

# IDENTIFIKATION VON LEISTUNGSERBRINGER-CLUSTERN IN GKV-ROUTINEDATEN DURCH PATIENT SHARING NETWORKS

Dominik Domhoff, Kathrin Seibert, Susanne Stiefler, Dirk Peschke

AGENS-Methodenworkshop

Magdeburg, 28.02.2019

# Hintergrund

- Studie *Heimeintritt vermeiden – Beginn stationärer Langzeitpflege und seine Prädiktoren in der Versorgungs-, Wohn- und Unterstützungssituation*
- Ziel: Populationsbezogene Analyse zu Einflussfaktoren für Pflegeheimeintritt
- Population = Leistungserbringer (SGB V)

# Wie lassen sich Gruppen von Leistungserbringern voneinander abgrenzen?

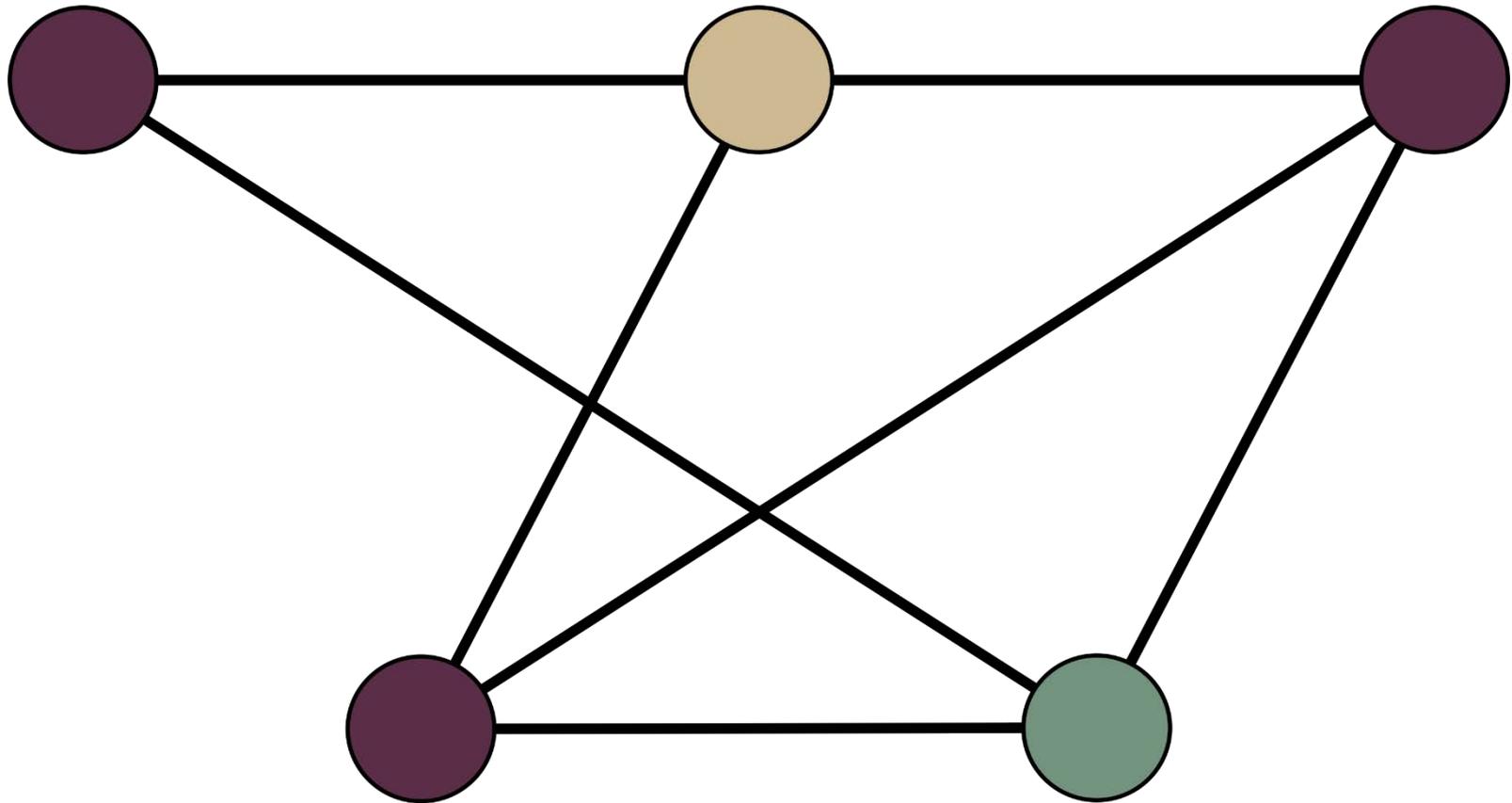
# Patient Sharing Networks

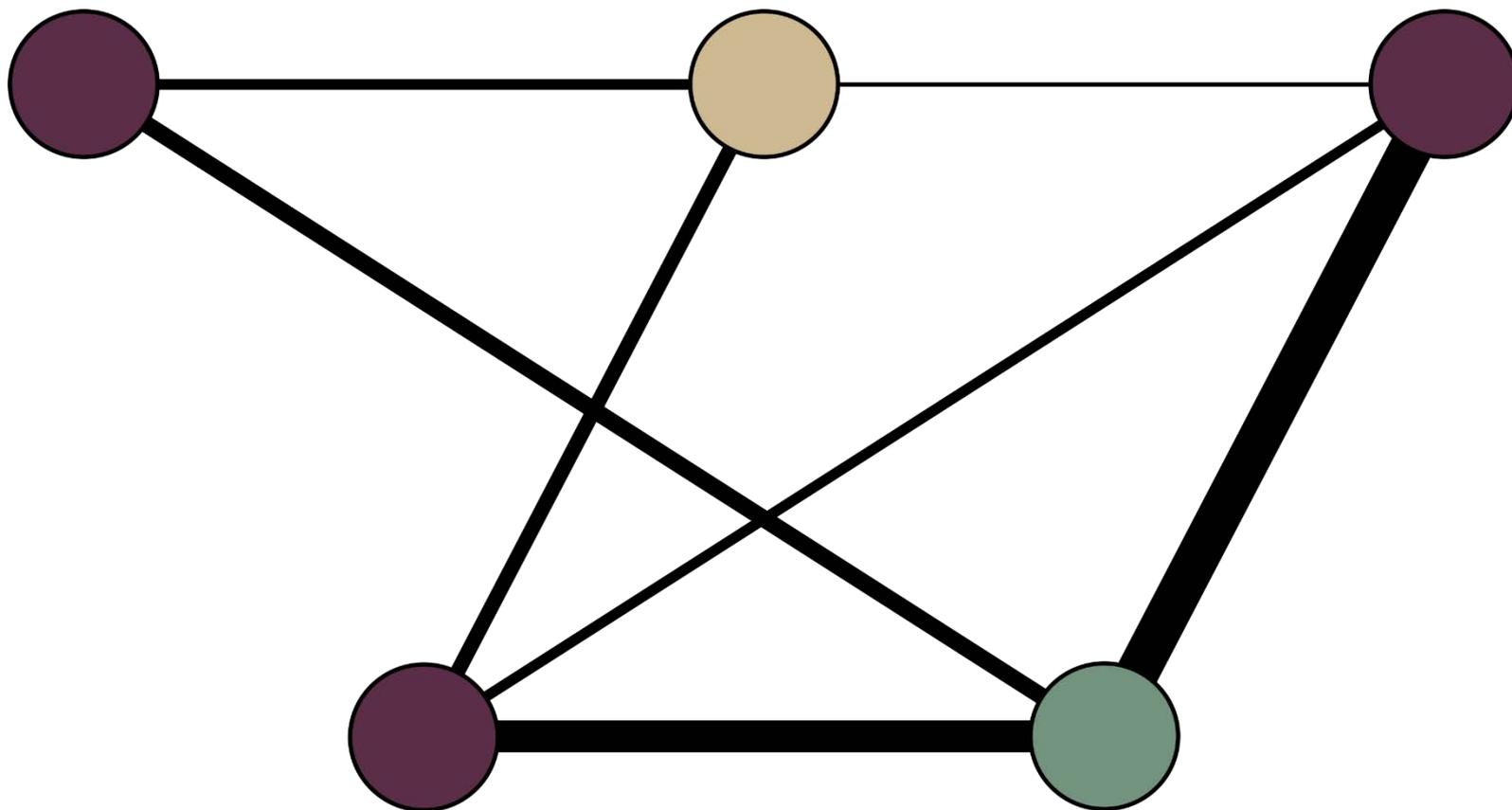
## Soziales Netzwerk

- Knoten: Leistungserbringende
- Kanten: gemeinsamen Patient\*innen

Annahme:

Mehr geteilte Patient\*innen → engere Verbindung





# Daten

- 7 Millionen AOK-Versicherte über 65 Jahren
- ca. 168.000 Leistungserbringende (Arztpraxen, Krankenhäuser, Heilmittelerbringende, Reha)
- Ausschluss: Labormedizin, Pathologie, Anästhesie
- ca. 25 Mio. Kanten

# Datenaufbereitung

- Reduktion der Datenmenge notwendig:  
80/85/90%-Perzentile
- Resultat:
  - gerichtetes, gewichtetes Netzwerk
  - 168.000 Knoten
  - 5–13 Mio. Kanten

# Community Detection

Anforderungen:

- „Wahre“ Gruppenzugehörigkeit unbekannt
- Zugehörigkeit zu mehreren Clustern (*overlapping communities*)
- Für gerichtete, gewichtete Netzwerke geeignet
- Programmlaufzeit

# Algorithmen

- SLPA, COPRA
- Auswahlkriterien:
  - Algorithmus
  - Parametereinstellungen
  - Netzwerkeigenschaften

# Ergebnisse

# Grundsätzliche Eigenschaften

- Anzahl/Größe der Cluster
- Zusammensetzung
- Hypothesen:
  - Mehrere tausend Akteure je Cluster eher unwahrscheinlich
  - Alle Arten/Sektoren in jedem Cluster vertreten

Modellparameter			Anzahl Cluster	Anzahl Leistungserbringer					
Algorith- mus	Kanten	Quantil		Min.	Q1	Median	Mittelw.	Q3	Max.
<b>COPRA</b>	Patienten	15 %	531	1	119	191	343	329	8023
<b>COPRA</b>	Patienten	10 %	631	1	99	171	290	297	7758
<b>COPRA</b>	Fälle	10 %	669	1	102	158	274	271	7740
<b>SLPA</b>	Patienten	20 %	870	4	72	121	190	211	1792
<b>SLPA</b>	Patienten	10 %	1047	2	58	104	157	189	1465
<b>SLPA</b>	Fälle	10 %	1226	2	47	92	135	161	1417

Modellparameter			Anteile von Arten an Leistungserbringern					
Algorith- mus	Kanten	Quantil	KH		niedergel. Ärzt*innen		Heilmittelerbringende	
			Median	Mittelwert	Median	Mittelwert	Median	Mittelwert
<b>COPRA</b>	Patienten	15%	1,6%	1,8%	65,6%	65,3%	32,1%	32,4%
<b>COPRA</b>	Patienten	10%	1,5%	1,8%	65,2%	63,5%	32,2%	34,2%
<b>COPRA</b>	Fälle	10%	1,5%	2,0%	65,1%	64,1%	32,4%	33,2%
<b>SLPA</b>	Patienten	20%	1,5%	1,6%	66,6%	66,4%	31,4%	31,6%
<b>SLPA</b>	Patienten	10%	1,4%	1,6%	66,7%	66,7%	31,4%	31,3%
<b>SLPA</b>	Fälle	10%	1,2%	1,4%	66,7%	66,3%	31,4%	31,9%

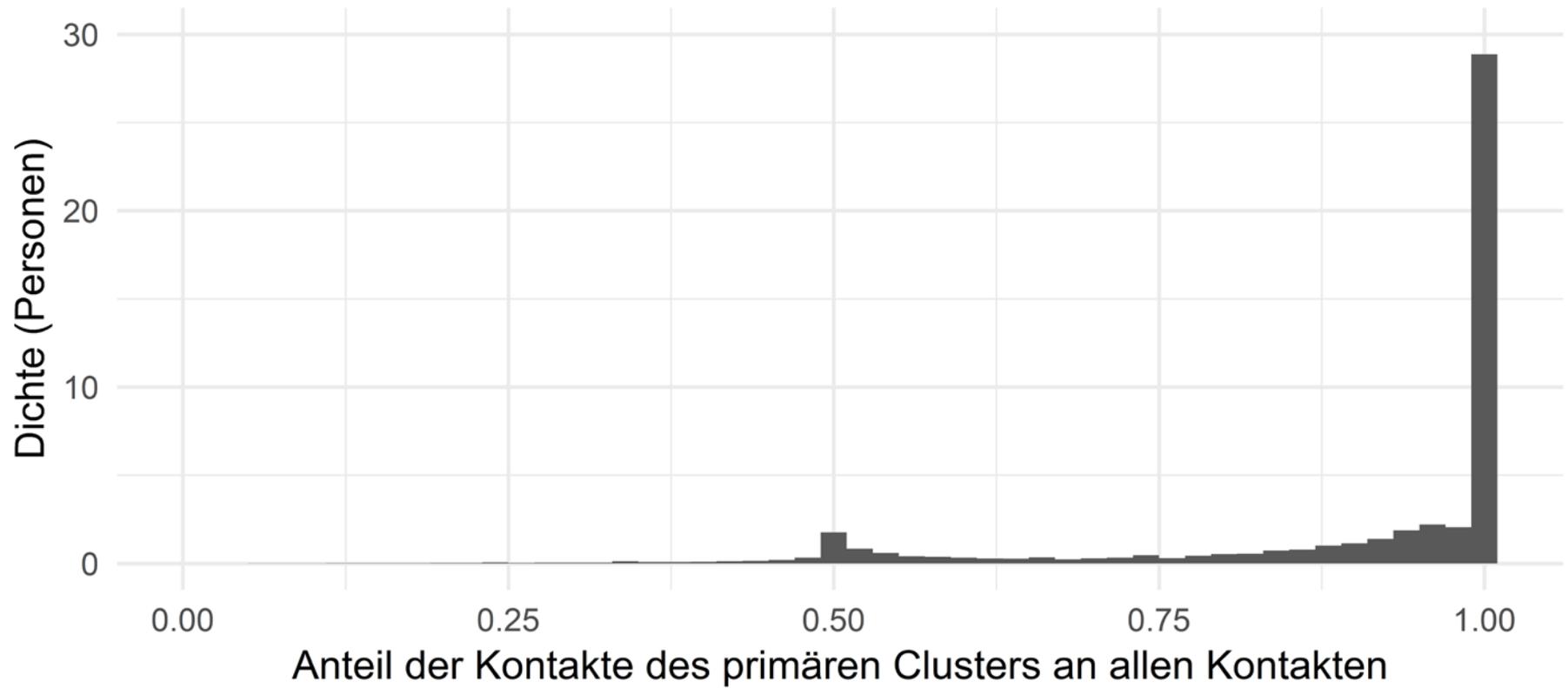
# Versorgte Personen

Zuordnung der Patient\*innen zu dem Cluster, bei dessen Leistungserbringenden die Mehrzahl der Behandlungsfälle stattfanden

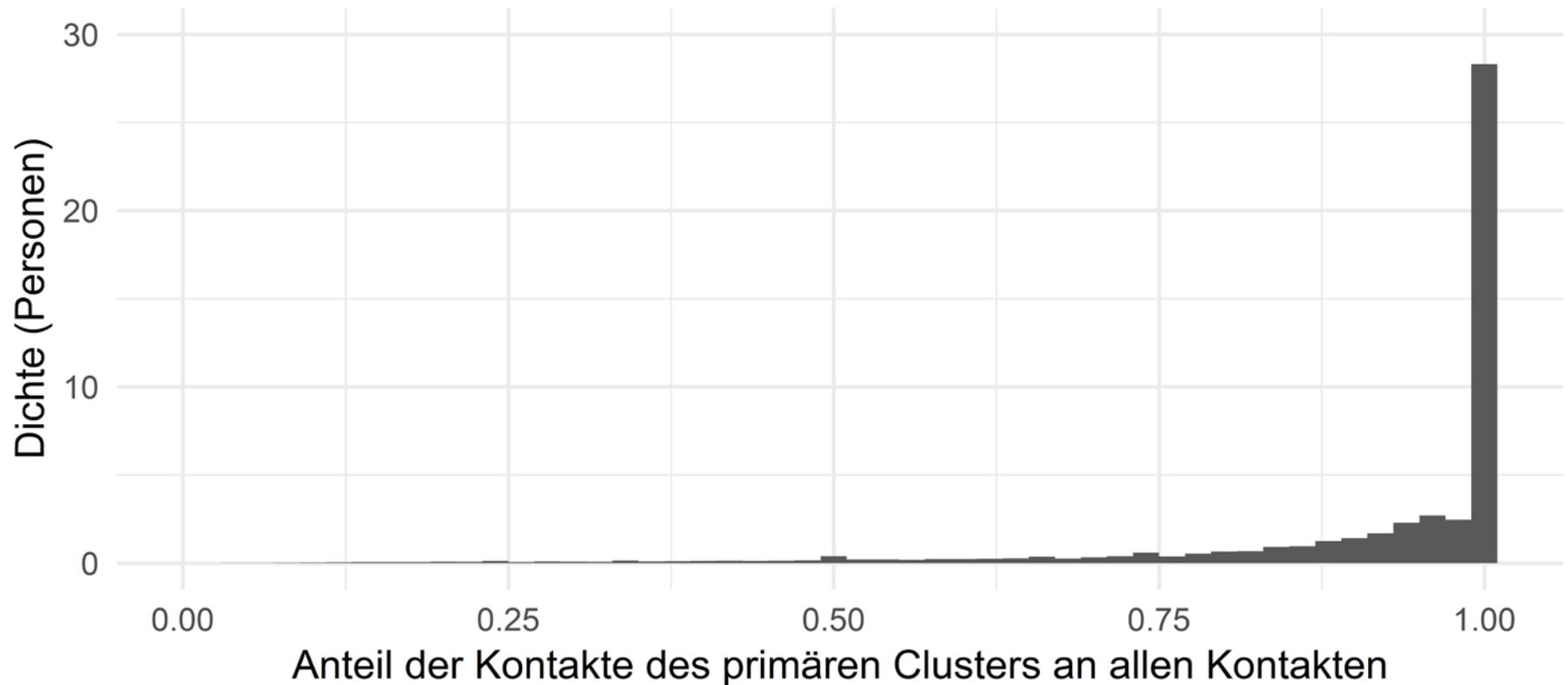
Hypothese:

Überwiegende Anzahl aller Kontakte/Fälle je Person sollten durch primäres Cluster erfolgen

## COPRA, Gewichtung: Patienten, 10%



### SLPA, Gewichtung: Patienten, 15%



# Geografische Verteilung

Leistungserbringende als geografische Mittelpunkte  
der versorgten Personen

Hypothesen:

- Überlappung stärker bei geografisch benachbarten Clustern
- Akteure innerhalb eines Clusters weisen üblicherweise keine hohen Distanzen auf



# Diskussion

- Erste Ergebnisse: nicht unplausibel
- Validierung von Clustern → Metriken?
- Übertragbarkeit?
- Potenzielle Alternative zu geografisch definierten Abgrenzungen

# Ausblick

- Anwendung zur Darstellung von Gruppenunterschieden
- Weitere Überprüfung des Konzeptes von PSN in Deutschland

**VIELEN DANK!**