Inequity of NIH cancer funding in the United States: an ecological study predicting funding based on disease burden from 2008 through 2023

Eli Berglas,[a] David Musheyev,[a] Aaron B. Lavi,[a] Rachel S. Berglas,[b] Rachel Berglas,[a] and Abdo E. Kabarriti[a,*]

[a]Department of Urology, State University of New York Downstate Health Sciences University, New York City, USA
[b]Albert Einstein College of Medicine, Bronx, NY, USA

The Lancet Regional Health - Americas 2025;45: 101081

Published Online xxx

https://doi.org/10.1016/j.lana.2025.101081

Summary

Background Disease burden has been used to predict National Institutes of Health (NIH) funding but included diseases with little underlying relationship. Here we focus on cancers to create a more appropriate model to allow for more targeted scrutinization of funding allocation.

Methods An ecological study using NIH funding data (2008–2023) was performed. Inclusion of cancers was based on their presence in the NIH Research Portfolio Online Reporting Tool and the 2021 Global Burden of Disease (GBD) study. Disability-adjusted life years (DALY) were collected and to evaluate the impact of public interest, Google Trends data was used. Multivariable linear regression determined appropriate funding based on disease burden and public interest. To quantify how each cancer's funding differed from model predictions residual values were used to calculate the percent over/under funding.

Findings Fifteen cancers met inclusion criteria. Neuroblastoma had the greatest ratio of funding to DALYs per 100,000 people (US$14,000,000) while lung cancer had the lowest (US$300,000). Stomach cancer was the most underfunded (197.9% [95% CI: 136.0%, 276.2%]) while brain cancer was the most overfunded (64.1% [95% CI: 53.8%, 72.1%]). Even at their lowest funding values in the study period brain, breast, and colorectal cancer all had greater than 40% overfunding. Contrarily, the lowest annual funding for leukemia, uterine, and stomach cancer received less than 150% of expected funding. Despite its overfunding brain cancer had an increase in DALYs in the study period.

Interpretation Modeling by disease category demonstrated disparities in funding indicating the need for reevaluation for possible funding inequities. The year-by-year approach taken in this study will drive the ability for future research to better understand NIH funding decisions. Additionally, the role of public interest in research funding needs to be further evaluated to ensure that popularity does not override disease burden, in funding decisions.

Funding No Funding.

Keywords: Medical ethics; Equity; Health disparity; NIH; National institute of health; Research funding; Grant funding; Health economics; Health policy; Global health; Burdens of disease; DALY; Cancer

Introduction

The National Institutes of Health (NIH) is an important source of funding for researchers with roughly US$45 billion in grants being distributed in 2022.[1] It has set categories for funding using the Research, Condition, and Disease Categorization (RCDC) system to data mine for keywords of disease, conditions, and topics.[2] For each keyword category, funding allocation is based on the requests and recommendations of the scientific community and NIH staff.[3] It is inevitable that under the current framework, certain pathologies will not receive research funding representative of their disease burden. Within the RCDC reporting system, the NIH did not include disease burden metrics, such as disability-adjusted life year (DALY), years of healthy life lost due to disability (YLD), or years of life lost (YLL).

*Corresponding author. South Brooklyn Health, SUNY Downstate Health Sciences University, Department of Urology, 2601 Ocean Parkway, Suite 7N38, Brooklyn, NY 11235, USA.
E-mail address: abdo.kabarriti@downstate.edu (A.E. Kabarriti).

Research in context

Evidence before this study

We searched PubMED for studies using disease burden to predict NIH funding using a regression analysis. The following keywords were searched: ("disease burden" OR "incidence" OR "mortality" OR "prevalence" OR "disability-adjusted life years" OR "years of life lost" OR "years lived with disability" OR "advocacy" OR "lobbying" OR "popularity" OR "Google trends") AND ("funding" OR "National Institute of Health"). Additionally, given our interest in assessing the public's role in cancer funding, we used the Google search engine to identify material that pertained to advocacy and lobbying of cancer funding. Articles published prior to September 15, 2024 were included.

A majority of studies only used a single year of data which led to inconsistent conclusions. Furthermore, all the studies included a wide array of unrelated diseases. Analyzing disease burden and research funding trends over a multiple year period for a single disease category would provide more actionable data to identify inequities in funding based on disease burden.

Added value of this study

Our study sets a new standard for the methodological rigor required to ascertain inequities in funding allocation for various pathologies. We identified several cancers with over/underfunding that requires further scrutinization. We also demonstrated the impact public interest can have on disease funding, signifying the importance of including this parameter in future analyses. Our longitudinal approach enabled a robust analysis of cancer trends and whether NIH funding matched disease burden variations. In addition, our model depicted cancer funding inequities. Thus, we supplied the initial equity warnings for funding allocations and set the stage for future studies to devise the best funding strategies for research and innovation.

Implications of all the available evidence

The trends we identify within this study provide actionable data for health experts and policymakers. Swift reevaluation is required across several of the included cancers. The present study provides a multitude of potential starting points for future researchers to further analyze our findings on an individual grant-level analysis. This study also acts as a primer for future research to further explore the funding disparities we identified as well as the temporal effects NIH funding can have on disease burden. Furthermore, this study should act as a precedent for future studies to model potential NIH funding inequities using a single disease category. Given our identification of the correlation between public interest and NIH funding, the need for more transparency behind funding decisions becomes paramount. While advocacy and lobbying for disease funding can have positive effects, the findings of this study indicate funding agencies must be diligent in funding diseases equitably.

These metrics can inform budget appraisers for each funding category and explain any current disparities.

Studies investigating the burden of disease categories, incidence, prevalence, mortality, YLL, and DALY, showed that only DALY was a predictor of funding dispersal.[4,5] However, this finding is inconsistent and is likely related to year-to-year changes.[6] Nevertheless, DALY measures have been a staple of previous studies and have shown funding disparities for a wide variety of diseases.[7,8]

Our model is unique as it compares cancers across two variables (DALYs and Google searches) to project annual NIH funding over a sixteen-year interval. Evaluation of funding over an extended period enables the formation of contextual hypotheses to describe funding disparities and the impact of NIH funding. We also quantify overall funding disparities amongst cancers and assess the effect of public interest on NIH funding allocations. The goal of our study was to identify inappropriately funded cancers that require individual grant-level scrutinization.

Methods

This cross-sectional epidemiological study used the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines.[9] This study did not require institutional review board approval since all information is publicly available and no patients were involved. To determine which cancers were included in the analysis, a survey of the 2021 Global Burden of Disease (GBD) study was performed.[10] Neoplasms category in the GBD study was queried and results are shown in Supplementary Table S1. The GBD study quantifies burden of disease, one measure of which is DALYs. The 2021 GBD study estimates DALYs from 1990 through 2021 using data from disease registries, surveys, and open-source databases. To calculate DALYs, YLD is first calculated using disability weights which measure how disease outcomes affect the level of health.[11] YLL is calculated by multiplying estimated deaths by predicted life expectancy based on demographic factors.[12] DALYs is then a summation of YLD and YLL.

To determine funding, the NIH Research Portfolio Online Reporting Tool (RePORT) was used. NIH RePORT is an online repository of NIH funding sorted by disease.[2] This is done using RCDCs which group funding efforts based on keywords within project descriptions. Collection of this data began in 2008; however new categories have been added and thus have fewer years of data. As outlined in Supplementary

Table S1, all cancers included in the GBD study were evaluated for matching RCDC categories to allow a link between disease burden and NIH funding.

Google Trends[13] data, using the Glimpse[14] extension, was included to determine the relationship between public interest and NIH funding. This extension tracks the number of searches that relate to a keyword. Glimpse has been used as a measure of public interest in prior studies.[15–18] The keywords inputted for each cancer are shown in Supplementary Table S1. Construction of keywords prioritized simple keywords to enable "the best matching" to all possibly related search terms.

Assuming that NIH funding is predicated upon disease burden and/or public interest, there is a need to lag these variables. Following a prior study's precedent, a lag of two years was considered appropriate.[7] Given the GBD study includes data through 2021 and NIH Report includes data from 2008 to 2023, DALYs were collected from the years 2006–2021 and the same was done for Google Searches. DALY estimates were calculated as rates per 100,000 persons including only patients within the United States. Google searches were collected as total counts of searches. All NIH funding values were inflation-adjusted to represent a dollar value representative of January 1st, 2024.[19] For cancers that did not have NIH funding data available for the entire period (2008–2023) DALYs and Google search values were included from two years before the first reported funding amount.

Data analysis

GBD provides annual DALY estimates with a 95% confidence interval (CI). Given the stability of the provided CIs, we averaged DALY point estimates, and a CI was calculated based on the standard error of the mean (SEM). Hodgkin and non-Hodgkin lymphoma data were combined. Lymphoma data combination was done on a year-by-year basis. Point estimates were summed. Standard error (SE) was calculated based on the original CIs. The square root of the summed squared SEs were calculated and multiplied by 1.96 to obtain a new 95% CI. DALYs were graphed with the original 95% CIs determined by the 2021 GBD study.[10]

To assess the effect of DALYs and Google searches on funding, univariable linear regression was performed. Three data points from each cancer were used. The year with the highest and lowest NIH funding was found for each cancer and was paired with the independent variable from two years prior, resulting in a two-year lag. The third data point was the cancer's average NIH funding and the average of DALYs or Google searches. This approach mitigated factors like autocorrelation while capturing data variability. Sample size was comparable to previous studies.[6,7] Univariable linear regression assessed if both DALYs and Google searches met an $\alpha < 0.2$ to validate their inclusion in multivariable regression. Multivariable regression was performed using DALYs and Google searches as predictors of NIH funding. Multicollinearity was assessed using variance inflation factor (VIF) with an a priori limit of 5 considered acceptable. The model met this requirement (VIF = 1.5).

All data was log-transformed. Scatter plots (linearity), Q-Q plots and Shapiro–Wilk test (normality), Residual vs. fitted value plots and Breusch–Pagan (heteroscedasticity), and Breusch-Godfrey (autocorrelation) evaluated assumptions of linear regression. Adjusted R^2 were obtained and 95% CIs were calculated using non-parametric bootstrapping with 1000 resamples summarized at the 2.5th and 97.5th percentiles. Calculation of coefficient 95% CIs used SEs of the coefficient calculated from the model.

Predicted values were calculated using multivariable linear regression. The remainder of the dataset was log-transformed and added to the multivariable linear equation. 95% CIs of predicted values were calculated from the standard error of estimate of the predictions. CIs were calculated by multiplying standard error of estimate of the predictions by 1.96 (assuming normality). Predictions and CIs were back-transformed using exponentiation. Using actual funding minus predicted, we calculated percent over/underfunding by dividing the difference by actual funding. The same procedure was done with the CIs. Positive percentages indicated overfunding and negative percentages indicated underfunding. Given the estimative nature of DALYs, sensitivity testing was performed. We modeled two multivariable regressions (one for upper and lower CIs as estimated by the 2021 GBD study) using the methods described above.

Joinpoint regression determined temporal trends of DALYs and Google searches (response variables) across time. Log link function was applied, and constant variance was assumed. Breusch–Pagan tests were used to evaluate heteroscedasticity. Two models of Google searches (prostate: P = 0.031 and pancreatic: P = 0.040) were significant. Ordinary least squares method was still used given evaluation of residual vs. fitted plots showed little evidence of heteroscedasticity (Supplementary Fig. S1). Joinpoints were determined by finding trends that resulted in a significant change (P < 0.05). A joinpoint could not occur within two data points of another or two points from the beginning or end of the available data. The maximum number of joinpoints are outlined in Supplementary Table S2.[20] Model selection was done using weighted Bayesian information criterion (BIC).[21] The model with the lowest weighted BIC was chosen. Annual percent change (APC) was calculated between joinpoints and average annual percent change (AAPC) was calculated for the entire period. APC was calculated as the exponential of the slope of the fitted regression minus one multiplied by 100 to represent a percentage change. AAPC was determined by averaging the APCs which were weighted based on duration.[22] Uncorrelated error models were used. For sensitivity analysis, we fit the error models with first-order autocorrelation of 0.3.[23] APC and AAPC CIs were estimated using the empirical quantile method with 1000 resamples summarizing results at the 2.5th and 97.5th percentiles.[24]

The role of institution-specific funding is shown proportionally based on total funding. R (version 4.3.3) was used for all analysis except joinpoint regression which was done using Joinpoint Regression Program Statistical Research and Applications Branch, National Cancer Institute (Version 5.2.0.0). P values < 0.05 were considered statistically significant. NIH funding values were presented as United States dollars with standard deviation (SD). DALYs were presented as a count per 100,000 people with 95% CI. Google searches were presented as total counts with SD. In the Supplementary materials, cancers were stratified into three groups based on average funding.

Role of funding source
No funding.

Results
Fifteen cancers were included in the study. On average, breast cancer was highest (US$929.6, SD = US$97.0), and stomach cancer was lowest (US$43.3, SD = US$10.1) funded. Breast cancer had the highest funding throughout but there was annual variability regarding the lowest-funded cancer (Supplementary Fig. S2). Lung cancer contributed the greatest average DALYs (1200.2 [95% CI: 1151.0, 1249.4]) and was the greatest across all years while the opposite was true for neuroblastoma (5.0 [95% CI: 4.8, 5.2]) (Supplementary Fig. S3). Breast cancer had the greatest public interest recording an average of 102,922,648.8, SD = 20,705,034.2 Google searches with esophageal cancer generating the fewest (3,949,404.0, SD = 434,454.3). Breast cancer consistently had the greatest annual Google search counts (Supplementary Fig. S4). Neuroblastoma had the greatest funding ratio to DALYs (US$14,000,000) while lung cancer had the lowest (US$300,000). Colorectal cancer had the highest (US$60.7) and brain cancer had the lowest (US$3.9 per Google search) ratio of funding dollars per Google search (Table 1).

To determine if DALYs can independently predict NIH funding, a univariable regression was performed with assumption testing (Supplementary Fig. S5). The regression was significant ($P < 0.0001$) with an $R^2 = 0.35$ [95% CI: 0.06, 0.52]. The coefficient of Log DALYs was 0.49 [95% CI: 0.29, 0.69] and the intercept was 2.69 [95% CI: 1.66, 3.73], $P < 0.0001$. Supplementary Fig. S6 shows that the univariable regression of Google searches predicting NIH funding met assumptions. This model was significant ($P < 0.0001$) with an $R^2 = 0.39$ [95% CI: 0.14, 0.82]. The coefficient of Log Google searches was 0.64 [95% CI: 0.40, 0.87], and the intercept was −5.40 [95% CI: −9.36, −1.45], $P = 0.0090$. As both univariable models met inclusion criteria ($\alpha < 0.2$), a multivariable regression using DALYs and Google searches (independent variables) and NIH funding (dependent variable) was performed which met assumptions (Supplementary Fig. S7). The regression was significant ($P < 0.0001$) with an $R^2 = 0.47$ [95% CI: 0.24, 0.85]. The coefficient of Log DALYs was 0.29 [95% CI: 0.07, 0.51],

	NIH funding[a] [SD] ($ Millions)	DALY [95% CI] (per 100,000 people)	Google Searches [SD] (Count)	Millions of Dollars/DALYs per 100,000 People	Dollar/Google Search
Breast cancer	929.6 [97.0]	437.6 [431.2, 444.0]	102,922,648.8 [20,705,034.2]	2.1	9.0
Lung cancer	406.2 [111.1]	1200.2 [1151.0, 1249.4]	41,329,546.5 [5,784,932.6]	0.3	9.8
Brain cancer	405.2 [49.7]	173.5 [170.3, 176.8]	10,341,742.3 [967,070.8]	2.3	3.9
Colorectal cancer	387.2 [25.8]	502.6 [500.3, 504.8]	6,372,840.0 [655,194.6]	0.8	60.7
Prostate cancer	359.0 [53.2]	248.8 [242.5, 255.1]	32,769,447.3 [2,965,197.2]	1.4	11.0
Lymphoma	316.6 [30.6]	198.9 [193.6, 204.2]	41,120,357.5 [4,550,102.8]	1.6	7.7
Pancreatic cancer	233.6 [52.2]	328.2 [316.2, 340.3]	19,023,394.8 [2,299,765.6]	0.7	12.3
Ovarian cancer	184.5 [22.8]	126.0 [122.5, 129.6]	20,109,181.3 [2,770,076.1]	1.5	9.2
Leukemia	170.3 [79.2]	202.9 [198.7, 207.1]	38,045,355.9 [3,360,136.2]	0.8	4.5
Cervical cancer	140.0 [17.6]	68.9 [67.9, 69.9]	22,338,145.7 [3,233,446.5]	2.0	6.3
Liver cancer	133.5 [32.0]	146.1 [135.2, 157.0]	15,583,726.1 [1,351,094.1]	0.9	8.6
Neuroblastoma	70.0 [31.9]	5.0 [4.8, 5.2]	4,172,551.4 [531,238.1]	14.0	16.8
Esophageal cancer	54.9 [8.2]	141.3 [137.8, 144.8]	3,949,404.0 [434,454.3]	0.4	13.9
Uterine cancer	49.2 [14.2]	74.4 [68.3, 80.6]	4,064,790.2 [467,690.9]	0.7	12.1
Stomach cancer	43.3 [10.1]	108.6 [107.7, 109.5]	10,705,192.5 [635,773.0]	0.4	4.0

Values represent available funding data from 2008 to 2023, DALY and Google search data from 2006 to 2021. CI: Confidence Interval, DALY: Disability Adjusted Life Years, SD = Standard Deviation. [a]Inflation adjusted to January 2024.

Table 1: **Average funding, disability adjusted life years, google searches and ratios among included cancers.**

P = 0.010. The coefficient of Log Google searches was 0.44 [95% CI: 0.17, 0.70], P = 0.0020 and the intercept was −3.55 [95% CI: −7.50, 0.40], P = 0.078.

Back-transformation of the residuals are shown as percent differences from actual funding in Fig. 1. When considering each cancer's highest annual funding value, neuroblastoma was the most overfunded (74.4% [95% CI: 50.0%, 86.9%]). Stomach cancer was the most underfunded receiving 147.7% [95% CI: 95.7%, 213.4%] less than predicted. At average funding value, stomach cancer was the most underfunded (197.9% [95% CI: 136.0%, 276.2%]) while brain cancer was the most overfunded, receiving 64.1% [95% CI: 53.8%, 72.1%] greater funding than predicted. At the lowest funding values brain, breast, and colorectal cancer all showed greater than 40% overfunding. Leukemia, uterine, and stomach cancer all had funding levels that, even when considering uncertainty, received less than 150% of predicted funding. Sensitivity testing was performed and results are summarized in Supplementary Table S3. Substitution with the upper and lower DALY estimates from the GBD study did not result in significant aberrations.

Applying the multivariable linear regression equation to all the data yields the results in Fig. 2. Amongst the top five funded cancers (listed in descending order

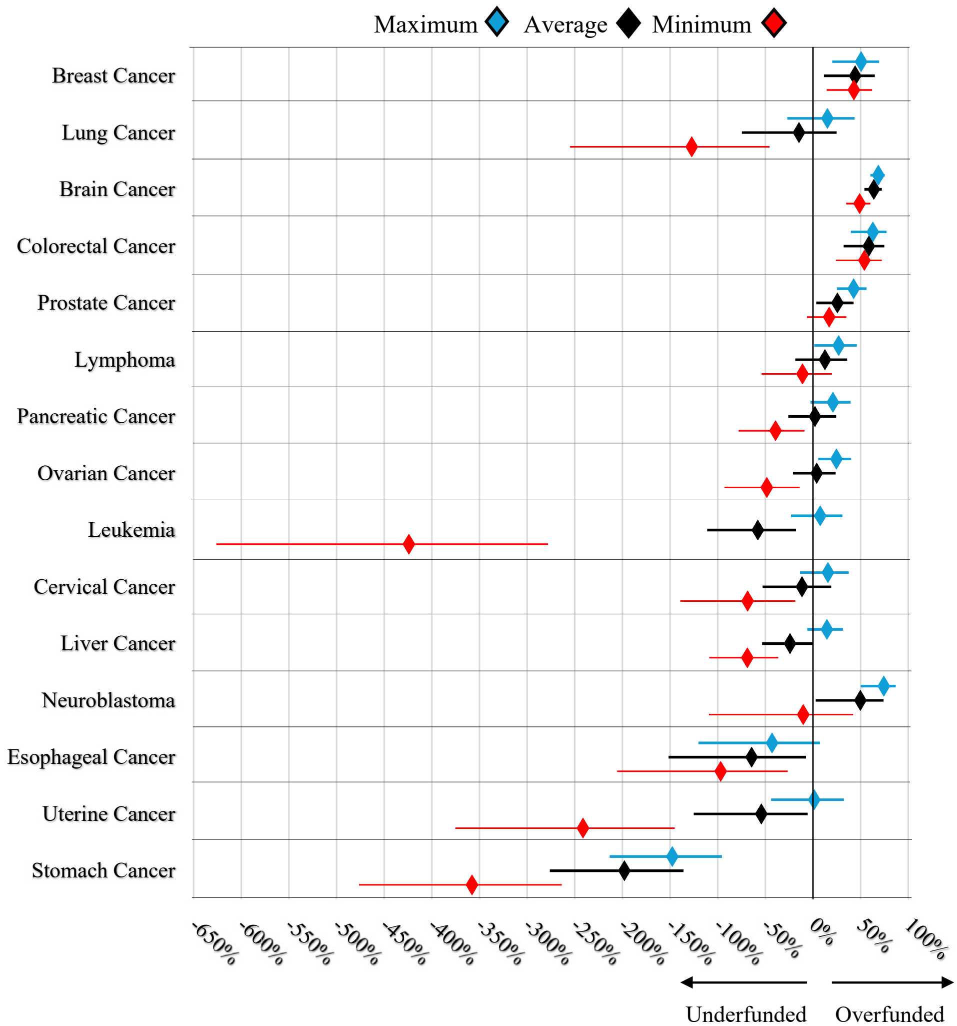

Fig. 1: Percent error of residuals for years with maximum funding, average funding, and minimum funding. Positive values indicate overfunding compared to predicted while negative values indicate underfunding. The three values presented are the maximum funding year (Blue), average funding (Black), and minimum funding year (Red). Values represent residual (actual funding–predicted funding) divided by actual funding. Confidence Intervals (CI) were made by determining the uncertainty of predicted values based on the standard error of estimate of the predictions. Following exponentiation these CIs were subtracted from the actual funding. These values were then divided by actual funding to represent CIs in the figure. For overfunding the theoretical maximum was 100% overfunded. Therefore percent funding discrepancies should be interpreted in the context of actual funding values. Cancers are presented in descending order based on average NIH funding dollars.

Overfunded

Underfunded

	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Breast	41 [-1, 66]	44 [7, 67]	48 [14, 69]	44 [8, 66]	51 [20, 69]	40 [4, 62]	42 [7, 64]	42 [9, 64]	41 [8, 63]	44 [14, 64]	47 [19, 65]	45 [15, 64]	49 [22, 67]	44 [13, 64]	44 [17, 63]	43 [14, 62]
Lung	-127 [-255, -45]	-100 [-209, -29]	-72 [-165, -11]	-56 [-140, -1]	-45 [-123, 5]	-65 [-153, -8]	-38 [-111, 10]	3 [-49, 36]	-4 [-59, 32]	0 [-52, 34]	11 [-35, 41]	13 [-32, 42]	15 [-27, 44]	14 [-28, 43]	18 [-23, 45]	18 [-23, 46]
Brain	49 [35, 60]	59 [47, 68]	63 [53, 71]	64 [53, 72]	64 [53, 73]	64 [52, 72]	63 [52, 71]	64 [53, 72]	63 [53, 71]	64 [54, 72]	67 [57, 74]	64 [55, 72]	66 [57, 73]	69 [60, 75]	68 [58, 76]	68 [58, 76]
Colorectal	56 [31, 72]	60 [35, 75]	61 [36, 76]	63 [40, 77]	62 [37, 77]	59 [32, 75]	57 [29, 75]	63 [38, 78]	58 [29, 75]	54 [24, 72]	61 [34, 77]	55 [27, 73]	58 [31, 74]	57 [32, 73]	56 [30, 73]	55 [27, 72]
Prostate	33 [10, 49]	38 [19, 53]	43 [25, 56]	31 [9, 47]	23 [0, 41]	31 [10, 47]	23 [0, 40]	32 [12, 47]	24 [2, 40]	17 [-6, 35]	22 [0, 39]	19 [-5, 37]	14 [-13, 34]	16 [-9, 36]	14 [-10, 33]	13 [-12, 33]
Lymphoma	-7 [-49, 23]	-11 [-54, 20]	-3 [-40, 25]	-1 [-36, 26]	8 [-22, 31]	13 [-16, 36]	18 [-10, 39]	28 [5, 46]	26 [3, 44]	24 [-1, 43]	27 [1, 46]	9 [-25, 33]	10 [-26, 35]	12 [-25, 37]	17 [-14, 39]	18 [-12, 40]
Pancreatic	..	..	..	-39 [-78, -9]	-23 [-59, 4]	-34 [-72, -5]	-35 [-74, -5]	3 [-25, 24]	-1 [-30, 22]	12 [-14, 32]	17 [-8, 36]	14 [-11, 34]	13 [-12, 33]	7 [-21, 28]	21 [-3, 39]	19 [-5, 38]
Ovarian	-48 [-93, -14]	-30 [-67, -2]	-3 [-29, 18]	2 [-24, 23]	10 [-14, 28]	0 [-25, 21]	-1 [-26, 20]	-15 [-45, 9]	7 [-17, 26]	8 [-16, 27]	11 [-13, 29]	16 [-5, 33]	25 [5, 40]	19 [-2, 36]	21 [1, 36]	11 [-11, 28]
Leukemia	-424 [-626, -278]	-315 [-465, -205]	-254 [-376, -163]	-238 [-357, -151]	-161 [-249, -94]	-205 [-310, -127]	-99 [-168, -47]	-30 [-73, 2]	-32 [-75, 0]	-16 [-53, 13]	-4 [-37, 21]	-19 [-57, 10]	-23 [-66, 8]	8 [-23, 31]	1 [-31, 24]	4 [-27, 27]
Cervical	-68 [-139, -19]	-36 [-92, 4]	-21 [-68, 13]	-7 [-55, 26]	-3 [-42, 25]	-20 [-65, 13]	0 [-37, 26]	-15 [-55, 15]	-18 [-59, 13]	-8 [-47, 21]	-12 [-53, 18]	-15 [-55, 15]	-12 [-53, 18]	-7 [-48, 22]	15 [-14, 36]	16 [-14, 38]
Liver	-29 [-63, -2]	-19 [-49, 5]	-10 [-37, 12]	-51 [-87, -21]	-56 [-94, -26]	-69 [-109, -36]	-64 [-104, -33]	-42 [-76, -15]	-48 [-83, -19]	-47 [-82, -19]	-20 [-49, 3]	-11 [-37, 10]	-13 [-39, 9]	-17 [-45, 5]	15 [-6, 31]	13 [-8, 30]
Neuroblastoma	..	..	..	-10 [-109, 42]	21 [-49, 59]	21 [-51, 59]	25 [-45, 61]	11 [-70, 54]	23 [-50, 60]	47 [-2, 73]	45 [-7, 72]	49 [1, 74]	57 [15, 78]	74 [50, 87]	73 [48, 86]	75 [51, 87]
Esophageal	..	..	..	..	..	..	..	..	..	..	..	..	-97 [-206, -27]	-81 [-168, -23]	-48 [-130, 5]	-43 [-120, 7]
Uterine	-241 [-375, -145]	-104 [-192, -43]	-93 [-182, -33]	-33 [-91, 8]	-31 [-89, 9]	-45 [-109, -1]	1 [-44, 32]	-9 [-59, 25]	-17 [-70, 20]	-31 [-94, 11]	-32 [-95, 11]	-75 [-159, -18]	-100 [-199, -34]	-127 [-238, -53]	-81 [-178, -18]	-71 [-162, -11]
Stomach	..	..	..	..	..	..	..	..	..	..	..	..	-358 [-476, -264]	-195 [-271, -135]	-159 [-229, -104]	-148 [-213, -96]

Fig. 2: Annual funding discrepancies from predicted values. Values represent actual funding–predicted funding divided by actual funding. Positive values indicate overfunding (blue values) compared to predicted while negative values indicate underfunding (red values). Confidence Intervals (CI) were made by determining the uncertainty of predicted values based on the standard error of estimate of the predictions. Following exponentiation these CIs were subtracted from the actual funding. These values were then divided by actual funding to represent CIs in the figure. For overfunding the theoretical maximum was 100% overfunded. Therefore percent funding discrepancies should be interpreted in the context of actual funding values. Cancers are presented in descending order based on average NIH funding dollars.

in Fig. 2), lung cancer appears to be the only anomaly which was limited to the years 2008–2014. The middle-funded cancers (6–10) demonstrated similar trends with leukemia showing a similar trajectory to lung cancer. Amongst the lowest-funded cancers, there are limitations given stomach and esophageal cancer have limited data. Stomach cancer was most underfunded for all years of available data.

Putting in context these funding trends, we display DALY and Google search trends in Supplementary Figs. S3 and S4. Joinpoint regression results are summarized in Table 2. Lung cancer, which showed a corrected underfunding trend, also demonstrated a consistent decrease in DALYs. This decrease was most evident from 2014 to 2019 (APC = −2.5% [95% CI: −2.9%, −2.2%], P < 0.0001). Liver cancer had the largest percent increase of DALYs over the study period (AAPC = 3.0% [95% CI: 2.8%, 3.1%], P < 0.0001). This increase was most pronounced at its two earlier joinpoints (2006–2011 & 2011–2017) which coincided with liver cancer being underfunded. Neuroblastoma had the largest overall decrease in DALYs (AAPC = −1.9% [95% CI: −2.0%, −1.7%], P < 0.0001) and was consistently overfunded (Fig. 2). Google trends showed larger percentage changes compared to DALYs. Most cancers had an overall decrease in Google searches. This was most apparent for breast and lung cancer (Table 2). Sensitivity analysis showed the results were relatively consistent when adjusting for possible autocorrelation (Supplementary Table S4).

The "cancer" category on the NIH reporter tool had 15,887.4 average annual grants from 2008 to 2023. Breast cancer received the greatest number of grants (Fig. 3A). All cancers received most of their funding from the National Cancer Institute (NCI). The "cancer" NIH category received 81.8% of its funding from the NCI. The National Institute of General Medical Sciences (NIGMS) contributed the second most to cancer funding (2.7% of total). Neuroblastoma received the greatest proportion of its funding dollars from the NCI (91.6%) while liver cancer received the lowest proportion (69.0%) (Fig. 3B). Among the top ten funded cancers, only the National Institute of Neurological Disorders and Stroke (NINDS) funding of brain cancer and the National Institute of Allergy and Infectious Diseases (NIAID) funding of lymphoma exceeded 5% of a cancer's funding (Supplementary Tables S5 and S6). This was more frequent amongst the lowest five funded cancers as the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) funding of liver, esophageal, and stomach cancer met this 5% threshold for both the number of grants and dollars contributed (Supplementary Table S7).

Discussion

Several studies attempted to quantify the predictors of NIH funding.[4–8] To our knowledge, no study has done a year-by-year analysis of NIH funding on one category of disease, across measures of disease burden and public interest. The advantages of such an approach allow the analysis of trends and speculation of causative factors in a contextual framework.

	Cancer	First Joinpoint			Second Joinpoint			Third Joinpoint			Total	
		Joinpoint (Year)	APC [95% CI] (%)	P Value	Joinpoint (Year)	APC [95% CI] (%)	P Value	Joinpoint (Year)	APC [95% CI] (%)	P Value	AAPC [95% CI] (%)	P Value
DALY	Breast cancer	2006–2013	−0.8 [−1.7, −0.5]	0.030	2013–2021	−0.3 [−0.5, 0.6]	0.28	–	–	–	−0.5 [−0.7, −0.4]	<0.0001
	Lung cancer	2006–2014	−1.3 [−1.4, −1]	<0.0001	2014–2019	−2.5 [−2.9, −2.2]	<0.0001	2019–2021	1.4 [0.4, 2.2]	0.008	−1.3 [−1.4, −1.2]	<0.0001
	Brain cancer	2006–2010	0.5 [−0.4, 1]	0.14	2010–2016	1.3 [1.0, 1.9]	0.020	2016–2021	−0.3 [−0.8, 0.1]	0.084	0.5 [0.4, 0.6]	<0.0001
	Colorectal cancer	2006–2010	−0.5 [−1.5, 0.0]	0.082	2010–2021	0.1 [−0.1, 0.8]	0.078	–	–	–	0.0 [−0.1, 0.1]	0.57
	Prostate cancer	2006–2013	−0.3 [−0.6, −0.1]	0.0020	2013–2021	1.8 [1.6, 2.0]	<0.0001	–	–	–	0.8 [0.7, 0.9]	<0.0001
	Lymphoma	2006–2013	−1.5 [−1.9, −1.3]	<0.0001	2013–2019	−0.8 [−1.1, −0.4]	<0.0001	2019–2021	1 [0.0, 1.5]	0.056	−0.9 [−1.0, −0.8]	<0.0001
	Pancreatic cancer	2009–2013	1.4 [0.8, 1.6]	<0.0001	2013–2016	2.2 [1.8, 2.4]	<0.0001	2016–2021	1.2 [0.9, 1.3]	<0.0001	1.5 [1.4, 1.5]	<0.0001
	Ovarian cancer	–	–	–	–	–	–	–	–	–	−1.1 [−1.2, −0.9]	<0.0001
	Leukemia	2006–2019	−0.9 [−1.3, −0.7]	0.020	2019–2021	0.2 [−0.8, 0.7]	0.86	–	–	–	−0.7 [−0.9, −0.7]	<0.0001
	Cervical cancer	2006–2016	0.4 [−0.1, 0.9]	0.070	2016–2019	−2.6 [−3.2, 0.9]	0.21	2019–2021	−0.3 [−2.1, 0.8]	0.47	−0.3 [−0.5, −0.2]	<0.0001
	Liver cancer	2006–2011	4.3 [3.9, 5.6]	<0.0001	2011–2017	3.0 [2.5, 3.5]	<0.0001	2017–2021	1.2 [0.0, 1.8]	0.048	3.0 [2.8, 3.1]	<0.0001
	Neuroblastoma	–	–	–	–	–	–	–	–	–	−1.9 [−2.0, −1.7]	<0.0001
	Esophageal cancer	–	–	–	–	–	–	–	–	–	1.2 [0.9, 1.5]	<0.0001
	Uterine cancer	2006–2013	3.0 [2.2, 3.3]	<0.0001	2013–2017	4.8 [4.0, 5.8]	<0.0001	2017–2021	1.4 [0.4, 2.1]	0.026	1.4 [0.4, 2.1]	<0.0001
	Stomach cancer	–	–	–	–	–	–	–	–	–	0.0 [−0.9, 1.1]	0.91
Google Searches	Breast cancer	2006–2010	−6.9 [−11.4, −4.4]	<0.0001	2010–2021	−3.1 [−3.7, −1.4]	0.022	–	–	–	−4.1 [−4.7, −3.5]	<0.0001
	Lung cancer	2006–2010	−9.7 [−17.9, −5.5]	0.0040	2010–2019	0.8 [−0.4, 7.2]	0.088	2019–2021	−6.5 [−14.6, −0.5]	0.034	−3.1 [−4.1, −2.2]	<0.0001
	Brain cancer	2006–2011	−2.3 [−7.8, 0.0]	0.056	2011–2018	4 [2.6, 8.6]	<0.0001	2018–2021	−9.9 [−14.9, −5.6]	<0.0001	−1.0 [−1.8, −0.4]	0.0060
	Colorectal cancer	2006–2013	−4.2 [−7.9, −2.4]	<0.0001	2013–2021	3.7 [2.1, 6.8]	<0.0001	–	–	–	0.0 [−0.8, 0.7]	0.88
	Prostate cancer	2006–2015	−2.9 [−5.5, −1.8]	<0.0001	2015–2021	2.0 [−0.1, 6.9]	0.070	–	–	–	−0.9 [−1.7, −0.2]	0.022
	Lymphoma	2006–2014	−3.9 [−5.2, −3.0]	<0.0001	2014–2019	7.6 [5.7, 10.9]	<0.0001	2019–2021	−9.3 [−14, −3.4]	<0.0001	−1.0 [−1.6, −0.5]	<0.0001
	Pancreatic cancer	2009–2016	−0.6 [−8.4, 1.6]	0.37	2016–2019	−0.6 [−8.4, 1.6]	0.016	2019–2021	−13.3 [−22.5, −2.7]	0.024	−0.2 [−1.9, 1.1]	0.62
	Ovarian cancer	2006–2008	−12.8 [−17.2, −4.0]	<0.0001	2008–2018	−0.7 [−1.4, 5.0]	0.62	2018–2021	−7.2 [−14.8, −2.7]	<0.0001	−3.7 [−4.7, −2.7]	<0.0001
	Leukemia	2006–2008	−7.2 [−10.7, −1.2]	0.0020	2008–2021	−1.1 [−3.6, 1.6]	0.16	–	–	–	−1.9 [−2.5, −1.1]	<0.0001
	Cervical cancer	–	–	–	–	–	–	–	–	–	−2.3 [−3.5, −1.2]	0.0020
	Liver cancer	2006–2010	−7.0 [−10.2, −4.7]	<0.0001	2010–2019	1.4 [0.8, 3.6]	0.0040	2019–2021	−5.7 [−9.8, −0.8]	0.006	−1.8 [−2.4, −1.3]	<0.0001
	Neuroblastoma	2009–2017	0.1 [−1.9, 12.8]	0.86	2017–2021	−8.5 [−23.1, −3.3]	<0.0001	–	–	–	−2.9 [−5.1, −0.2]	0.042
	Esophageal cancer	–	–	–	–	–	–	–	–	–	−0.8 [−11.5, 12.1]	0.84
	Uterine cancer	2006–2008	−9.3 [−12.8, −2.9]	<0.0001	2008–2019	−1.2 [−1.6, 3.0]	0.35	2019–2021	−6.7 [−12.0, −2.1]	<0.0001	−3.0 [−3.7, −2.3]	<0.0001
	Stomach cancer	–	–	–	–	–	–	–	–	–	−3.2 [−8.7, 3.1]	0.28

APC values represent percent change in specified time interval. AAPC is a weighted average of the calculated APC. Weighting was based on duration of APC. Esophageal and stomach cancer did not have enough data points to consider joinpoints and thus only a total AAPC was calculated for those cancers. The number of joinpoints was determined based on the model with lowest weighted Bayesian information criterion. APC: Annual percent change, AAPC: Average annual percent change, CI: Confidence Interval, DALY: Disability Adjusted Life Years.

Table 2: **Joinpoint regression results for disability adjusted life years and google searches.**

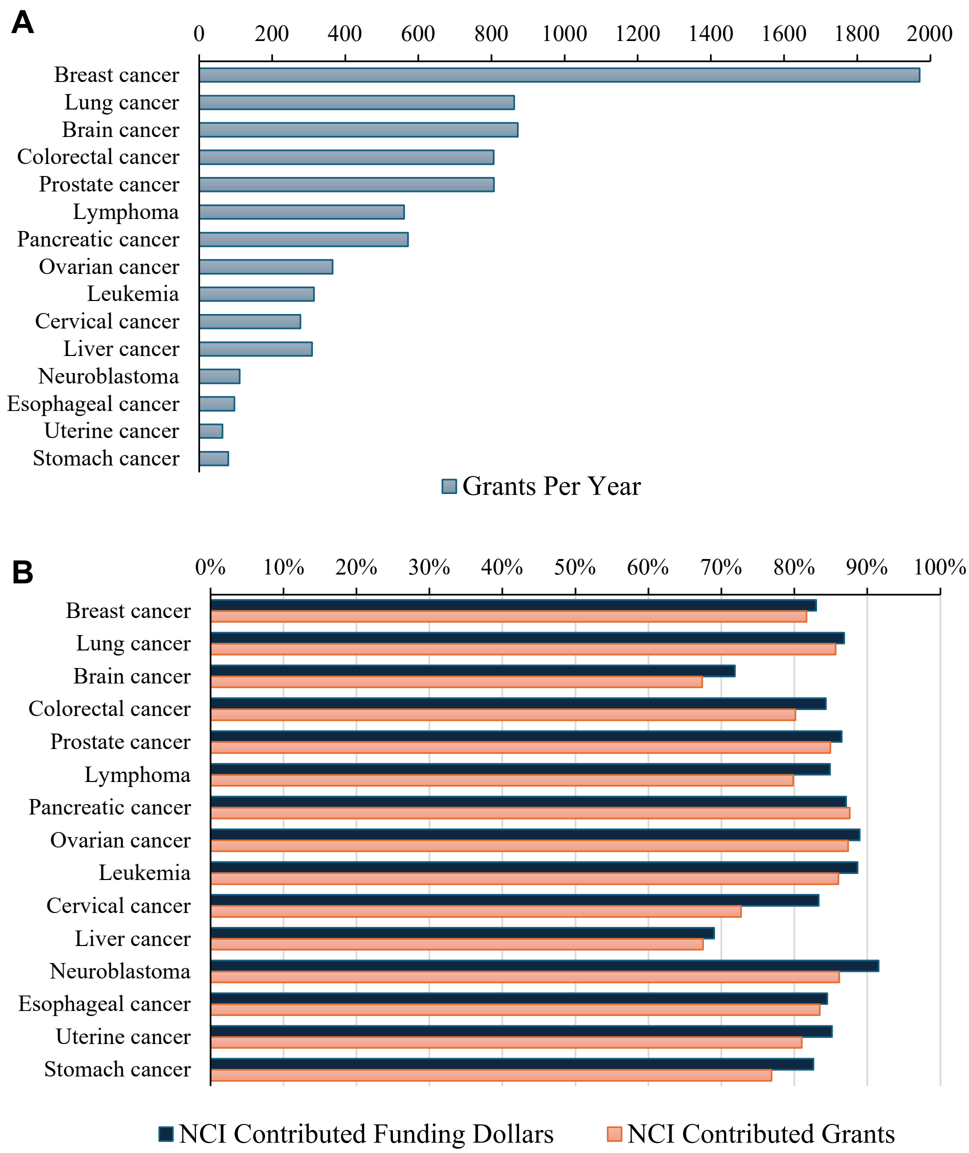

Fig. 3: National institutes of health funding shown by **A.** Number of Grants per Year and **B.** Proportion of total grants and total funding from the National Cancer Institute (NCI). Values represent funding data ranging from 2008 through 2023 although not all cancers have full data. Cancers are presented in descending order based on average NIH funding dollars in descending order. NCI = National Cancer Institute.

Like previous studies, we find that DALYs significantly predicted NIH funding.[4–8] The effect of incremental DALY changes on NIH funding as quantified by the beta coefficient was consistent with a prior study that applied similar methods.[7] Focusing on the presented multivariable model, the beta coefficient of DALYs was equal to 0.29 [95% CI: 0.07, 0.51]. Given the log scale, this indicates that for every one percent increase in DALYs, NIH funding would be expected to rise by 0.29%. Comparatively, the beta coefficient of log Google searches was 0.44 [95% CI: 0.17, 0.70]. While this suggests Google searches are more impactful in determining funding levels, this must be taken into context with the scaling and units of the original value. As shown in Table 1, Google search values are in the order of millions which is not the case for DALYs. Therefore, evaluation of the beta values should be used in the context that a one percent increase of DALYs represents a much smaller true change.

Our results regarding Google searches align with a previous study's findings that demonstrated Google searches to be predictive of cancer funding in South Korea.[25] A causative explanation for this finding is that the number of Google searches may directly indicate public interest. With increased public interest comes greater awareness and the development of more lobbying groups. One study found that advocacy was a direct driver of NIH funding making the relationship of Google searches and NIH funding compared to more obvious measures of disease burden, reasonable.[26] Similar to our findings, a separate study looking specifically at non-profit organizations and advocacy group funding, found that the revenue of these organizations was correlated with disease burden characteristics.[27]

Breast cancer had more than double the funding of any other cancer. However, its average DALY value was comparable to colorectal cancer which had an average funding of less than US$400,000,000 annually. Considering Google searches, breast cancer once again more than doubles any other cancer (Table 1). Both values for breast cancer have been declining with the decline in Google searches being far greater than that of DALYs. These trends seem to indicate a strong relationship between public interest in breast cancer and its subsequent funding. Colorectal cancer, on the other hand, did not experience the same decline in its funding value despite having a similar trend in Google searches from 2006 to 2013 (Table 2). This finding necessitates a closer look at how lobbying efforts for breast cancer have changed. Another outlier shown in Table 1 is neuroblastoma. Despite contributing the fewest average DALYs, its funding-to-disease burden ratio far outweighs all the other cancers. This is shown in the regression models as neuroblastoma demonstrated a trend of overfunding particularly in more recent years (Fig. 2). Promisingly, this overfunding coincided with a decrease in DALYs greater than any other cancer (Table 2). The context of this decrease should be noted in accordance with the low value of DALYs contributed by this disease, therefore the larger percentage drop does not represent a large overall impact. Thus, neuroblastoma can be highlighted as a cancer that requires further scrutiny.

Regarding overfunding, brain cancer, even at its lowest annual funding value, still received greater than predicted funding (Fig. 1). However, brain cancer DALY trends increased moderately throughout the study period (Table 2). This is unlike breast and colorectal cancer, which also were overfunded at their minimum funding values. Demonstrating that even with increasing disease burden, brain cancer still was relatively overfunded. Brain cancer needs to be further evaluated because it receives large sums of funding from non-NCI institutes (Fig. 3B & Supplementary Table S5). The importance of this fact is predicated on the need for coordination between funding agencies. Given that brain cancer is funded disproportional to its disease burden and public interest, there is a need to evaluate the communication between institutions. Stomach cancer was the only cancer that demonstrated a clear pattern of underfunding regarding its highest, average, and lowest NIH funding values (Fig. 1). However, as shown in Fig. 2 there was a clear decrease in its funding discrepancy from 2020 to 2021. This trend has continued although at a slower pace. The number of grants it has received is close to cancers of similar funding and, given the limited time it has been included as a RCDC category, more time is likely needed to clarify the meaning of the trend (Fig. 3A).

A limitation of the model building is that DALYs and Google searches do not comprise all the factors determining funding decisions. Therefore, there is inherent model misspecification. Furthermore, the use of three data points for each cancer which was necessary to avoid deleterious transformations of the data further limits the generalizability of this study. While the present study quantified NIH spending, we did not organize the investments into their respective areas of research, as has been done previously.[28] Lastly, this study evaluated NIH funding which is not the only means of financial investment into cancer research. While a previous study analyzed other sources of funding among cancers, they only did so for a single year, limiting the applicability of their findings.[27]

Future studies should be performed to understand NIH funding temporally by disease category, as comparisons across many different disease categories for a single year yield results that are less actionable. We conclude that brain, breast, and colorectal cancer had greater funding than predicted while leukemia, uterine, and stomach cancer received less than the expected funding. This study provides steppingstones for future studies to examine years of over/underfunding as determined by this study, to uncover the rationale for research funding of each respective cancer. Focusing on a specific cancer would enable a more detailed analysis of factors responsible for funding allocation including identifying each disease's funding sources (not just NIH), modes of research being funded (e.g., basic science, clinical, epidemiological), recent advancements, and political lobbying efforts.

Contributors
Conceptualisation: EB, DM, AEK. Data curation: EB, ABL, RSB. Formal analysis: EB, DM, RSB, RB. Funding acquisition: N/A. Investigation: EB. Methodology: EB. Project administration: EB. Resources: N/A. Software: N/A. Supervision: AEK. Validation: EB, DM. Visualisation: EB, DM. Writing—original draft: EB, DM, ABL. Writing—review & editing: EB, DM, ABL, AEK. EB and RB had full access to the study data. Approval of manuscript submission: All authors.

Data sharing statement
All data is available to be shared for research purposes upon request to eli.berglas@downstate.edu. All data used is publicly available from Institute for Health Metrics and Evaluation Global Burden of Disease 2021 and NIH Research Portfolio Online Reporting Tools.

Declaration of interests
None to declare.

Acknowledgements
Not applicable.

Funding Source: No funding.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.lana.2025.101081.

References

1 NIH, *Grants & funding*. National Institutes of Health (NIH); 2019. Available from: https://www.nih.gov/grants-funding.
2 National Institute of Health, *RePORT*. *report.nih.gov*. National Institute of Health; 2020 [cited 2024 Jul 31]. Available from: https://report.nih.gov/funding/categorical-spending#/.
3 NIH Research Planning, *National institutes of health (NIH)*. U.S. Department of Health and Human Services; 2015 [cited 2024 Jul 31]. Available from: https://www.nih.gov/about-nih/nih-research-planning.
4 Gross CP, Anderson GF, Powe NR. The relation between funding by the national institutes of health and the burden of disease. *N Engl J Med*. 1999;340(24):1881–1887.
5 Gillum LA, Gouveia C, Dorsey ER, et al. NIH disease funding levels and burden of disease. Ross J, editor *PLoS One*. 2011;6(2):e16837.
6 Ballreich JM, Gross CP, Powe NR, Anderson GF. Allocation of national institutes of health funding by disease category in 2008 and 2019. *JAMA Netw Open*. 2021;4(1):e2034890 [cited 2022 Feb 21]; Available from: https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2775642.
7 Nguyen AT, Aris IM, Snyder BD, et al. Musculoskeletal health: an ecological study assessing disease burden and research funding. *Lancet Reg Health Am*. 2024;29:100661 [cited 2024 Jul 31]; Available from: https://www.sciencedirect.com/science/article/pii/S2667193X23002351?via%3Dihub#bib9.
8 Nimgaonkar A, Mughal AY, Heimer H, Nimgaonkar V, Greenstein D, Wright A. Exploring static and dynamic relationships between burden of disease and research funding in the United States. *Health Res Pol Syst*. 2022;20(1).
9 von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ*. 2007;335(7624):806–808. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2034723/.
10 Institute for Health Metrics and Evaluation, *Global health data exchange (GHDx)*. Institute for Health Metrics and Evaluation. University of Washington; 2021. Available from: https://vizhub.healthdata.org/gbd-results/.
11 GBD 2021 Diseases and Injuries Collaborators, Global incidence, prevalence, years lived with disability (YLDs), disability-adjusted life-years (DALYs), and healthy life expectancy (HALE) for 371 diseases and injuries in 204 countries and territories and 811 subnational locations, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet*. 2024;403(10440):2133–2161. Available from: https://doi.org/10.1016/S0140-6736(24)00757-8.
12 GBD 2021 Causes of Death Collaborators, Global burden of 288 causes of death and life expectancy decomposition in 204 countries and territories and 811 subnational locations, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet*. 2024;403(10440):2100–2132. Available from: https://doi.org/10.1016/S0140-6736(24)00367-2.
13 Google, *Google trends*. Google Trends; 2024. Available from: https://trends.google.com/trends/.
14 Glimpse - discover trends before they're trending. Glimpse [cited 2024 Aug 1]. Available from: https://meetglimpse.com/.
15 Amusa LB, Twinomurinzi H, Okonkwo CW. Modeling COVID-19 incidence with Google trends. *Front Res Metr Anal*. 2022;7:1003972. Available from: https://doi.org/10.3389/frma.2022.1003972.
16 Kaleem T, Malouff TD, Stross WC, et al. Google search trends in oncology and the impact of celebrity cancer awareness. *Cureus*. 2019;11(8):e5360. Available from: https://doi.org/10.7759/cureus.5360.
17 Schootman M, Toor A, Cavazos-Rehg P, et al. The utility of Google Trends data to examine interest in cancer screening. *BMJ Open*. 2015;5(6):e006678.
18 Wang D, Lang JC, Chen Y-H. Assessment of using Google Trends for real-time monitoring of infectious disease outbreaks: a measles case study. *Sci Rep*. 2024;14(1):9470. Available from: https://doi.org/10.1038/s41598-024-60120-8.
19 Bls.gov [cited 2025 Feb 10]. Available from: https://www.bls.gov/data/inflation_calculator.htmhttps://surveillance.cancer.gov/joinpoint/.
20 Number of Joinpoints, Joinpoint help system. Available from: https://surveillance.cancer.gov/help/joinpoint/setting-parameters/method-and-parameters-tab/number-of-joinpoints.
21 Kim HJ, Chen HS, Midthune D, et al. Data-driven choice of a model selection method in joinpoint regression. *J Appl Stat*. 2022:1–22.
22 Clegg LX, Hankey BF, Tiwari R, Feuer EJ, Edwards BK. Estimating average annual per cent change in trend analysis. *Stat Med*. 2009;28(29):3670–3682.
23 *Estimating the autocorrelation parameter*. *Joinpoint help system*; 2025 [cited 2025 Mar 5]. Available from: https://surveillance.cancer.gov/help/joinpoint/tech-help/frequently-asked-questions/estimating-the-autocorrelation-parameter.
24 Kim HJ, Luo J, Chen HS, et al. Improved confidence interval for average annual percent change in trend analysis. *Stat Med*. 2017;36(19):3059–3074.
25 Jung YL, Yoo HS, Kim ES. The relationship between government research funding and the cancer burden in South Korea: implications for prioritising health research. *Health Res Pol Syst*. 2019;17(1).
26 Best RK. Disease politics and medical research funding. *Am Sociol Rev*. 2012;77(5):780–803.
27 Kamath SD, Kircher SM, Benson AB. Comparison of cancer burden and nonprofit organization funding reveals disparities in funding across cancer types. *J Natl Compr Cancer Netw*. 2019;17(7):849–854. Available from: https://jnccn.org/view/journals/jnccn/17/7/article-p849.xml.
28 Hagstrom EL, Patel S, Karimkhani C, et al. Comparing cutaneous research funded by the US National Institutes of Health (NIH) with the US skin disease burden. *J Am Acad Dermatol*. 2015;73(3):383–391.e1.