Milo Schield, University of New Mexico

Abstract: The choice of a denominator can change the size and direction of an association of rates. It can change the direction over time; it can change the order or direction of group rates; it can change the size of group prevalences. Examples presented include comparing Covid deaths per capita vs per case, the prevalence of men and blacks among those arrested versus among those in the population. Statistical disparities don't prove the presence or absence of discrimination. Statistical disparities are evidence in making such conclusions.

INTRODUCTION:

Three of the most basic questions one can ask about a statistic are these[1]:

1. How big (how many, how much)?
2. Compared to what?
3. Per what?

This paper deals with the third question: "per what?"

We know that a comparison of counts can be reversed by converting the counts into ratios. For example, in June 2021 there were more unemployed people in the US (9.5 million)[2] than in South Africa (7.2 million)[3]. But the unemployment rate was greater in South Africa (32.6%) than in the US (5.4%).[4]

We know that changing the denominator can change the size of a ratio. In 2019, the US general fertility rate was 58 per 1,000 women ages 15-44.[5] The birth rate per capita was 11 per 1,000 population.

This paper makes a different point. The choice of the denominator can change the size and direction of an association of rates:

1. A change in direction over time.
2. A change in direction between group rates
3. A change in the size of a difference or disparity

The choice of a denominator is typically the user's choice. That choice can be innocent or naive. But it can also be diabolical: evil, wicked or cruel. Either way, the choice of the denominator has statistical consequences.

1. CHANGE IN DIRECTION OVER TIME

Consider the left side of Figure 1. Per household, microwave expense per year increased by 43%. But per household reporting a microwave expenditure, microwave expense per year decreased by 48%.

Table 1: Compare cost of microwave ovens and babysitting in 1987 versus 1980 per HH and per HH reporting

Microwave ($/year)	1980	1987	Change
per Household (HH)	$14	$20	43%
Percentage reporting	2.9%	8.0%	176%
per HH Reporting	$483	$250	-48%
www.bls.gov/opub/mlr/1992/05/art3full.pdf			

Baby sitting ($/year)	1980	1987	Change
per Household (HH)	$76	$75	-1%
Percentage reporting	6.8%	6.4%	-6%
per HH Reporting	$1,118	$1,172	5%
www.bls.gov/opub/mlr/1992/05/art3full.pdf			

[1] Schield, M. (2021). Statistical Literacy for Policy Makers. Copy at www.statlit.org/pdf/2021-Schield-ISI.pdf
[2] Accessed August 11, 2021 https://www.bls.gov/news.release.pdf
[3] https://www.news24.com/fin24/economy/sas-jobless-grows-to-72-million-as-unemployment-rate-breaches-new-record-20210223
[4] Accessed August 11, 2021 https://countryeconomy.com/unemployment
[5] www.cdc.gov/nchs/data/vsrr/vsrr-8-508.pdf

Now look at the right side. Per household, babysitting expense per year decreased by 1%. But per household reporting payments for babysitting, the rate increased by 5%.

So what explains this reversal? In both cases, it is the percentage reporting. The expense per household reporting always equals the expense per household divided by the fraction of households reporting.

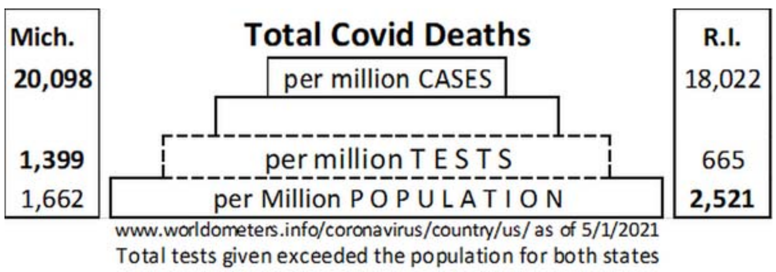
If the percentage reporting increases (left side), the expense per household reporting will decrease.
If the percentage reporting decreases (right side), the expense per household reporting will increase.

2. CHANGE IN DIRECTION BETWEEN GROUP RATES

2.1 COVID DEATH RATES BY STATE

Consider Covid deaths by state as of May 1, 2021. Per capita (per million population), Rhode Island (RI) had more Covid deaths than Michigan (MI).

Figure 1: Compare Covid Deaths in two States per capita, per test and per case



But per test and per case, Michigan had more Covid deaths than Rhode Island. Once again we have a reversal. Before it was a reversal over time. Now it is a reversal in rank or order.

What explains the reversal? The ratios between the denominators vary by state.

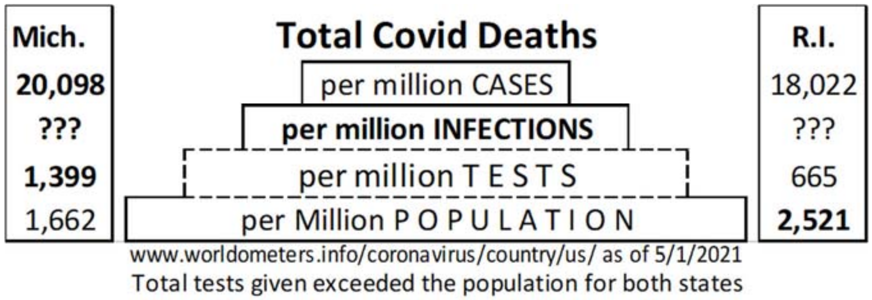
Figure 2: Covid Count and Ratio Data for Two States

---------- Covid Totals as of 5/01/2021 ----------					
Michigan	Counts	Rhode Island	Michigan	Ratios	Rhode Island
18,893	Deaths	2,671	1.35	Tests/Capita	3.79
940,175	Cases	148,186	14.37	Tests/Case	27.12
13,509,456	Tests	4,019,021	www.worldometers.info/coronavirus/country/us		
9,986,857	Population	1,059,361	Count data recorded on 5/1/2021 by Schield		

- Michigan had 1.4 times as many tests as people (Rhode Island at 3.8 times as many).
- Michigan at 14 times as many tests as cases (Rhode Island had 27 times as many).

Now we have a second question. Which denominator is better? All three comparisons are true. But one may be better – less misleading – than the other. Which denominator is better? Per capita, per test or per case? Maybe none of these. Consider a fourth denominator: per 1,000 infections.

Figure 3: Compare Covid Deaths in Two States per Capita, per Test, per Case and per Infection



Unfortunately we don't have US deaths per thousand infections. Does that mean it is worthless?

No. If you believe that infections is the best denominator, then you need to argue about whether per case, per test, or per capita is closer to per infection.

2.2 AUTO DEATH RATES BY STATE

Here is another example of a reversal in rank or order. Consider the 1996 auto death rate for two states: Arkansas (AR) and Hawaii (HI).

The 1996 auto-death rate was
- four times as high in HI (13) as in AR (3) *per 1,000 miles of road.* [See Appendix.]
- 50% higher in AR(2.4) than in HI (1.6) per 100 million vehicle miles[6]
- 75% higher in AR (35) than in HI (20) per 100,000 drivers
- Twice as high in AR (38) as in HI (19) *per 100,000 registered vehicles*. [7]
- Twice as high in AR (25) as in HI (13) per 100,000 population

Once again the rank or order is reversed by choosing a different denominator. What explains this reversal? The difference between the two states in the number of vehicles per mile of road.

To get the vehicles per mile of road, divide auto deaths per mile by auto deaths per vehicle.
- Take Hawaii. Divide the 35 deaths per 1,000 miles of road by the 19 deaths per 100,000 vehicles. That gives 184 vehicles per 100 miles of road for Hawaii.
- For Arkansas, we divided the 7 auto deaths per 1,000 miles of road by the 38 auto deaths per 100,000 vehicles. That gives 20 vehicles per 100 miles of road in Arkansas.

The vehicles per mile of road were about nine times as many in Hawaii as in Arkansas.

The two states have different relationships between the two variables: vehicles and miles of road. This difference is what creates the mixed-fruit comparison: the apples and oranges comparison.

To make this an apples and apples comparison, we need to take this difference into account (to control for it). For more on this see Schield (2021).

---

[6] US Statistical Abstract, 1999. Table 1042. Copy in Appendix. Last half of 1999 currently missing online.
[7] US Traffic Safety Facts, 1997. Table 108. Copy in Appendix.
https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/809100

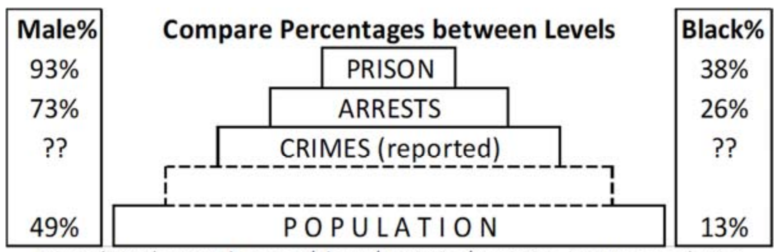
3. CHANGE IN SIZE OF ASSOCIATION BETWEEN GROUPS

The size of an association is often described as the effect size. Effect size is important in two ways. The bigger the size of the association,

1. the stronger support is gives for a disputable conclusion
2. the less likely it is to be influenced by confounders

Effect size can be influenced by confounders. Effect size can be influenced by the choice of the denominator.

Figure 4: Prevalence of Males and Blacks per capita, per arrest and per prisoner



Prison (2019): www.bop.gov/about/statistics/statistics_inmate_gender.jsp
Arrest M+F (2019): www.ojjdp.gov/ojstatbb/crime/ucr.asp?table_in=1

On the left side:
Men are 9 times as prevalent among prisoners as are women. Men are three times as prevalent among those arrested as are women. Whereas men are about as common as women in the population. These are big disparities.

Men are 73% of those arrested (49% of the population). Men are 93% of prisoners (49% of population).

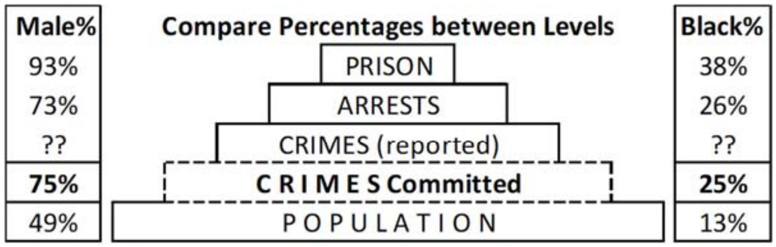
On the right side:
Blacks are 26% of those arrested (13% of the population). Blacks are 38% of those in prison (13% of the population).

Both of these are big disparities. Does this prove discrimination against men or against blacks in the criminal justice system? Maybe. But both of these comparisons can be influenced by the choice of the denominator.

What other denominator might be relevant? Crimes committed. If men are committing more crimes than women, then it makes sense that men are more likely to be arrested than women.

The police know about crimes reported, but not about crimes unreported. However the US DOJ conducts a victimization survey. They ask those surveyed whether they have been victimized and if so what was the sex and race of the victimizer. Figure 5 presents that data.

Figure 5: Prevalence of Males and Blacks per capita, per arrest, per prisoner and per victimization



2019: https://bjs.ojp.gov/library/publications/list?series_filter=Criminal%20Victimization

If we compare the prevalence of men among those arrested with the prevalence of men among those committing crimes, the gender disparity all but disappears. The same is true for the racial disparity.

Does this prove that the criminal justice system does not discriminate against men or blacks?

No. Statistical disparities – or their absence – don't prove the presence – or absence – of discrimination. They provide evidence.

The crimes committed data is self-reported by the victims. Maybe the victims are misremembering or giving a stereotypical answer when they weren't sure of the sex or race of the victimizer.

Maybe the criminals in one race commit more crimes than the criminals in the other race? To illustrate, suppose that all the crimes by whites were committed by one person. And suppose that person is never arrested. Then having 73% of the arrests involve someone white would be totally wrong.

Is this new denominator helpful? Yes, it may help focus the conversation on some of the underlying causes of what appear to be gender or racial disparities.

SUMMARY:

The choice of a denominator can be diabolical. It can change the size and direction of an association of rates. It can change the direction over time; it can change the order or direction of group rates; it can change the size of group prevalences. Examples presented include comparing Covid deaths per capita vs per case, the prevalence of men and blacks among those arrested versus among those in the population. Statistical disparities don't prove the presence or absence of discrimination. Statistical disparities are evidence in making such conclusions.

So, take CARE in moving from an association to causation: from disparity to discrimination. Many – if not most – such associations are crude associations. They may be true, but they don't take into account (they don't control for) related factors. Often there is a story behind the statistical disparity. These disparities can be confusing: they can be confounding or confounded. To learn more about teaching confounding, see Schield (2021)

Bibliography:
Schield, Milo (2021). Statistical Literacy: Teaching Confounding. USCOTS Workshop. Paper at www.StatLit.org/pdf/2021-Schield-USCOTS.pdf

APPENDIX A

Table 108 of the 1997 US Safety Facts

Table 108
Persons Killed, Licensed Drivers, Registered Vehicles, Population, and Fatality Rates by State, 1997

State	1997 Licensed Drivers (Thousands)	Fatalities per 100,000 Drivers	1997 Registered Vehicles (Thousands)	Fatalities per 100,000 Registered Vehicles	1997 Population (Thousands)	Fatalities per 100,000 Population	1997 Total Killed
AL	3,387	35.19	3,708	32.15	4,322	27.58	**1,192**
AK	446	17.26	556	13.85	610	12.62	**77**
AZ	3,120	30.48	3,218	29.55	4,553	20.89	**951**
AR	1,879	35.13	1,648	40.05	2,523	26.16	**660**
CA	20,385	18.09	25,399	14.52	32,182	11.46	**3,688**
CO	2,836	21.61	3,618	16.94	3,892	15.75	**613**
CT	2,270	14.93	2,708	12.52	3,267	10.38	**339**
DE	536	26.68	624	22.92	735	19.46	**143**
DC	356	16.85	235	25.53	530	11.32	**60**
FL	11,749	23.70	11,084	25.13	14,677	18.98	**2,785**
GA	5,063	31.15	6,318	24.96	7,490	21.05	**1,577**
HI	739	17.73	714	18.35	1,192	10.99	**131**
ID	844	30.69	1,116	23.21	1,209	21.42	**259**
IL	7,692	18.16	8,625	16.20	11,989	11.65	**1,397**
IN	3,924	23.83	5,444	17.17	5,865	15.94	**935**
IA	1,953	23.96	2,983	15.69	2,854	16.40	**468**
KS	1,825	26.41	2,200	21.91	2,601	18.53	**482**
KY	2,575	33.28	2,819	30.40	3,910	21.92	**857**
LA	2,678	34.76	3,449	26.99	4,354	21.38	**931**
ME	901	21.31	1,087	17.66	1,242	15.46	**192**
MD	3,347	18.26	3,825	15.97	5,095	11.99	**611**
MA	4,393	10.04	5,159	8.55	6,114	7.21	**441**
MI	6,751	21.42	8,179	17.68	9,780	14.79	**1,446**
MN	2,839	21.13	4,051	14.81	4,687	12.80	**600**
MS	1,723	49.97	2,265	38.01	2,732	31.52	**861**
MO	3,744	31.84	4,406	27.05	5,408	22.04	**1,192**
MT	662	40.03	1,001	26.47	879	30.15	**265**
NE	1,179	25.61	1,525	19.80	1,657	18.23	**302**
NV	1,186	29.26	1,169	29.68	1,679	20.67	**347**
NH	883	14.16	1,175	10.64	1,172	10.67	**125**

Note: 1998 data for state licensed drivers and registered vehicles not available at time of publication.

Arkansas (AR): 660 killed with 210,729 lane miles of road: 3.131 deaths per thousand lane miles of road.
Hawaii (HI): 131 killed with 9,799 lane miles of road: 13.369 deaths per thousand lane miles of road.

Table 1042 of the 1999 U.S. Statistical Abstract

No. 1042. Motor Vehicle Deaths, by State: 1990 to 1997

[Includes both traffic and nontraffic motor vehicle deaths. See source for definitions]

State	1990	1995	1996	1997	Mileage rate [2] 1990	Mileage rate [2] 1997	State	1990	1995	1996	1997	Mileage rate [2] 1990	Mileage rate [2] 1997
U.S. [1]	46,814	43,363	43,300	43,200	2.2	1.7	MO	1,174	1,110	1,148	1,192	2.3	2.0
AL	1,234	1,111	1,142	1,181	2.9	2.3	MT	225	215	198	265	2.7	2.8
AK	102	86	80	77	2.6	1.8	NE	289	254	293	302	2.1	1.9
AZ	947	1,040	993	961	2.7	2.4	NV	405	312	348	347	4.0	2.4
AR	625	631	615	660	3.0	2.4	NH	154	118	134	125	1.6	1.2
CA	5,411	4,165	3,972	3,377	2.1	1.2	NJ	908	776	818	(NA)	1.5	(NA)
CO	583	645	434	516	2.1	1.4	NM	534	485	481	484	3.3	2.2
CT	395	318	310	337	1.5	1.2	NY	2,318	1,668	1,562	1,625	2.2	1.4
DE	151	123	120	147	2.3	1.9	NC	1,489	1,438	1,492	1,484	2.4	1.9
DC	91	(NA)	(NA)	(NA)	2.7	(NA)	ND	128	74	85	105	2.2	1.6
FL	3,049	2,812	2,813	2,847	2.8	2.2	OH	1,708	1,357	1,393	1,439	2.0	1.4
GA	1,659	1,494	1,578	1,584	2.3	1.8	OK	684	674	775	842	2.1	2.1
HI	186	127	145	131	2.3	1.6	OR	608	572	524	521	2.3	1.7
ID	259	263	258	259	2.6	2.1	PA	1,767	1,480	1,470	1,562	2.1	1.6
IL	1,650	1,589	1,475	1,404	2.0	1.5	RI	104	69	69	75	1.5	1.1
IN	1,097	960	981	(NA)	2.0	(NA)	SC	987	882	930	903	2.9	2.3
IA	481	527	465	468	2.1	1.8	SD	166	158	175	148	2.4	1.9
KS	453	438	491	483	2.0	1.9	TN	1,312	1,240	1,211	1,223	2.8	2.1
KY	850	856	844	865	2.5	2.1	TX	3,381	3,172	3,738	3,476	2.1	1.9
LA	1,023	880	809	840	2.7	2.1	UT	296	321	321	367	2.0	1.9
ME	215	189	168	192	1.8	1.5	VT	94	106	88	96	1.6	1.5
MD	741	682	614	609	1.8	1.3	VA	1,091	900	869	981	1.8	1.4
MA	655	448	417	443	1.4	0.9	WA	875	654	690	659	2.0	1.3
MI	1,633	1,537	1,505	1,446	2.0	1.7	WV	502	376	344	372	3.3	2.1
MN	644	597	576	598	1.7	1.3	WI	825	739	759	721	1.9	1.4
MS	863	868	811	861	3.5	2.9	WY	130	170	143	137	2.2	1.9

NA Not available. [1] Source: National Center for Health Statistics. [2] Deaths per 100 million vehicle miles.

Source: Except as noted, National Safety Council, Itasca, IL, *Accident Facts*, annual (copyright).

Road Miles by State: Sorted from Most to Least

STATE	TOTAL LANE MILES
Texas	683,533
California	396,540
Illinois	306,658
Minnesota	290,618
Kansas	286,606
Missouri	277,504
Florida	275,376
Georgia	272,662
Ohio	262,492
Michigan	256,579
Pennsylvania	251,708
New York	240,489
Wisconsin	239,318
Oklahoma	238,754
Iowa	235,549
North Carolina	229,011
Arkansas	210,729
Alabama	210,531
Tennessee	203,850
Indiana	202,707
Nebraska	193,996
Colorado	185,486
North Dakota	178,845
Washington	167,632
Kentucky	166,971
South Dakota	166,635
South Carolina	166,594
Virginia	164,132
Oregon	162,101
Mississippi	162,088
Montana	150,446
New Mexico	150,216
Arizona	146,465
Louisiana	134,115
Idaho	107,568
Utah	102,031
Nevada	100,805
New Jersey	85,108
West Virginia	80,167
Massachusetts	77,730
Maryland	71,129
Wyoming	62,620
Maine	46,736
Connecticut	45,916
Alaska	36,009
New Hampshire	33,391
Vermont	29,273
Delaware	14,069
Rhode Island	12,664
Hawaii	9,799
DC	3,445

Source: https://blog.cubitplanning.com/2010/02/road-miles-by-state/