

# **CROSS DATA SET ANALYSIS**

**Data are  
Everywhere**

# **Cross Data Set Analysis**

**Why?**

**How?**

**What will you find?**

**Why ?**

**Data Infrastructure Grant**

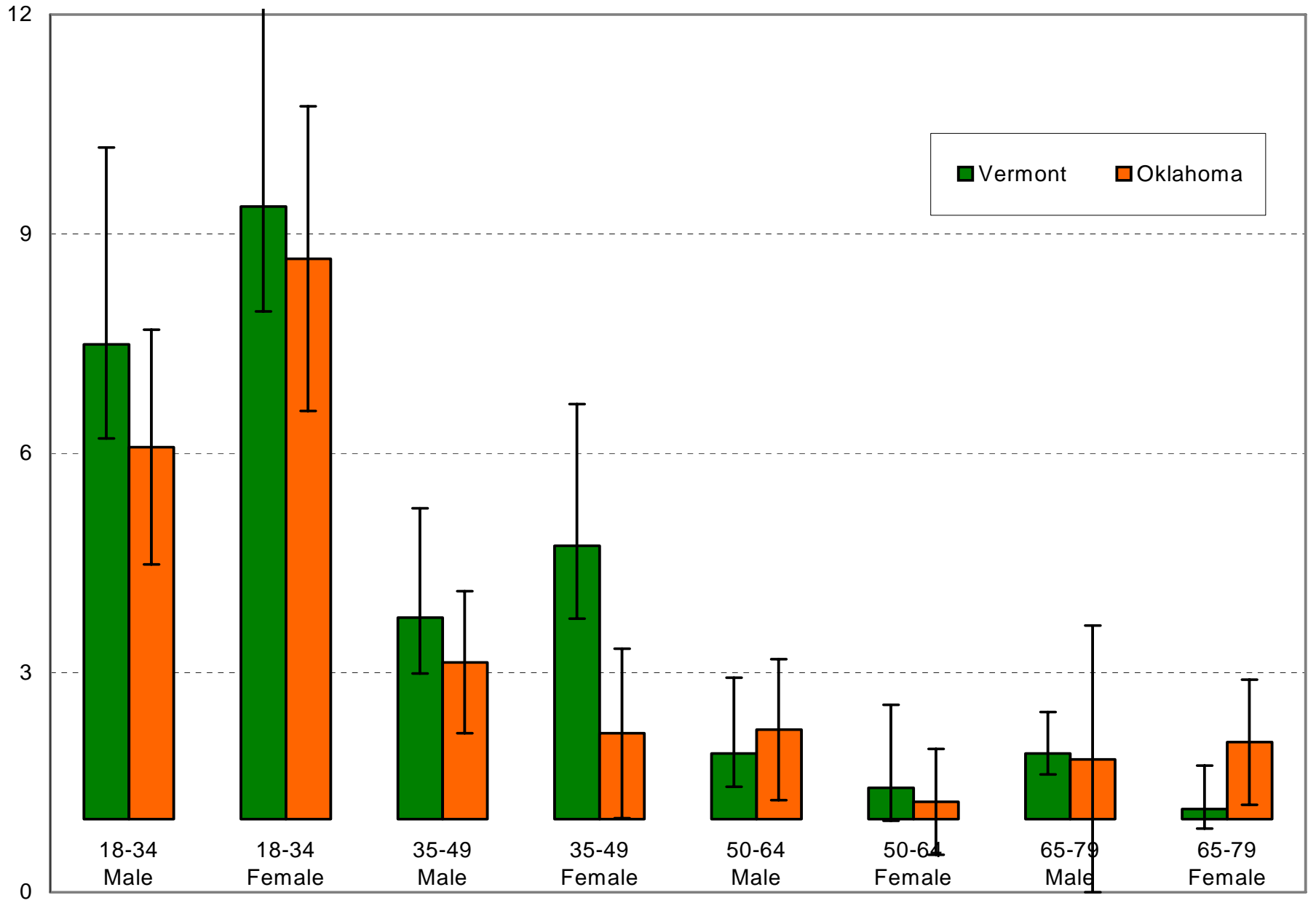
**Employment**

**Trouble with the Law**

**School Indicators**

**Mortality**

# Relative Risk of Mortality



# **Why ?**

**Less clinician burden**

**Less consumer burden**

**More valid and reliable**

**More cost efficient**

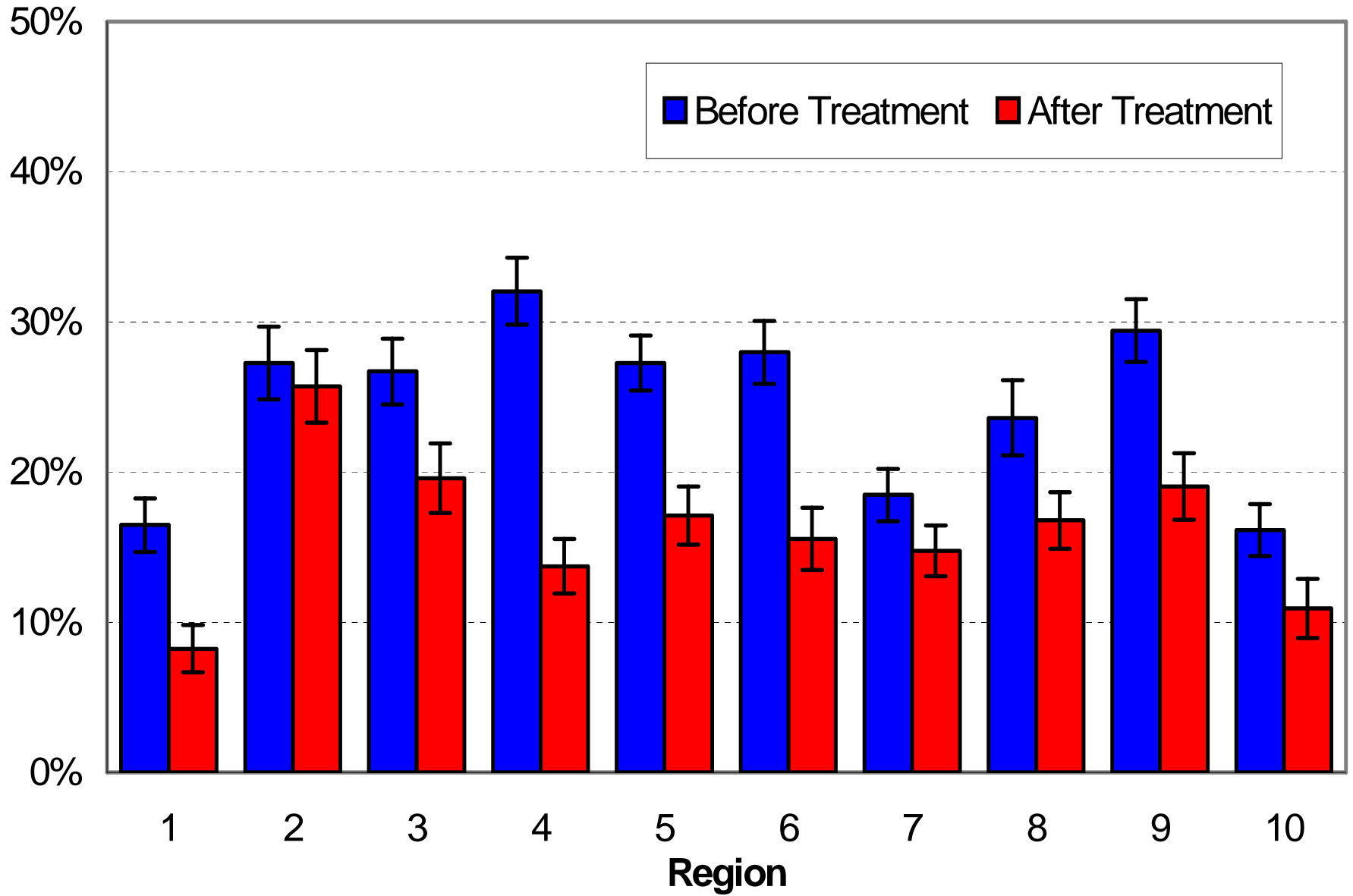
**Why ?**

**Ongoing & Sustainable**

**Longitudinal**

**Pre and post studies**

# Percent Arrested



**Why ?**

**Answer Other Questions**

**Unmet need**

**Access to health care**

**Service system integration**

**Why ?**

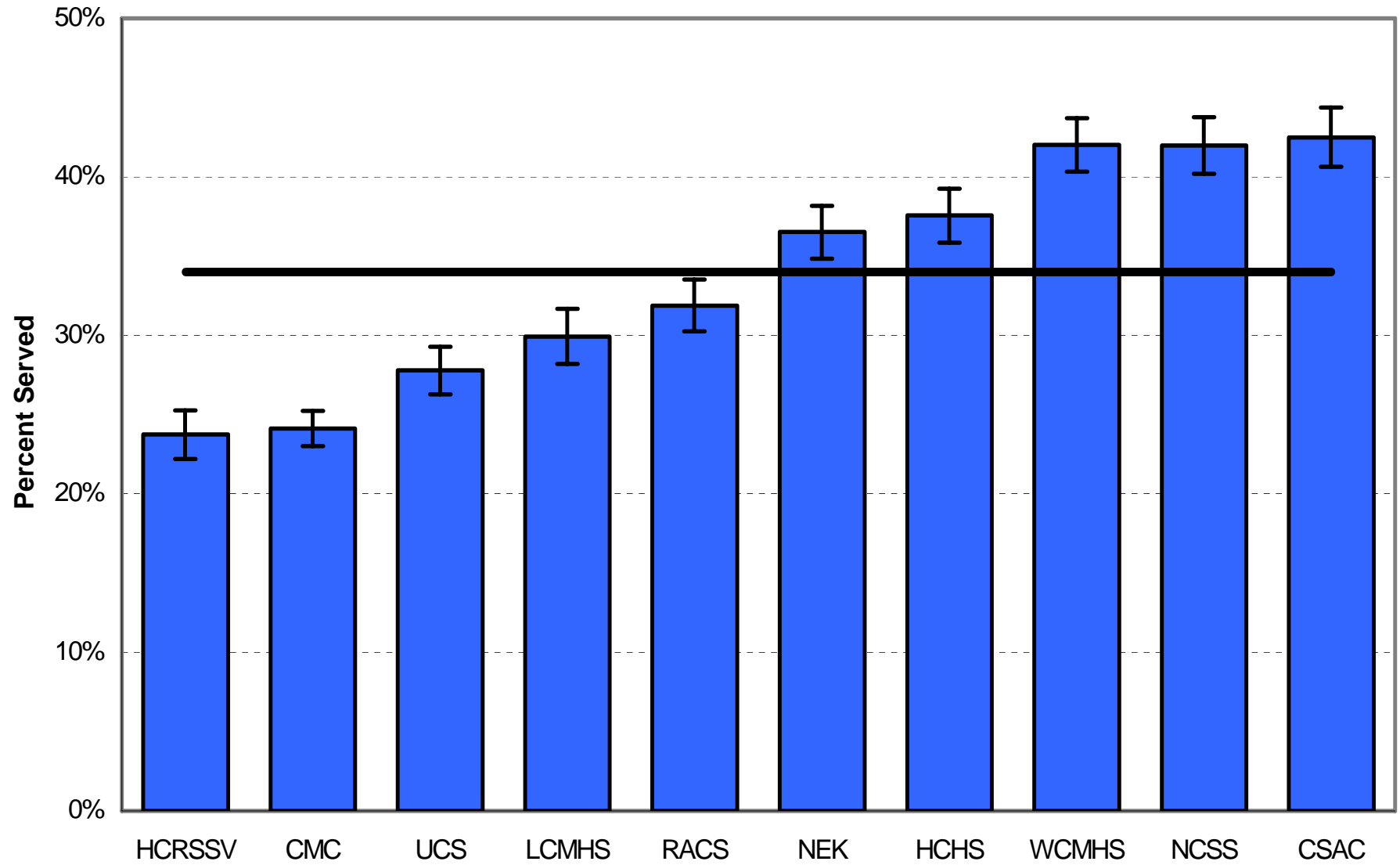
**Answer Other Questions**

**Treatment Outcomes**

**Hospitalization**

**Economic Dependency**

# Access to Mental Health Services Adults Hospitalized During the Previous Year



**How ?**

**Direct Record Linkage**

**Constructed Identifiers**

**Probabilistic**

**Population Estimation**

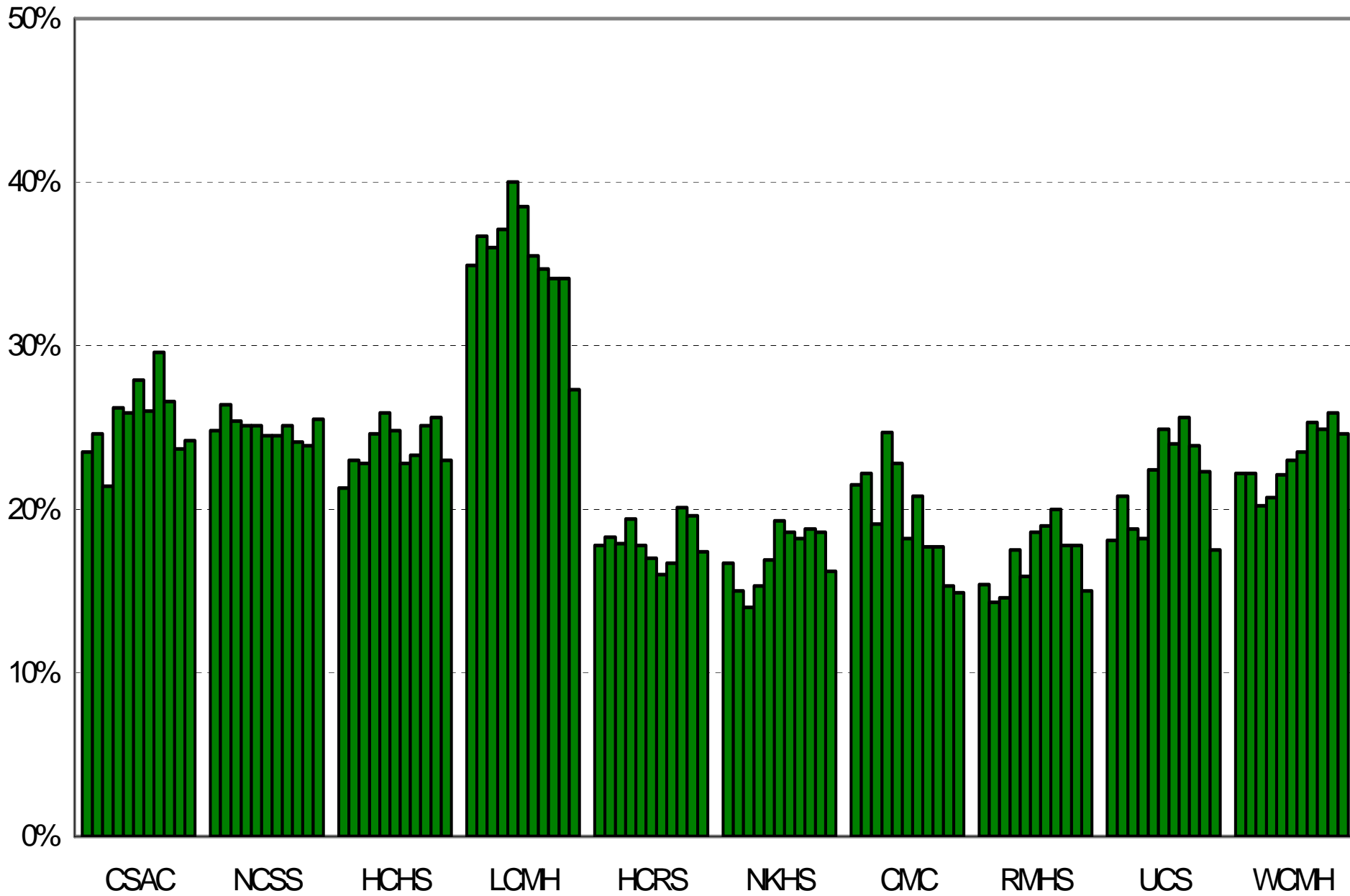
**How ?**

**Direct Record Linkage**

**Gold Standard**

**Confidentiality Issues**

# Percent Employed



**\$4,664,164**

**How ?**

**Constructed Identifiers**

**Confidentiality issues**

**False positives**

**False negatives**

# Oklahoma's Probabilistic Matching Process

Block by Sex

Join by DOB Components

Assign Weights  
to All Identifying Variables

# **Join by DOB Components**

Month and Year

Day and Year

Month and Day

# Assign Weights to All Identifying Variables

Perfect Match = 49 pts.

SSN (all digits) = 22.55 or -2.38 pts.

SSN (8 digits) = 14.51 pts.

SSN (7 digits) = 6.07 pts.

Last or Maiden Name = 9.58 or -3.62 pts.

First Name = 6.69 or -3.27 pts.

Middle Initial = 3.65 pts.

Date of Birth = 6.22 pts.

Positive & Negative Weights

Jaro, Matthew. Statistics in Medicine, Vol. 14, 491-498 (1995).

Determine Threshold of Likelihood

Gray Area - Manual Inspection

Matching Algorithm Changes Depending on Available Variables

**How ?**

**Probabilistic  
Population Estimation**

**Not for clinical applications**

**Tells "how many?"**

**Not "who?"**

# **Data Elements**

**Required:**

**Date of Birth & Gender**

**Recommended:**

**Race/Ethnic & Geographical**

$$P_j(l_j) = \sum_{i=1}^{l_j} \frac{365}{365 - i}$$

# Probabilistic Population Estimates

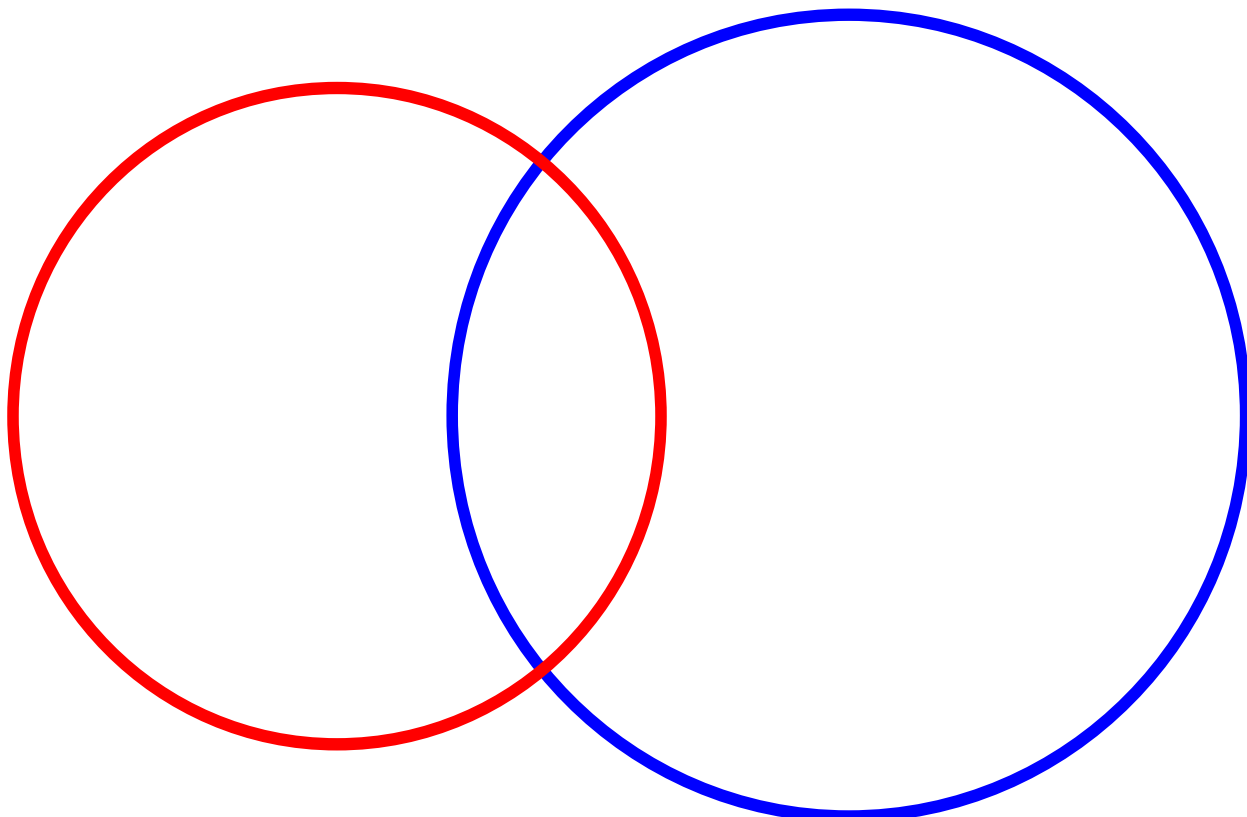
for Specified Numbers of Birth Dates Within a Given Year

## Birth Dates

## Number of People

1	1.003	±	.103
10	10.15	±	.776
100	117	±	9
200	290	±	22
300	632	±	64
350	1177	±	171
360	1603	±	325

# Caseload Overlap

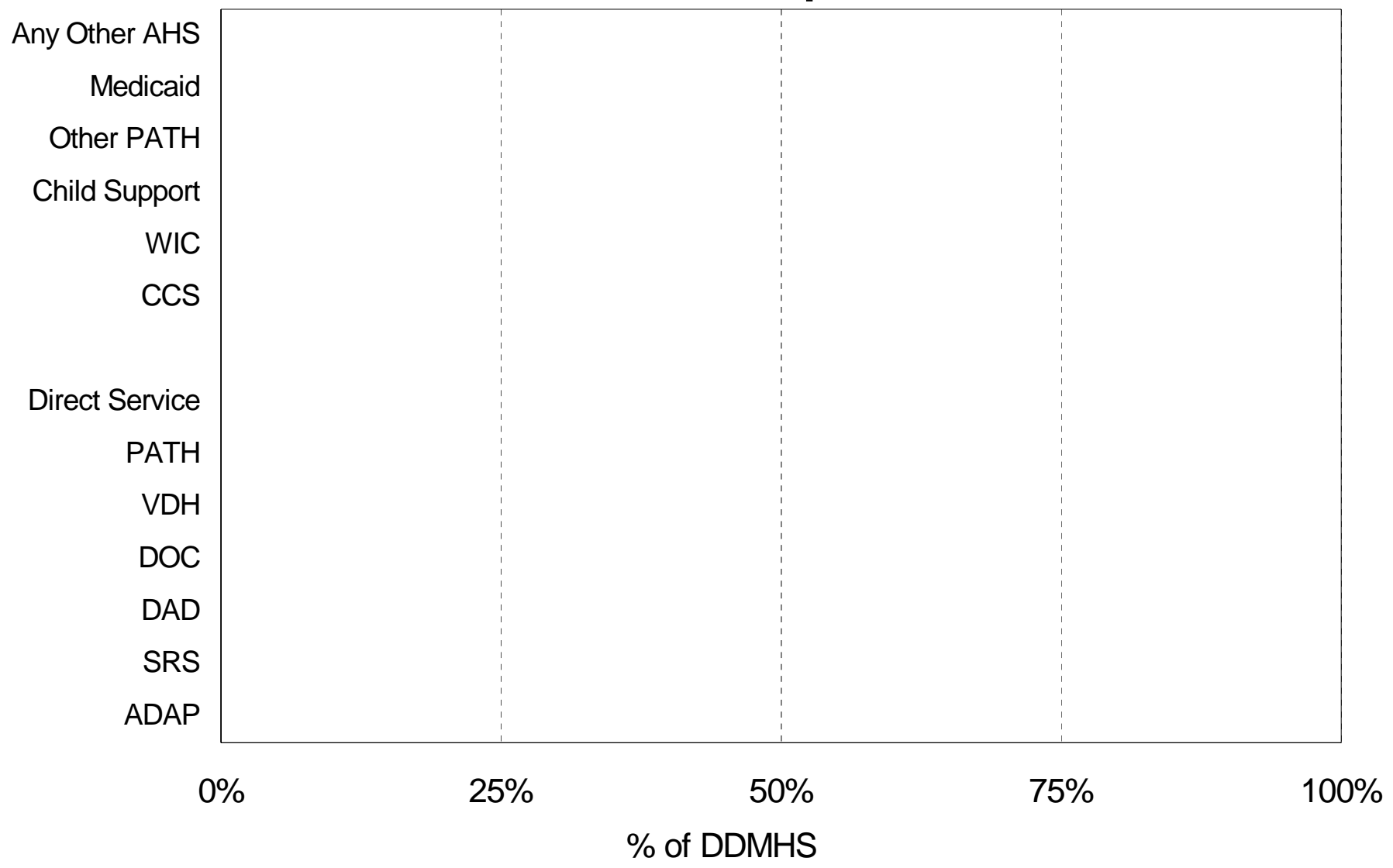


$$(A \cap B) = A + B - (A \cup B)$$

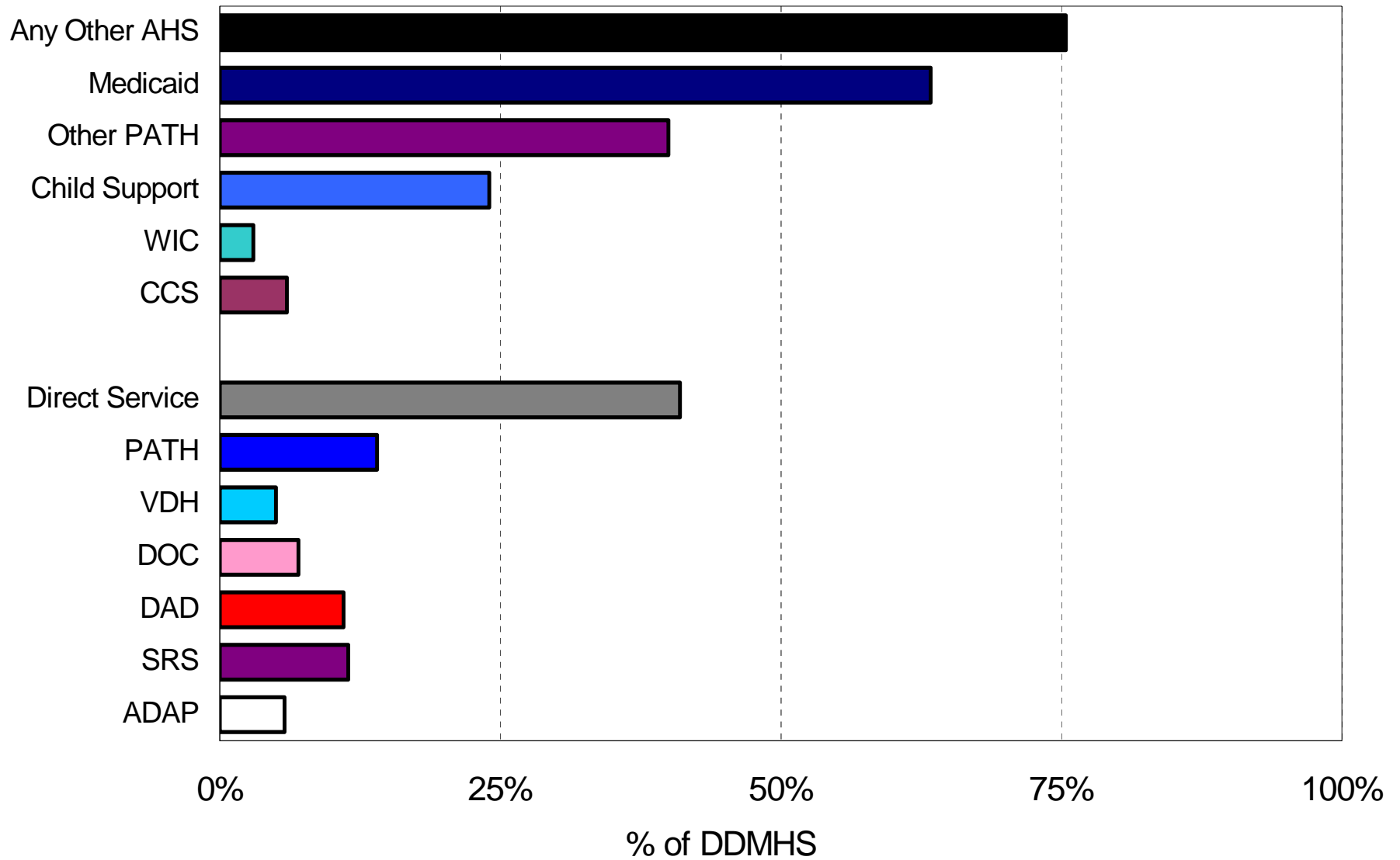
**Vermont**  
**Agency of Human Services**  
**42 Data Sets**

<b>Medicaid</b>	<b>(1)</b>
<b>PATH</b>	<b>(7)</b>
<b>Child Support</b>	<b>(1)</b>
<b>VDH</b>	<b>(8)</b>
<b>DOC</b>	<b>(3)</b>
<b>DDMHS</b>	<b>(6)</b>
<b>SRS</b>	<b>(6)</b>
<b>DAD</b>	<b>(9)</b>
<b>ADAP</b>	<b>(1)</b>
	<b>(42)</b>

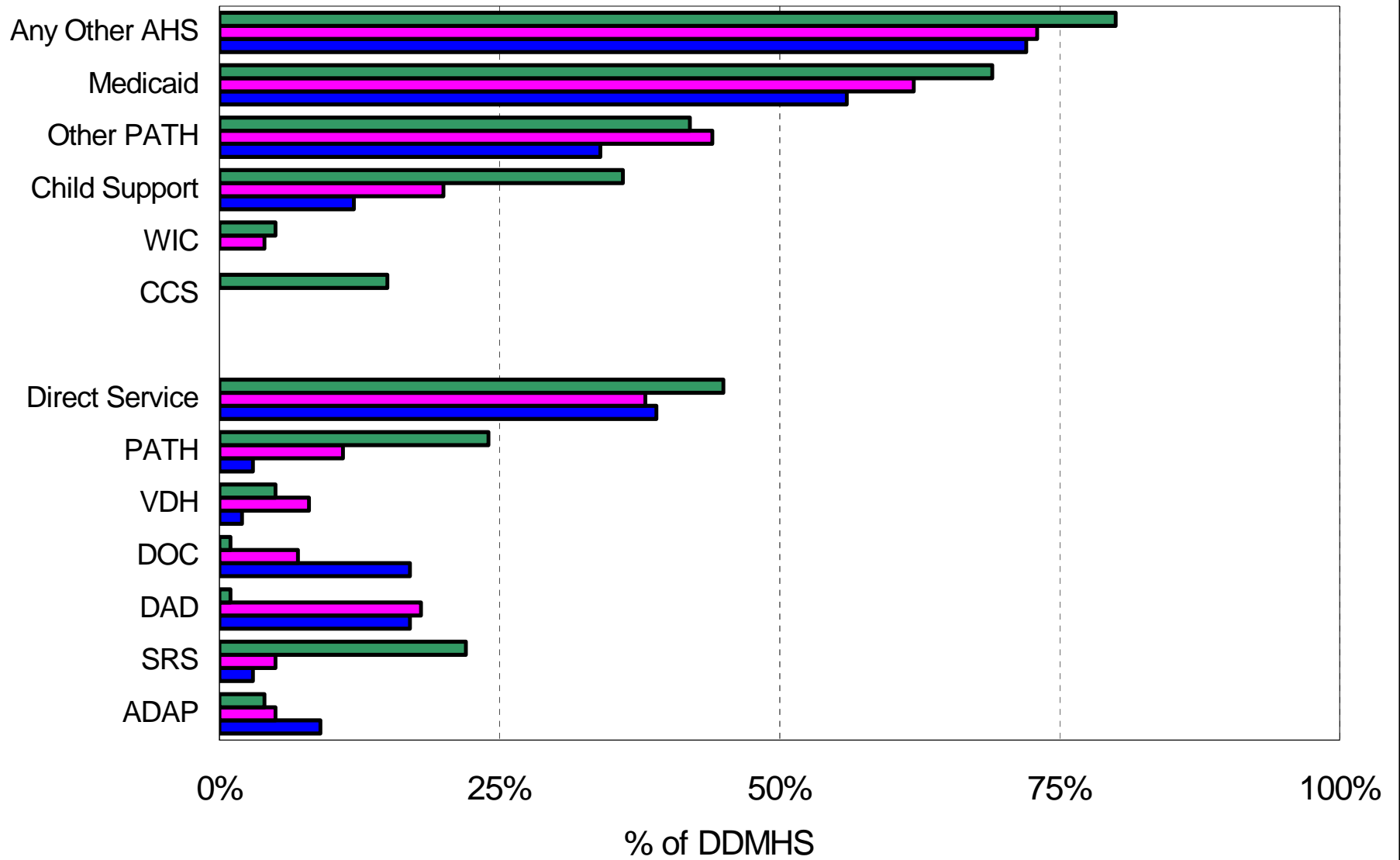
# AHS Caseload Overlap DDMHS and Other Departments



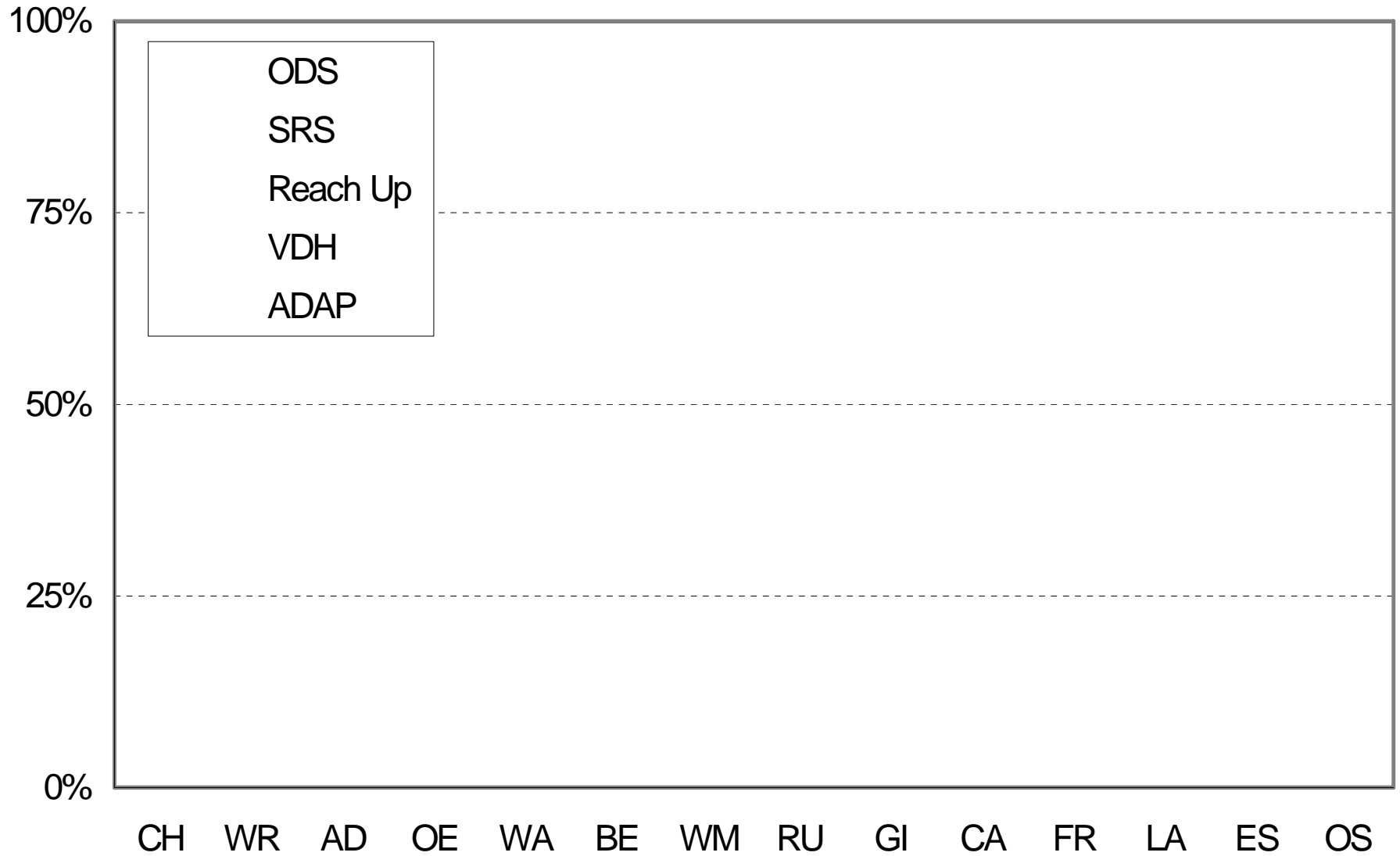
# AHS Caseload Overlap DDMHS and Other Departments



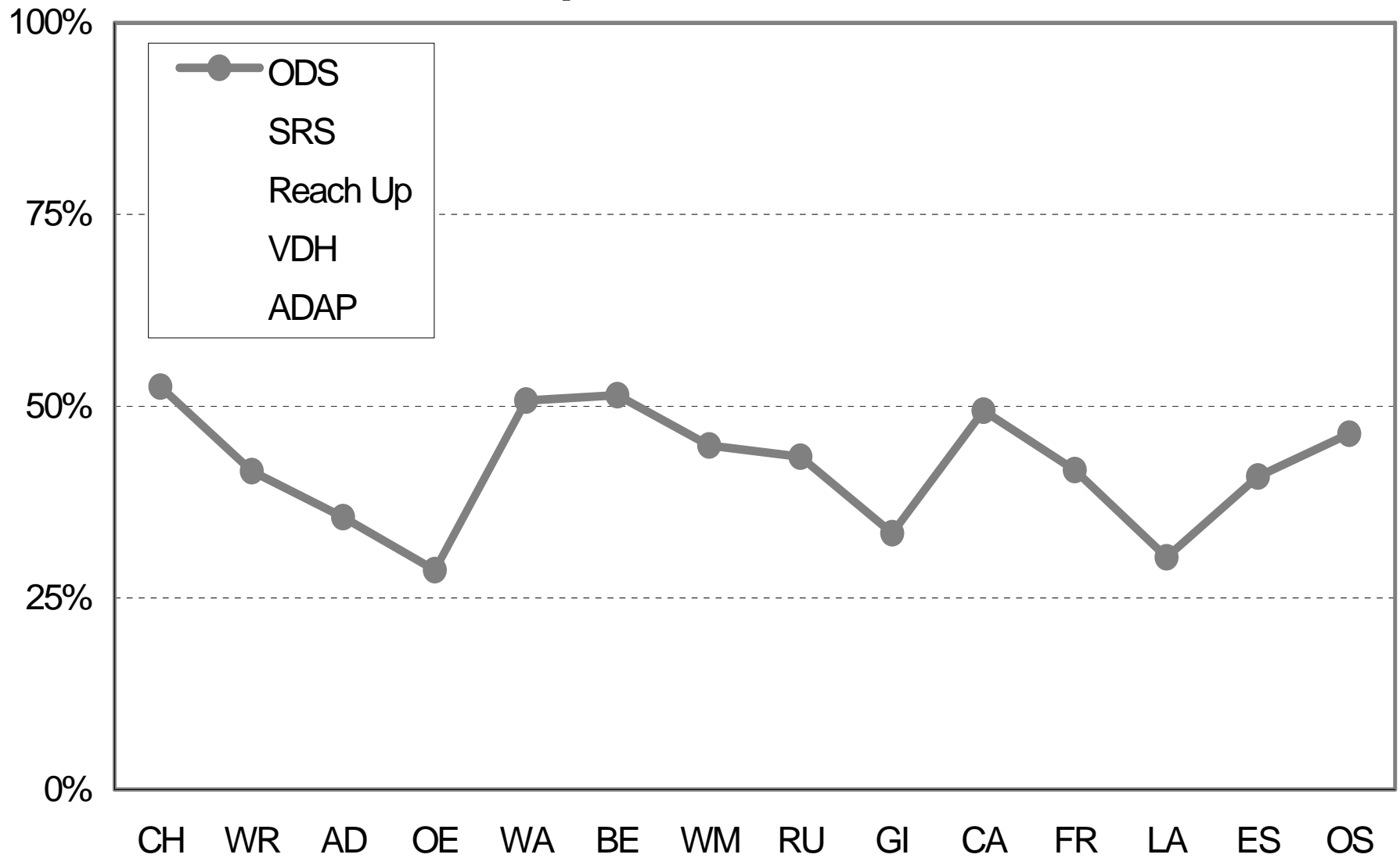
# AHS Caseload Overlap DDMHS and Other Departments



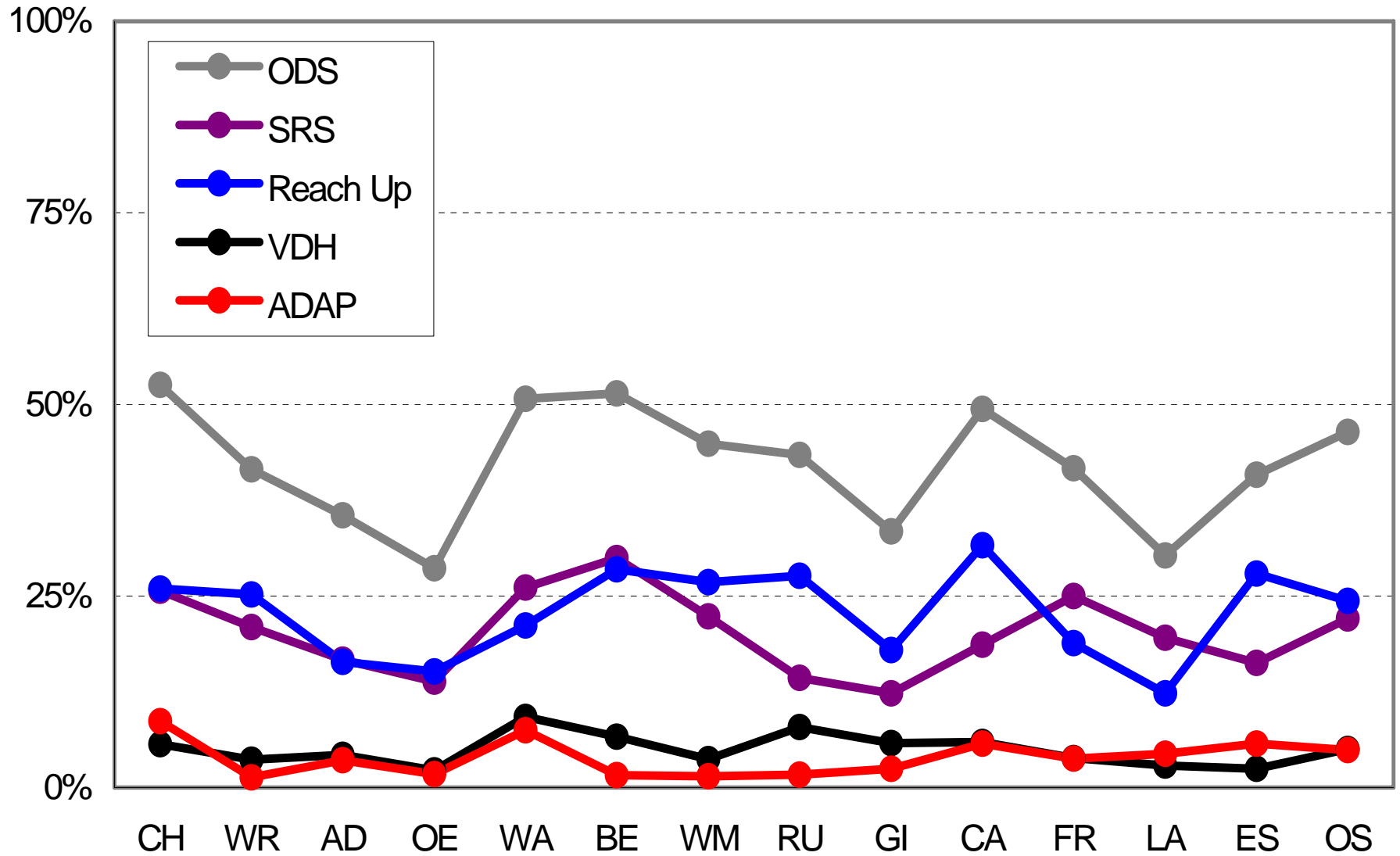
# DDMHS Children's Services Caseload Overlap with Other Direct Services



# DDMHS Children's Services Caseload Overlap with Other Direct Services



# DDMHS Children's Services Caseload Overlap with Other Direct Services



# **Question:**

Which departments have the  
greatest caseload overlap  
and how much?

**Jeff Cohen, OCS**

# AHS Caseload Overlap FY 2001

## Children

### Direct Services

	# Served	Reach Up	VDH	DOC	DDMHS	DAD	SRS	ADAP
Reach Up	15,159		2,204	79	1,850	50	1,282	108
VDH	8,462	2,204		23	416	25	365	23
DOC	607	79	23		101	22	142	134
DDMHS	7,830	1,850	416	101		97	1,710	335
DAD	283	50	25	22	97		60	32
SRS	5,330	1,282	365	142	1,710	60		264
ADAP	974	108	23	134	335	32	264	

# AHS Caseload Overlap FY 2001

## Children

### Direct Services

	# Served	Reach Up	VDH	DOC	DDMHS	DAD	SRS	ADAP
Reach Up	15,159		2,204	79	1,850	50	1,282	108
VDH	8,462	2,204		23	416	25	365	23
DOC	607	79	23		101	22	142	134
DDMHS	7,830	1,850	416	101		97	1,710	335
DAD	283	50	25	22	97		60	32
SRS	5,330	1,282	365	142	1,710	60		264
ADAP	974	108	23	134	335	32	264	

**Children's MH**

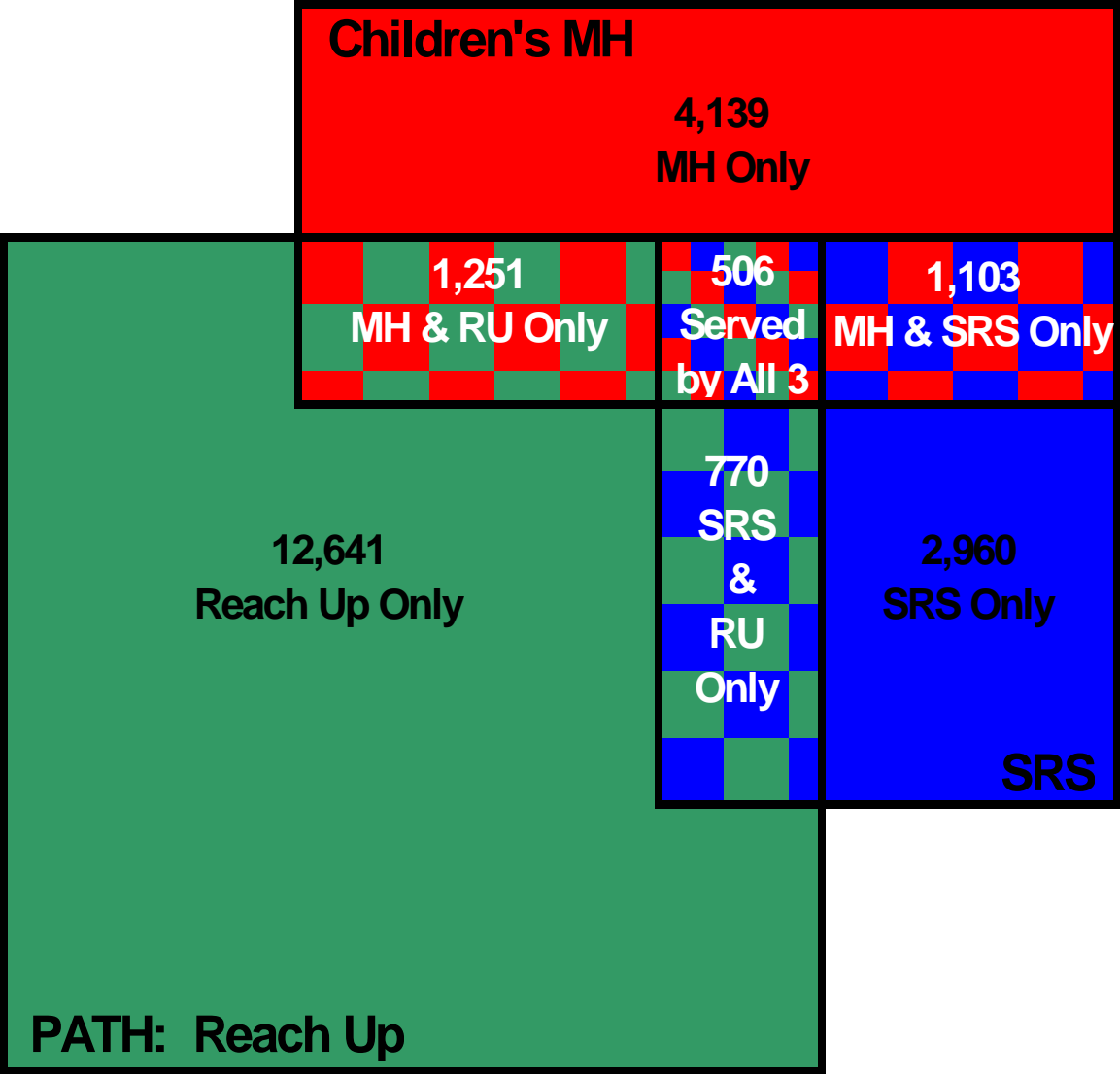
**6,991 Served**

**15,159 Served**

**PATH: Reach Up**

**5,330 Served**

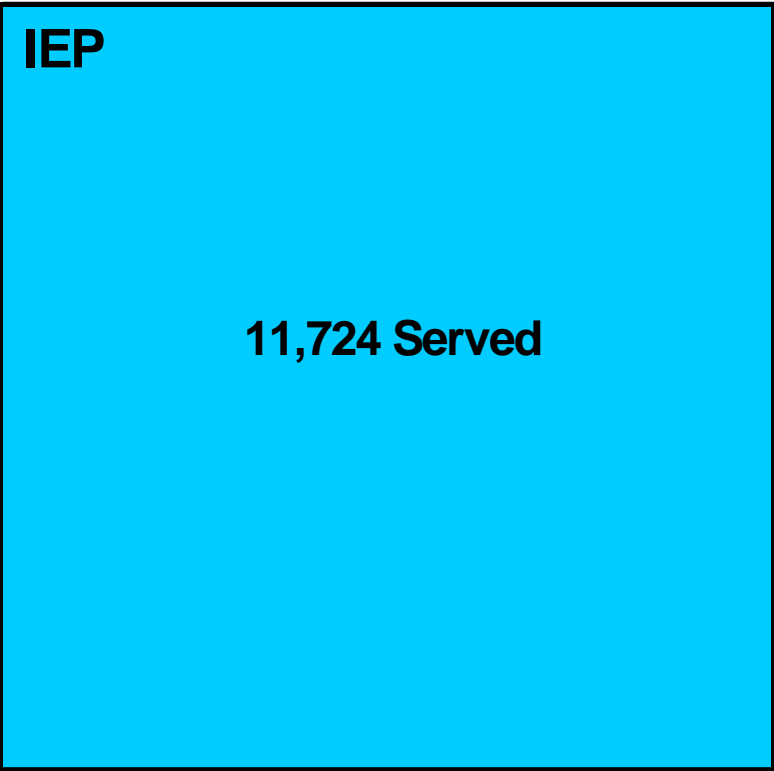
**SRS**



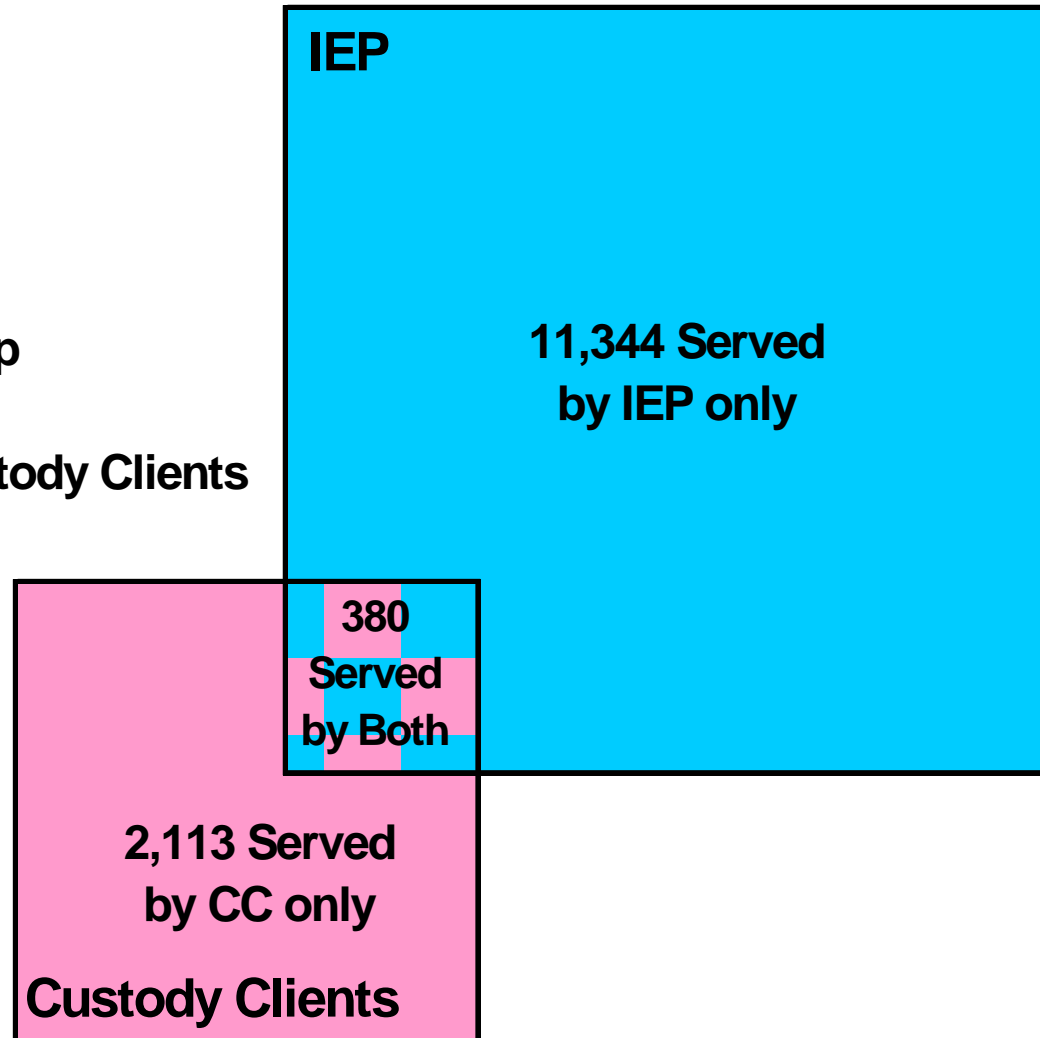
# **Question:**

How many kids in SRS custody  
are on an IEP in school?

**David Murphy, AHS**



**Overlap**  
**3% of IEP**  
**15% of Custody Clients**



# **Process**

**Getting Buy-In**

**Interagency Agreements**

**Disseminating Results**

**SAVE**

**THE DATA**

