousnhanes/overview.aspx?BeginYear=2017

13. STROBE Initiative. STROBE checklists [Internet]. [cited 2026 Oct 9]. Available from: https://www.strobe-statement.org/checklists/

14. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Epidemiology. 2007;18(6):800-4. doi:10.1097/EDE.0b013e3181577654

15. National Center for Health Statistics. Guidelines for high quality analyses of NHANES data [Internet]. [cited 2026 Oct 9]. Available from: https://wwwn.cdc.gov/nchs/nhanes/QualityAnalysesGuidelines.aspx

16. National Center for Health Statistics. NHANES 2017–2018 demographic variables and sample weights [Internet]. [cited 2026 Oct 9]. Available from: https://wwwn.cdc.gov/Nchs/Data/Nhanes/Public/2017/DataFiles/DEMO_J.htm

17. National Center for Health Statistics. NHANES 2017–2018 hospital utilization and access to care [Internet]. [cited 2026 Oct 9]. Available from: https://wwwn.cdc.gov/Nchs/Data/Nhanes/Public/2017/DataFiles/HUQ_J.htm

18. National Center for Health Statistics. NHANES 2017–2018 medical conditions [Internet]. [cited 2026 Oct 9]. Available from: https://wwwn.cdc.gov/Nchs/Data/Nhanes/Public/2017/DataFiles/MCQ_J.htm

19. National Center for Health Statistics. NHANES 2017–2018 diabetes [Internet]. [cited 2026 Oct 9]. Available from: https://wwwn.cdc.gov/Nchs/Data/Nhanes/Public/2017/DataFiles/DIQ_J.htm

20. National Center for Health Statistics. NHANES 2017–2018 blood pressure and cholesterol [Internet]. [cited 2026 Oct 9]. Available from: https://wwwn.cdc.gov/Nchs/Data/Nhanes/Public/2017/DataFiles/BPQ_J.htm

21. National Center for Health Statistics. NHANES tutorials: weighting [Internet]. [cited 2026 Oct 9]. Available from: https://wwwn.cdc.gov/nchs/nhanes/tutorials/weighting.aspx

22. National Center for Health Statistics. NHANES tutorials: variance estimation [Internet]. [cited 2026 Oct 9]. Available from: https://wwwn.cdc.gov/nchs/nhanes/tutorials/varianceestimation.aspx

23. Barros AJ, Hirakata VN. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. BMC Med Res Methodol. 2003;3:21. doi:10.1186/1471-2288-3-21

24. Lumley T. Analysis of complex survey samples. J Stat Softw. 2004;9(8):1-19. doi:10.18637/jss.v009.i08

25. National Center for Health Statistics. NHANES tutorials: reliability of estimates [Internet]. [cited 2026 Oct 9]. Available from: https://wwwn.cdc.gov/nchs/nhanes/tutorials/reliabilityofestimates.aspx

26. Parker JD, Talih M, Malec DJ, Beresovsky V, Carroll M, Gonzalez JF, et al. National Center for Health Statistics Data Presentation Standards for Proportions. Vital Health Stat 2. 2017;(175):1-22.

27. National Center for Health Statistics. NHANES ethics review board approval [Internet]. [cited 2026 Oct 9]. Available from: https://www.cdc.gov/nchs/nhanes/about/erb.html

28. National Center for Health Statistics. NCHS data user agreement [Internet]. [cited 2026 Oct 9]. Available from: https://www.cdc.gov/nchs/policy/data-user-agreement.html

29. National Center for Health Statistics. NHANES home interview consent, 2017–2018 [Internet]. [cited 2026 Nov 1]. Available from: https://wwwn.cdc.gov/nchs/data/nhanes/public/2017/documents/2017_household_interview_consent.pdf

Multidisciplinary Healthcare Journal · Vol. 1, No. 1 (2025) · CC BY 4.0 · https://creativecommons.org/licenses/by/4.0/