BMJ Open

A county-level cross-sectional analysis of positive deviance to assess multiple population health outcomes in Indiana

Michael Hendryx,[1] Lucia Guerra-Reyes,[2] Benjamin D Holland,[1] Michael Dean McGinnis,[3] Emily Meanwell,[4] Susan E Middlestadt,[5] Karen M Yoder[6]

To cite: Hendryx M, Guerra-Reyes L, Holland BD, *et al*. A county-level cross-sectional analysis of positive deviance to assess multiple population health outcomes in Indiana. *BMJ Open* 2017;**7**:e017370. doi:10.1136/bmjopen-2017-017370

► Prepublication history for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2017-017370).

Received 24 April 2017
Revised 11 August 2017
Accepted 23 August 2017

[1]Department of Environmental and Occupational Health, Indiana University, Bloomington, Indiana, USA
[2]Department of Applied Health Science, Indiana University, Bloomington, Indiana, USA
[3]Department of Political Science, Indiana University Bloomington, Bloomington, Indiana, USA
[4]Social Science Research Commons, Indiana University Bloomington, Bloomington, Indiana, USA
[5]Department of Applied Health Science, Indiana University, Bloomington, USA
[6]Office of Civic Engagement, Indiana University School of Dentistry, Indianapolis, Indiana, USA

Correspondence to
Dr Michael Hendryx; hendryx@indiana.edu

ABSTRACT

Objective To test a positive deviance method to identify counties that are performing better than statistical expectations on a set of population health indicators.

Design Quantitative, cross-sectional county-level secondary analysis of risk variables and outcomes in Indiana. Data are analysed using multiple linear regression to identify counties performing better or worse than expected given traditional risk indicators, with a focus on 'positive deviants' or counties performing better than expected.

Participants Counties in Indiana (n=92) constitute the unit of analysis.

Main outcome measures Per cent adult obesity, per cent fair/poor health, low birth weight per cent, per cent with diabetes, years of potential life lost, colorectal cancer incidence rate and circulatory disease mortality rate.

Results County performance that outperforms expectations is for the most part outcome specific. But there are a few counties that performed particularly well across most measures.

Conclusions The positive deviance approach provides a means for state and local public health departments to identify places that show better health outcomes despite demographic, social, economic or behavioural disadvantage. These places may serve as case studies or models for subsequent investigations to uncover best practices in the face of adversity and generalise effective approaches to other areas.

Strengths and limitations of this study

- ► The study offers an approach to identify places where health outcomes exceed expectations given traditional risk indicators.
- ► This information might be useful to identify characteristics of places that allow higher than expected performance so that others may learn and benefit from them.
- ► Limitations of the study include the cross-sectional and ecological design and the absence of direct data on what drives better than expected performance.

INTRODUCTION

'Positive deviance' has been viewed as a characteristic of individuals that allow them to achieve better health outcomes compared with other individuals with similar risk profiles.[1–4] Persons facing social or economic obstacles to good health may nevertheless practise health-promoting behaviours when those behaviours are viewed as acceptable, affordable, effective and consistent with local culture.[5–7] In this study, we extend the idea of positive deviance from characteristics of individuals to those of communities[8] and hypothesise that communities may also possess collective attributes that allow them to outperform expectations given traditional risk indicators.

The majority of research on positive deviance has focused on characteristics of individuals that allow them to outperform expectations. For example, families living in impoverished conditions in Vietnam who nevertheless had children with good nutritional status were found to engage in certain behaviours and food choices that promoted better nutrition.[5] Interventions were subsequently developed and tested to promote the adoption of these healthy practices in other households (see also Bisets Bullen[9] for a review of child nutrition interventions based on positive deviance approaches). Another study, conducted in a largely Hispanic community in Texas at high risk for childhood obesity, found unique feeding behaviours and healthy habits among parents with normal weight children.[1]

However, it may be the case that features of communities can influence collective personal decisions about health behaviours, leading to positive deviant outcomes at the level of the group. Research on community-level positive deviance is sparse. Zullig *et al*[10] reported the results of a national county-level analysis of health-related quality of life and used spatial analysis techniques to identify county clusters that were underperforming or overperforming statistical expectations. Rust *et al*[7] have argued that attention should be given to identifying the critical elements present in some communities that allow them to achieve greater health equity under conditions of poverty or social vulnerability. They cite the well-known case of the sudden and unexpected elimination of an African American disparity in infant mortality outcomes in Dane County, Wisconsin, a phenomenon occurring at the level of the county. They also report results from a county-level analysis in Georgia showing that some counties with major social inequalities had paradoxically equitable health outcomes. Klaiman *et al*[11] used health expenditure data to identify Local Health Departments in Florida, New York and Washington that outperformed their peers on maternal and child health indicators. They found that contextual variables, such as type of community and funding, do not fully explain the positive deviant health departments and suggest further research into broader mechanisms, community partnerships and engagement as possible strategies that foster positive deviance.

Overall, it is not clear what forces may lead people in some communities collectively to pursue these positive approaches while other communities do not, although a number of theoretical approaches to understanding health behaviour at the level of the group or community (eg, social learning theory, theory of triadic influence)[12] may be useful to generate and perhaps test possible explanations. Differences in safety net programme, public policies, financial resources or individual leadership may also play roles, although evidence on these points is limited. The question arises as to whether there may be general characteristics (social, cultural, historical, geographic, etc) of counties or other aggregate units that allow them to outperform expectations across multiple health dimensions or whether positive deviance characteristics are specific to outcome.

In the current study, we examine a set of health outcome measures and a set of 'traditional' risk indicators (eg, age, race, education, smoking, obesity, etc) of those outcomes to identify counties in one state, Indiana, performing statistically better an expected given those traditional risk indicators. We test the hypothesis that the same counties will be identified as top performers across multiple health outcomes. If this hypothesis is supported, it provides evidence that there are general forces within counties that enable them to outperform expectations that is unrecognised but general characteristics that lead to better outcomes. If the hypothesis is not supported, it suggests that high-performing counties are outcome specific and that unique features may be operating in a county for a specific outcome that lead to its success on that particular measure. In either case, we may identify characteristics within counties that are useful for other counties to model and adapt for local purposes, targeted either to a specific outcome or multiple outcomes.

METHODS

Design

The study is a non-experimental, correlational investigation of the associations between a set of health outcome dependent variables and a set of social, economic, behavioural and demographic independent variables, measured at the county level for the state of Indiana. The study included anonymous, county-level aggregated data and university ethics review was not required.

Data

The choice of dependent variables was intended to represent a range of important and common health indicators. Likewise, the choice of independent variables was based on available data but was intended to represent many traditional risk indicators that have been studied in relation to these outcomes.[13–19] To obtain the data, we downloaded the most recent (2015) County Health Ranking Data (http://www.countyhealthrankings.org/.) The variables in this dataset are themselves collected from a variety of sources and sometimes represent different years. All of the variables represent population percentages or per capita rates. We also downloaded National Cancer Institute data on total age-adjusted colorectal cancer incidence rates for Indiana counties (http://statecancerprofiles.cancer.gov/incidencerates/index.php.). We chose colorectal cancer among cancer types because it is relatively common and reducible through prevention efforts. Finally, we downloaded Centers for Disease Control and Prevention (CDC) age-adjusted mortality rate data for all circulatory diseases for Indiana counties (http://wonder.cdc.gov/ucd-icd10.html.). From these data sources we extracted the variables provided in table 1.

Analysis

Indiana has 92 counties. Therefore, it was necessary to identify from the long list of potential independent variables a smaller number to include in regression models. We examined bivariate Pearson correlations and eliminated housing quality, food insecurity, household income and college education because they were highly correlated (>0.65) with one of the other independent variables. Specifically, low housing quality correlated with income inequality (r=0.67); food insecurity correlated with childhood poverty (r=0.77); household income correlated with childhood poverty (r=-0.83) and college education correlated with health uninsurance rate (r=-0.67). Next, we ran multiple linear regression models for each of the seven dependent variables using a stepwise inclusion approach so that only variables related to the dependent variable at $p<0.15$ were retained. This approach also allows us to individualise covariates from the larger set for each dependent variable based on empirical evidence. Using the selected variables for each dependent variable, we found for each county the predicted outcome, the actual outcome and the difference between predicted and actual (the residual). Across the seven models,

Table 1 List of study variables

Dependent variables	Years covered
Low birthweight delivery rate (per cent of live births <2500 g)	2006–2012
Diabetes prevalence rate	2011
Adult obesity rate	2011
Per cent of adults reporting fair or poor health (vs excellent, very good or good health)	2006–2012
Premature mortality (years of potential life lost before age 75 per 100 000 age-adjusted population)	2010–2012
Circulatory disease mortality rates per 100 000 age adjusted	2009–2013
Colorectal cancer incidence rates	2010–2014
Independent variables	**Years covered**
Adult smoking rate	2006–2012
Teen birth rate (number of births per 1000 young women aged 15–19)	2006–2012
Per cent without health insurance	2012
Supply of primary care physicians (ratio of population to primary care physicians)	2012
Supply of dentists (ratio of population to dentists)	2013
Mental health providers (ratio of population to mental health providers)	2014
High school graduation rate	2011–2012
Per cent adults with at least some college	2009–2013
Unemployment rate	2013
Median household income	2013
Child poverty rate	2013
Income inequality (ratio of household income at the 80th percentile to income at the 20th percentile	2009–2013
Adult obesity rate (used as an independent variable in the other six outcome models)	2011
Food insecurity (per cent who do not have access to reliable food source)	2012
Healthy food access (per cent of population that is low income and does not live close to a grocery store)	2010
Housing quality (per cent of households with one or more of: lacking complete kitchen facilities, complete plumbing facilities, severely overcrowded, severely cost burdened)	2007–2011
Social associations (number of associations per 10 000 population)	2012
Per cent population 65 and over	2013
Per cent population less than 18	2013
Per cent non-Hispanic African American	2013
Per cent Hispanic	2013
Per cent Asian	2013
Per cent non-Hispanic White	2013
Per cent female	2013
Per cent rural	2010

we examined the independent variables to determine whether there were common indicators across outcomes. We also sorted the counties by their residuals and examined the resulting lists to investigate whether the same counties were present at the extremes across the multiple outcomes. To quantify the level of agreement in the residuals, we found the Spearman rank correlations among the residuals, scoring the residuals in two alternative specifications. The first specification used the residual scores themselves. The second specification was based on standardising the residuals and categorising them into scores from 0 to 2, where 2=scored >1 SD above the mean, 1=midrange and 0=scored >1 SD below the mean; that is, each county received a score of 0, 1 or 2 for each outcome based on its residual. The majority of counties will be classified as 'normal' using this standardised approach, so that deviants on either side of the distribution will be fewer in number and stand out more clearly. The focus in this paper is on positive deviance counties, although some results are also presented for counties performing worse than expected.

RESULTS

A summary of the county statistics (mean, SD and range) for the dependent variables is summarised in table 2. This table also provides summary statistics for the selected independent variables that were significant in one or more of the regression models. The last two columns of this table provide the SD and range for the residuals that were obtained from the regression models; for example, the mean county-level obesity rate was 31.8% and the regression models underpredicted or overpredicted obesity rates within a range of −5.4% to 5.0%.

Table 3 provides a summary of the final regression models for the seven dependent variables. One variable, the teen birth rate, was included as an important risk for six of the seven outcomes. Hispanic population was an important indicator for five models, four of which indicated favourable outcomes and one unfavourable. Child poverty rates was an important risk in four models. (Finding teen birth rate in six of the seven models was unexpected, and we conducted an additional exploratory regression analysis using teen birth rate as the dependent variable. The final model (R^1=0.50, F (df=79, 8)=10.0, p<0.0001) indicated that higher teen birth rates were associated with higher per cent of the population aged <18, higher child poverty, less high school education, higher unemployment, fewer primary care physicians, higher obesity rates, urban locations and smaller per cent African American populations.)

Across models, the mean Spearman rank correlation among the residuals was only r=0.10 (NS). County performance ranks on one outcome were weakly and non-significantly related to county performance ranks on other outcomes. After standardising the residuals and categorising them from 0 to 2, the Spearman r among the scores remained poor (r=0.14 NS). Finally, we found for

Table 2 Summary statistics for dependent and selected independent variables, Indiana counties

Dependent variables	Mean	SD (observed)	Range (observed)	SD (residual)	Range (residual)
Per cent adult obesity	31.8	2.8	22.8–37.7	2.3	–5.4–5.0
Per cent fair/poor health	17.2	4.1	7.2–28.1	3.0	–6.4–7.7
Low birth weight per cent	7.6	1.0	5.1–9.8	0.7	–1.6–2.6
Per cent with diabetes	11.3	1.4	7.8–15.6	1.0	–1.7–3.0
Years potential life lost	7829.0	1453.3	3931.2–13 579.4	825.8	–2079.0–3326.5
Colorectal cancer incidence rate	44.8	7.4	32.4–68.0	6.3	–11.33–22.96
Circulatory disease mortality rate	259.6	34.7	152.7–397.5	27.3	–96.1–109.9
Independent variables					
Percent Hispanic	3.7	3.3	0.9–17.8		
Percent Black	2.7	4.4	0.2–26.9		
Percent Asian	0.9	1.2	0.2–6.8		
Percent female	50.2	1.1	45.8–53.2		
Percent population<18	23.6	2.4	16.0–33.8		
Percent population age>65	15.7	2.0	9.9–20.4		
Teen birth rate	39.2	9.9	12.6–62.8		
Per cent adult smokers	24.0	4.9	12.4–41.7		
Per cent adult obesity	31.8	2.8	22.8–37.7		
Income inequality	4.01	0.5	3.16–6.47		
Per cent children in poverty	20.6	5.7	6.1–33.3		
Unemployment rate	7.6	1.2	5.3–10.6		
Primary care physicians	50.1	28.1	7.1–193.4		
Mental health providers	80.2	69.0	7.1–486.1		
Dentists	36.3	16.3	7.1–81.0		
Per cent uninsured	16.4	2.5	9.5–28.1		
Per cent rural	54.5	26.9	0.6–100		
Per cent with limited access to healthy food	4.3	3.0	0–10.1		
Per cent with high school education	88.6	4.7	72.3–97.5		

Table 3 Results of linear multiple regression models

Independent variables	Dependent variables: Obesity rate	Per cent fair/poor health	Low birth weight per cent	Diabetes rate	Premature mortality	Colorectal cancer incidence rate	Circulatory disease mortality rate
Per cent Hispanic	0.16 (0.08)*	−0.22 (0.10)*	—	−0.07 (.04)	−5162 (2902)	—	−192.2 (85.6)*
Per cent Black	—	—	0.05 (.03)	0.08 (0.03)*	—	—	—
Per cent Asian	0−0.45 (0.27)	—	—	—	—	—	—
Per cent female	—	—	—	0.27 (0.10)*	12 349 (8278)	—	—
Per cent population<18	—	—	—	—	−10044 (4257)*	—	—
Per cent population age>65	0.46 (0.17)*	—	0.12 (0.06)*	0.33 (0.07)*	—	—	—
Teen birth rate	0.0006 (0.0003)*	0.002 (0.0004)*	0.0002 (0.0001)*	0.0003 (0.0001)*	72 (11)*	--	1.2 (0.3)*
Per cent adult smokers	—	0.31 (0.08)*	—	—	—	—	—
Per cent adult obesity	—	—	—	—	—	—	211.9 (113.2)
Income inequality	—	0.02 (0.008)*	—	—	—	—	—
Per cent children in poverty	—	0.16 (0.08)	0.04 (0.02)	—	6869 (2072)*	82.3 (17.5)*	—
Unemployment rate	—	−1.09 (0.36)*	—	—	23 786 (9717)*	−179.5 (74.0)*	702.1 (271.2)*
Primary care physicians	—	—	—	−0.0001 (0.00005)*	—	−0.06 (0.03)	—
Mental health providers	—	−0.0001 (0.00006)*	—	−0.00004 (0.00002)*	−3.0 (1.6)	—	—
Dentists	—	—	—	—	—		—
Per cent uninsured	—	—	−0.12 (0.04)*	—	—	−99.7 (36.1)*	—
Per cent rural	—	—	−0.01 (0.005)*	—	—	—	—
Per cent with limited access to healthy food	—	—	—	—	−7659 (3228)*	−74.2 (25.0)*	—
High school education	—	—	—	—	—		—
Model R2	0.34	0.51	0.49	0.55	0.72	0.23	0.39

All variables significant at p<0.15. All models have significant overall F value, p<0.0001. Cell values are unstandardised coefficients and standard errors.
*p<.05

Table 4 Selected demographic characteristics of the state and of top outperforming counties

	State	Outperforming					
		Parke	Brown	DeKalb	Elkhart	Union	Warren
Population 2014	6.6 m	17 233	14 962	42 383	201 971	7246	8352
Population per cent change 2010 to 2014	1.7	−0.7	−1.8	0.4	2.2	–3.6	−1.8
Per cent non-Hispanic White	80.3	94.7	95.9	95.1	75.9	95.9	97.0
Per cent college degree or higher	23.6	14.3	21.2	17.0	17.9	19.3	17.2
Per cent of population aged 16+ in labour force 2010–2014	64.0	52.9	62.2	64.4	66.5	64.5	62.5
Per cent in poverty	15.2	17.5	12.9	10.1	13.9	13.8	9.9
Per cent rural	27.4	75.0	100	42.3	20.1	100	77.1

each county the sum of the standardised scores across outcomes, with a possible range from 0 for the poorest performance across measures to 14 for the best possible performance. The mean score was 7.2 (SD=1.9). The highest score was 11 out of 14 for Parke County and Union County; the next highest score was 10 for four counties (Brown, DeKalb, Elkhart and Warren). The lowest score was two for Delaware County followed by three for Crawford County and Lawrence County. The sum score results suggest that most counties scored in the midrange, and there were very few counties that were positive deviants across measures generally.

Table 4 shows selected basic demographic characteristics of the top counties drawn from US Census data; note that table 4 figures are population averages while the figures in table 2 are county averages that weigh each county equally regardless of population. There are no obvious patterns among these variables that distinguish top performers.

Table 5 summarises the residuals and county ranks across models for Parke and Delaware Counties to show scores for selected counties with the most extreme positive and negative performance. Although Parke County and Union County received the top score across measures, they were not the highest scoring positive deviants on any of the seven individual measures. That distinction belonged to the following counties: Wayne (obesity); Brown (fair/poor health); Howard (low birth weight); Starke (diabetes); Ohio (premature mortality); Lawrence (colorectal cancer incidence) and Warren (circulatory disease mortality) (results not shown).

DISCUSSION

The analyses did not support the hypothesis that there would be strong, generalised non-traditional indicators that drive outcomes across the selected measures at the county level in Indiana. In retrospect, this may not be surprising given the variability in disease aetiology from one outcome to another. County performance that was unusually high was for the most part outcome specific. That performance was often outcome specific suggests that efforts to learn and disseminate best practices should also be outcome specific. Nevertheless, there were a few counties that seemed to be doing particularly well (Parke, Union) across most measures. A focus on teen birth rate as a general risk indicator is also potentially useful.

Some of the findings in table 3 seem counterintuitive or unexpected. Smoking rates, for example, measured at the level of the county, were not related to premature mortality or cancer incidence when other covariates were considered. Rates of health uninsurance were identified in two models in directions opposite to what one might suppose. This is a limitation of measures at the county level,but also should be interpreted with respect to other variables considered simultaneously; smoking at the county level, for example, may partially overlap with other risks such as education or poverty rates.

The appearance of teen birth rates as a significant independent variable in most models was also unexpected; this finding should not be interpreted as a direct causal link between teen birth rates and selected outcomes, but likely indicates that teen birth rates act as a proxy to other unmeasured risk conditions that have general health consequences. It may prove useful to investigate the conditions that promote lower teen birth rate outcomes and the extent to which successful interventions to reduce teen birth rates may or may not result in improved health conditions more generally. Teen birth rates have been in decline in the USA.[20] Studies of teen birth rates suggest national and state-level forces that contribute to higher or lower rates,[21 22] but studies of county or subcounty influences are less well developed.

Limitations of the study include the cross-sectional data, and the fact that some years for examining relationships between independent and dependent variables are not ideally ordered in time. Although the intent was to examine collective outcomes, the county-level, ecological data prohibit causal inference about individual behaviours in relation to independent variables. The study is also limited by the selected set of dependent and independent measures and is limited to one state over one time. The counterintuitive findings regarding smoking rates may reflect data limitations as well: these are self-reported data over a 7-year period which may contain inaccuracies,

Table 5 Observed, predicted and residual values for two extreme counties, Parke and Delaware

	Parke County	Delaware County
Per cent adult obesity		
Observed value	28.6	33.4
Predicted value	32.3	30.5
Residual	−0.3.7	2.9
Rank	4	86
Per cent fair/poor health		
Observed value	14.8	20.1
Predicted value	19.3	15.7
Residual	−4.5	4.4
Rank	3	83
Per cent low birth weight		
Observed value	7.1	9.3
Predicted value	7.4	8.2
Residual	−0.3	1.1
Rank	32	88
Per cent diabetes		
Observed value	11.7	11.8
Predicted value	12.8	10.7
Residual	−1.1	1.1
Rank	9	78
Years potential life lost (premature mortality)		
Observed value	7685	8839
Predicted value	9342	7627
Residual	−1658	1212
Rank	4	89
Colorectal cancer incidence rate		
Observed value	42.0	39.3
Predicted value	48.0	44.0
Residual	−6.0	−4.7
Rank	13	25
Circulatory disease mortality rate		
Observed value	276.3	255.0
Predicted value	263.0	257.0
Residual	13.3	−2.0
Rank	68	48

Also shown are the county ranks (1=best and 92=worst) for these counties.

and the rates for some counties had wide CIs which limits the ability to distinguish differences between counties. Finally, counties were used as the unit of analysis due to data availability; they are a somewhat arbitrary geographic unit and may fail to represent important subcounty variation from community to community, although the presence of county-level local health departments that exist in Indiana and many other states may provide a means in subsequent investigations for understanding health department impacts and for possible leveraging through county health department interventions.

It is important to keep in mind that the use of residuals to identify positive deviants identify counties that are performing outside of expectations given a set of traditional risk indicators. These are not necessarily the same counties that have the top observed values on the dependent variables. To cite one example, Monroe County had one of the lowest diabetes rates in the state at 8.5%, but it was predicted to have a rate of 8.46% and so it is not a positive deviant example, only a county performing as expected given the assessed risk indicators.

We suggest moving forward on the basis that outlier performance is not random but is based at least to some extent on unobserved forces that determine health outcomes. Important next steps include efforts to identify and understand the forces operating in these counties that are driving the outcomes. Qualitative studies to examine these forces may be especially useful,[23] and Canavan *et al*[24] provide an example of how positive deviant counties on adult obesity were identified, and qualitative interviews subsequently conducted to understand 'themes and strategies' that may drive the positive results. The approach employed in this study provides a technique to identify where to look and indicates that subsequent investigations in many cases will be outcome specific.

That being said, Parke County offers a good choice for focusing on a single county where performance was often better than expected. Among the state's 92 counties, this county ranked in the top 13 or better on 5 of the 7 measures. Parke County is a rural county in western Indiana near the Illinois border. It is known as the 'covered bridge capital of the world' and is a popular sightseeing place for visitors from Indianapolis, Chicago and other places. Selected Parke County demographic figures may be found in table 4. This county achieved frequent positive results despite relatively poor scores on education, labour force participation and poverty. Targeted follow-up qualitative studies to explore what may be happening here to promote positive health outcomes could be useful.

The methodological approach that we have described here is similar to other recently reported efforts.[11 25] Rust *et al*[26] used a different method to measure differences over time in county trends in colorectal cancer mortality to identify counties with reductions in race-related disparities. It may similarly be useful to attempt to replicate our findings with additional data points over time to investigate consistency of positive deviant counties. The findings in the current study suggest novel follow-up steps for local and state public health professionals to pursue. The typical evaluation of local public health indicators undertaken by state level officials involves direct identification of places where an outcome of interest is unfavourable, followed by efforts to improve that outcome. We do not suggest to replace these improvement efforts, but rather that state authorities can use the positive deviance approach as an additional method to identify role model counties where public health practices might be studied and perhaps replicated. In our results, this suggests that Parke or Union County might be investigated to understand positive performance across measures, that Starke County might be investigated to understand how it addresses diabetes or that Ohio County can be investigated to understand what makes it stand out with regard to lower than expected premature mortality. This does not guarantee that powerful new insights will be identified in these counties, but the possibility is there to generalise positive benefits to other places. Similarly, county health departments can try to understand those cases where their performance is particularly strong, or less than ideal given adjustments for traditional risk indicators, to improve areas of weakness or to tout accomplishments in areas where others may also learn and benefit.

Contributors MH led the design, analysis, interpretation and writing. LG-R and BH contributed to design, writing, interpretation, literature reviews and analysis. All other authors contributed to study conceptual design, writing and interpretation of findings.

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data available.

REFERENCES

1. Foster BA, Farragher J, Parker P, *et al*. A positive deviance approach to early childhood obesity: cross-sectional characterization of positive outliers. *Child Obes* 2015;11:281–8.
2. Tucker M, Harris GEAlcohol use among university students: considering a positive deviance approach.*J Health Psychol* 2016;21:1918–27.
3. Walker LO, Sterling BS, Hoke MM, *et al*. Applying the concept of positive deviance to public health data: a tool for reducing health disparities. *Public Health Nurs* 2007;24:571–6.
4. Wallace ME, Harville EW. Predictors of healthy birth outcome in adolescents: a positive deviance approach. *J Pediatr Adolesc Gynecol* 2012;25:314–21.
5. Mackintosh UA, Marsh DR, Schroeder DG. Sustained positive deviant child care practices and their effects on child growth in viet nam. *Food Nutr Bull* 2002;23:16–25.
6. Marsh DR, Schroeder DG, Dearden KA, *et al*. The power of positive deviance. *BMJ* 2004;329:1177–9.
7. Rust G, Levine RS, Fry-Johnson Y, *et al*. Paths to success: optimal and equitable health outcomes for all. *J Health Care Poor Underserved* 2012;23:7–19.
8. Rose AJ, McCullough MB. A practical guide to using the postive deviance method in health services research. *Health Serv Res* 2017;52:1207–22.
9. Bisits Bullen PA. The positive deviance/hearth approach to reducing child malnutrition: systematic review. *Trop Med Int Health* 2011;16:1354–66.
10. Zullig K, Hendryx M, Conley J. Geographic variation in the health-related quality of life (HRQOL) of US counties. American Public Health Association 142nd Annual Meeting. *New Orleans* 2014.
11. Klaiman T, Pantazis A, Chainani A. *et al*Using a positive deviance framework to identify local health departments in communities with exceptional maternal and child health outcomes: a cross sectional study,.*BMC Public Health* 2016;16:602.
12. Flay BR, Snyder F, Petraitis J. The Theory of Triadic Influence. In: DiClemente RJ, Kegler MC, Crosby RA, eds. *Emerging Theories in Health Promotion Practice and Research*. New York: Jossey-Bass. 2nd ed, 2009.
13. Chang T, Choi H, Richardson CR, *et al*. Implications of teen birth for overweight and obesity in adulthood. *Am J Obstet Gynecol* 2013;209:e1–7.
14. Franks PW, McCarthy MI. Exposing the exposures responsible for type 2 diabetes and obesity. *Science* 2016;354:69–73.
15. Gellert C, Schöttker B, Brenner H. Smoking and all-cause mortality in older people: systematic review and meta-analysis. *Arch Intern Med* 2012;172:837–44.
16. Remington PL, Catlin BB, Gennuso KP. The county health rankings: rationale and methods. *Popul Health Metr* 2015;13:11.
17. Starfield B, Shi L, Macinko J. Contribution of primary care to health systems and health. *Milbank Q* 2005;83:457–502.
18. Woolhandler S, Himmelstein DU. The relationshipm of health insurance and mortality: islack of insurance deadly? *Ann Intern Med* 2017:10.7326/M17-1403. [Epub ahead of print 27 June 2017].
19. Zonderman AB, Mode NA, Ejiogu N, *et al*. Race and poverty status as a risk for overallmortality in community-dwelling middle-aged adults. *JAMA Intern Med* 2016;176:1394–5.
20. Hamilton BE, Rossen LM, Branum AM. Teen birth rates for urban and rural areas in theUnited States 2017. *NCHS Data Brief* 2016;264:1–8.
21. Beltz MA, Sacks VH, Moore KA, *et al*. State policy and teen childbearing: a review of research studies. *J Adolesc Health* 2015;56:130–8.
22. Kearney MS, Levine PB. Investigating recent trends in the U.S teen birth rate. *J Health Econ* 2015;41:15–29.
23. Stoto MA, Nelson CD, Klaiman T. Getting from what to why: using qualitative methods in public health systems research. *AcademyHealth* 2013.
24. Canavan ME, Cherlin E, Boegeman S, *et al*. Community factors related to healthy eating & active living in counties with lower than expected adult obesity rates. *BMC Obes* 2016;3:49.
25. Klaiman TA, Pantazis A, Bekemeier B. A method for identifying positive deviant local health departments in maternal and child health. *Front Health Serv System Res*. 2014.
26. Rust G, Zhang S, Yu Z, *et al*. Counties eliminating racial disparities in colorectal cancer mortality. *Cancer* 2016;122:1735–48.