

# Comparing methods of grouping hospitals

Jingxuan Zhao

11/09/2022

# Motivation

- Grouping health care providers is critical to a range of health policy and health services research efforts.

# Study Question

- To compare the performance of approaches for defining groups of hospitals:  
HRRs, CBSAs and MSAs, CDA Communities

# Preview of findings

- Hospital groups defined by community detection algorithms (CDAs) were the most distinctive, were reliable to alternative specifications, and had greater generalizability than HRRs, MSAs, or CBSAs.
- Hospital referral regions (HRRs) performed reasonably well on several dimensions.

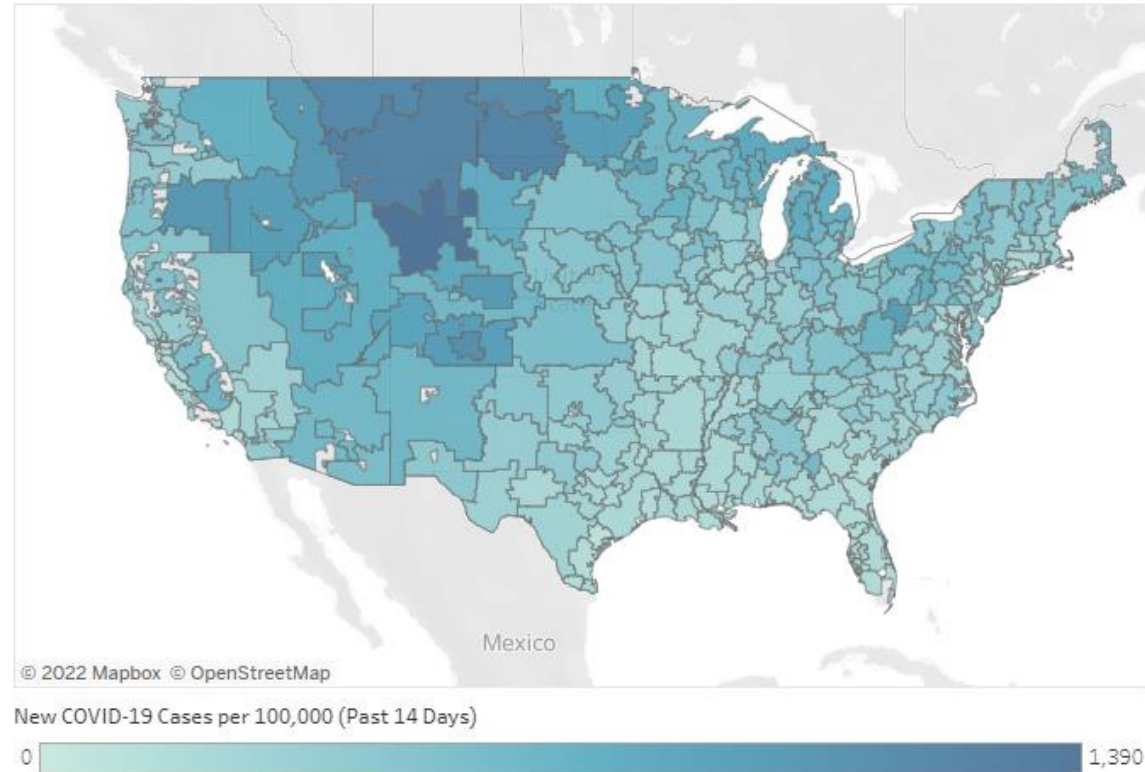
# Contributions

- Comparing different methods of grouping hospitals, providing information for researchers when they need to choose an approach to grouping hospitals.

# Hospital referral regions (HRRs)

- **Step1:** grouping at zip code level by which city the majority (>50%) of Medicare fee-for-service patients received hospital care at each zip code. (Hospital service areas, HSA)
- **Step2:** grouping at HAS level by which city the majority (>50%) of patients received major cardiovascular surgery and for neurosurgery at each HSA.
- Not updated since 1993.
- Patient travel patterns for specialty care.

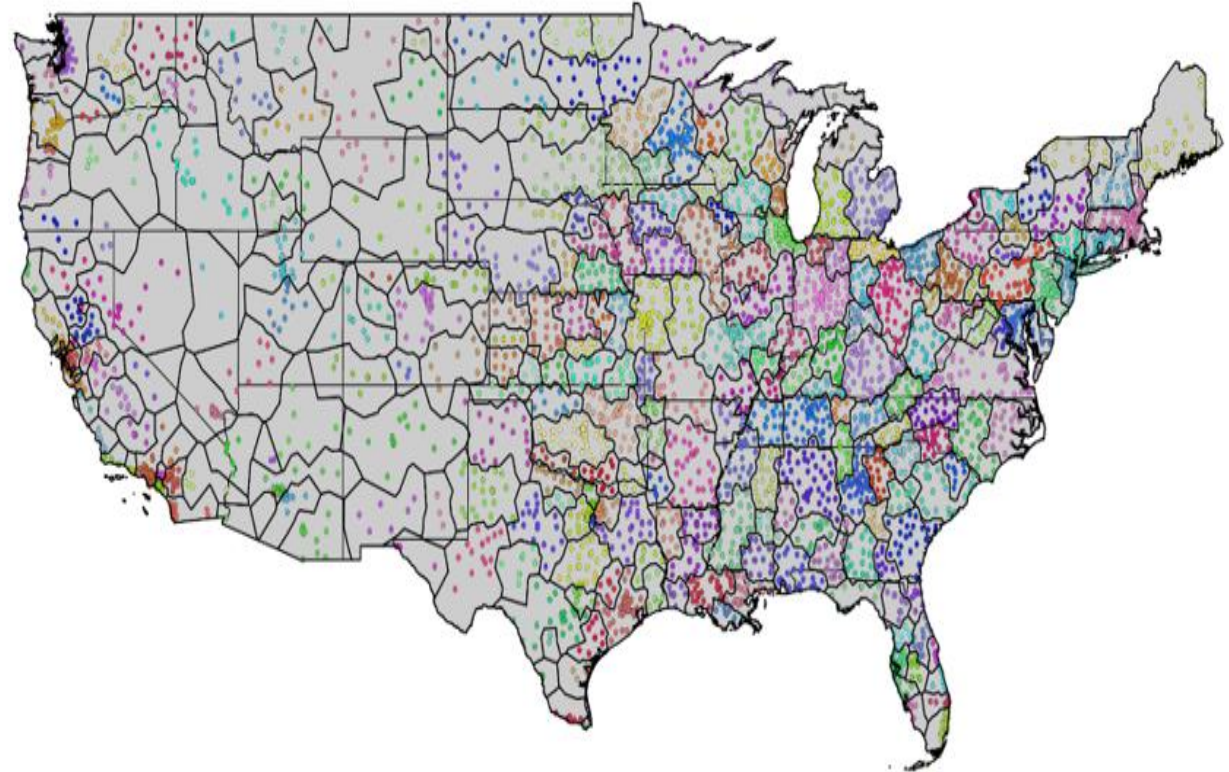
New COVID-19 Cases per 100,000 for the Past 14 Days (10/31/2021)  
by Hospital Referral Region (HRR)





# Community detection algorithms (CDA) communities

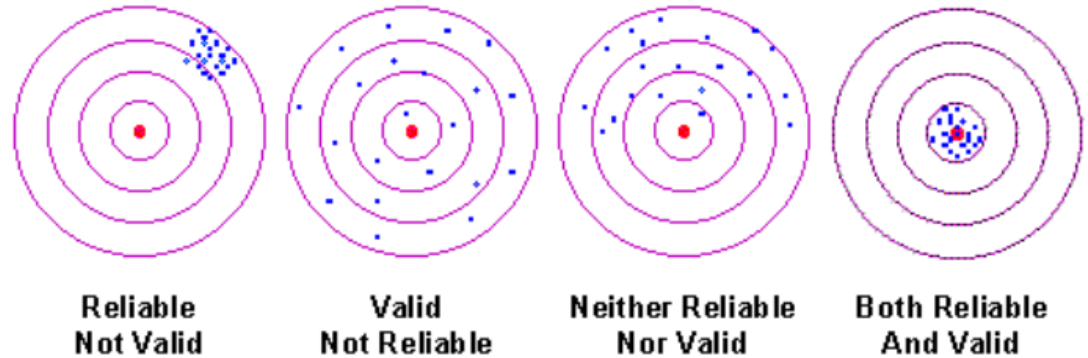
- Shared populations, referral patterns, unintentional interhospital travel patterns, and transfers.
- All fee-for-service Medicare patients
- Walktrap algorithm
- Maximizes the distinctiveness
- <https://www.youtube.com/watch?v=jIS5pZ8doH8>
- **Interhospital travel patterns**



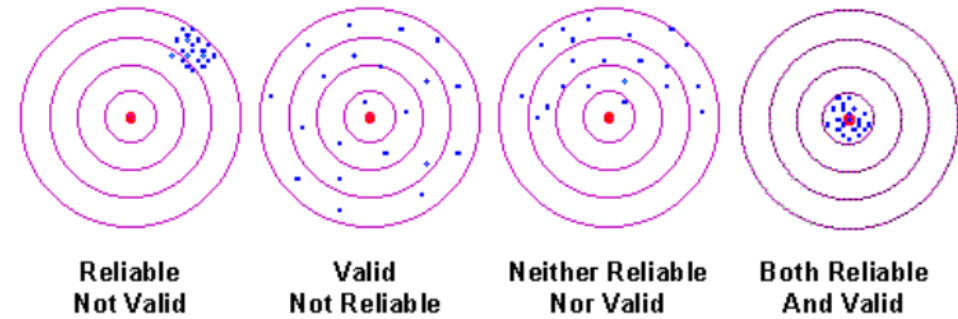


# Reliability, validity, and generalizability

- Reliability: reproducible— reliable health care groups should have stable membership over time and under varied assumptions about how to group hospitals.
- Validity: how well it reflects an underlying “true” value.
- Generalizability: the ability to extend measurement from the sample under study to a broader population



# Some additional notes



- Normalized mutual information (NMI): Reliability
  - The amount of joint information contained in group partitions
  - Ranges from a low of 0 (no mutual information) to a maximum of 1 (perfect correlation).
  - [https://course.ccs.neu.edu/cs6140sp15/7\\_locality\\_cluster/Assignment-6/NMI.pdf](https://course.ccs.neu.edu/cs6140sp15/7_locality_cluster/Assignment-6/NMI.pdf)
- Modularity: Validity
  - The proportion of patients that are shared within groups as opposed to between groups relative to what would be expected given the number of patients shared with hospitals in each group
  - Ranges from  $-1$  to  $+1$ , with 0 representing no better or worse than random. High modularity scores demonstrate high validity.
  - <https://www.cs.cmu.edu/~ckingsf/bioinfo-lectures/modularity.pdf>

# Data

- American Hospital Association (AHA) Annual Survey (2012-2014).
- 2007 and 2013 delineations of MSAs.
- 2015 Hospital Compare data.
- Physician Shared Patient Patterns (CMS)
  - All health care providers appearing on Medicare claims.
  - Providers who share common patients (providers appeared on a Medicare claim within 30 days)

# Reliability, validity, and generalizability

Measurement property	Criteria	Description
Reliability	Method-based reliability	Changes in methodology (eg, cutoff points) should not arbitrarily change hospital community membership.
Reliability	Reliability over time	Changes over time should not arbitrarily change hospital community membership.
Validity	Responsiveness to change	Communities should change to reflect changes in patient movement or risk becoming misleading over time.
Validity	Defined by patient travel patterns	Communities should be defined by patient travel patterns because they reflect referral patterns, and communication, competition for referrals and define coordination needs.
Validity	Highly distinct	Approaches in which hospitals within the community are highly connected, with few connections to outside hospitals, are more meaningful than approach that divide highly connected hospitals into separate communities.
Generalizability	Largest feasible population	Because community definitions are used in a range of research and policy applications, they should be based on the broadest possible population to reflect wide range of referral relationships and coordination needs, and should include as many hospitals as possible nationally.
Generalizability	Adaptable number of groups	Any delineation into a specific number of groups could be arbitrary; a hierarchical approach allows for division of communities into greater numbers.
Validity Extension	Members exhibit shared behaviors	Hospitals, and providers practicing at them, that are members of the same community should exhibit similar behavior because they learn from one another and develop communities of practice.

# Methodological reliability

- Changes in methodology (eg, cutoff points) should not arbitrarily change hospital community membership.

HRRs	Low. Referral patterns were largely based on major cardiovascular surgery and for neurosurgery.
MSAs and CBSAs	Low. Several numeric cut-offs in the definition
CDA Communities	Moderately high. A continuous measure of distance and all Medicare patients (rather than patients with specific diagnoses).

# Reliability over time

- Changes over time should not arbitrarily change hospital community membership.

HRRs	NA
MSAs and CBSAs	Moderately high. unlikely to be sensitive to change in health care travel patterns.
CDA Communities	Low?. NMI>0.93 for 2012-2014.

# Validity: Responsiveness to change in health care travel patterns

- Communities should change to reflect changes in patient movement.

HRRs	Low.
MSAs and CBSAs	Moderately high?. Unlikely to be sensitive to change in health care travel patterns.
CDA Communities	High. By including all fee-for-service Medicare patients, this method captures changes in patient travel patterns over time for a broader group of patients not limited to specific conditions.

# Validity: Distinctiveness

- hospitals within the community are highly connected, with few connections to outside hospitals.

**TABLE 1** Size and distinctiveness of grouping methods

	306 HRR	364 MSA	898 CBSA	266 CDA Communities	308 CDA Communities	863 CDA Communities
Average size	28	28	22	31	26	8
Minimum	2	1	1	2	2	2
Maximum	85	116	116	93	93	37
% of ties within	51.8	63.3	48.9	61.9	57.8	27.9
% of patients within	75.0	84.7	72.4	87.2	85.3	62.0
Modularity	0.75	0.83 <sup>a</sup>	0.72 <sup>b</sup>	0.86	0.84	0.63

- CDA>MSA>HRR>CBSA



# Generalizability: Inclusivity

- They should be inclusive of broad populations

HRRs	Moderately high. Only reflect referral patterns for the neurosurgical and cardiovascular surgery Medicare populations.
MSAs	Low. Cannot be applied to retired Medicare populations. Covers 56% hospitals.
CBSAs	Moderately low. Cannot be applied to retired Medicare populations. Covers 75% hospitals.
CDA Communities	High. All Medicare fee-for-service patients.

# Generalizability: Flexibility

- Offering flexibility in group sizes, such that the method can be generalized to the widest range of future analytic purposes

HRRs	Low. 3-level hierarchy with ZIP codes nested within HSAs nested within HRRs, hard to regroup.
MSAs and CBSAs	Low. They are hierarchical in that they are groups of counties.
CDA Communities	High. The communities can be split anywhere along the hierarchy.

# Similarity

- The resulting hospital groupings are sufficiently similar with these different approaches.

	266 CDA Communities	308 CDA Communities	863 CDA Communities	MSA	CBSA
HRR	0.88	0.88	0.86	0.92	0.88
266 CDA Communities		0.99	0.89	0.93	0.89
308 CDA Communities			0.9	0.93	0.9
863 CDA Communities				0.91	0.91
MSA					1

Note: Similarity measured by normalized mutual information (NMI). For MSA and CBSA, similarity is only assessed for hospitals that reside within the statistical area.

# Validity extension: Shared behavior

- Hospitals within same group share similar behaviors.
- Randomly splitting each community, HRR, and statistical area in half and testing the correlation between the mean scores of each half on five performance measures.
  - Medicare Spending Per Beneficiary
  - 30-day All-Cause Readmission Rates
  - Mammography Follow-Up Rates
  - MRI Lumbar Spine for Lower Back Pain
  - Total process scores

# Validity extension: Shared behavior

**Table A4. Split-Half Correlation Test of Similar Behavior**

	Median	Readmissions	MSPB	Mammography		Lumbar MRI		VBP Process			
	Performance			Correlation	Perf.	Correlation	Perf.	Correlation	Perf.	Score	
	Over Five	Correlation	Perf.	Correlation	Perf.	Correlation	Perf.	Correlation	Perf.		
	measures	Coefficient	Rank	Coefficient	Rank	Coefficient	Rank	Coefficient	Rank		
863 CDA Communities	1	0.5	1	0.7	1	0.27	4	0.13	6	0.11	1
266 CDA Communities	2	0.47	2	0.66	2	0.31	1	0.26	2	0.09	2
308 CDA Communities	3	0.37	4	0.63	3	0.29	3	0.22	4	0.07	3
HRR	3	0.42	3	0.62	4	0.31	2	0.24	3	0.02	5
MSA	5	0.29	6	0.59	6	0.26	5	0.14	5	0.04	4
CBSA	5	0.32	5	0.6	5	0.25	6	0.27	1	-0.02	6

The correlation coefficient reflects the similarity between the performance of two randomly selected halves of each Community, HRR and census area. Each grouping is then ranked by coefficient, with the greatest correlations receiving the lowest rank. The median of each grouping's rank is then determined to identify the overall success of the ranking method.

- CDA>HRR> MSA> CBSA

**TABLE 3** Summary of hospital grouping method performance on eight measurement properties

Measurement property	Criteria	Description	CDA	HRR	MSA	CBSA
Reliability	Method-based reliability	Changes in methodology (eg, cutoff points) should not arbitrarily change hospital community membership.				
Reliability	Reliability over time	Changes over time should not arbitrarily change hospital community membership.				
Validity	Responsiveness to change	Communities should change to reflect changes in patient movement or risk becoming misleading over time.				
Validity	Defined by patient travel patterns	Communities should be defined by patient travel patterns because they reflect referral patterns, and communication, competition for referrals and define coordination needs.				
Validity	Highly distinct	Approaches in which hospitals within the community are highly connected, with few connections to outside hospitals, are more meaningful than approach that divide highly connected hospitals into separate communities.				
Generalizability	Largest feasible population	Because community definitions are used in a range of research and policy applications, they should be based on the broadest possible population to reflect wide range of referral relationships and coordination needs, and should include as many hospitals as possible nationally.				
Generalizability	Adaptable number of groups	Any delineation into a specific number of groups could be arbitrary; a hierarchical approach allows for division of communities into greater numbers.				
Validity Extension	Members exhibit shared behaviors	Hospitals, and providers practicing at them, that are members of the same community should exhibit similar behavior because they learn from one another and develop communities of practice.				

- Quantitatively verified to have high performance, and, where comparisons are possible, the best performer of included approaches.
- Quantitatively tested with moderately high performance and 2nd best performer when comparison possible.
- Quantitatively tested with moderately low performance and 3rd best performer when comparison possible.
- Conceptual reasons to question performance but not quantitatively testable.
- Strong conceptual reasons to doubt performance, but not quantitatively testable.
- Quantitatively tested with low performance.

# Which approach to use?

- Reliability over time and responsiveness to change in health care travel patterns
  - Good reliability over time: make many analytic tasks easier. However, insensitive to changes in market conditions that should lead to differing groups (hospitals joining a multihospital system and changing their key referral partners)
- Large group or small group?
  - Large group: higher modularity.
  - Small group: more actionable, lower modularity (higher % of patients travel between groups).

# Reflections

- Providing a good framework for comparing different grouping approaches.
- Non-Medicare patients?
- Other hospital factors?