Phenotyping Evaluation Presentation

Data Types – Testing for hypertension

- **Blood pressure (BP) events** (from table mimic3_demo.CHARTEVENTS)
  - Any event containing the word “systolic”, VALUENUM >= 140, count > 2 times

  AND
  - Any event containing the word “diastolic”, VALUENUM >= 90, count > 2 times

- **ICD-9 diagnosis codes** (from table mimic3_demo.DIAGNOSES_ICD)
  - 401.0 OR 401.1 OR 401.9, for essential hypertension (EH)

  OR
  - Any ICD-9 code containing the word “hypertension”, which includes codes for EH

- **Prescriptions** (from table mimic3_demo.PRESCRIPTIONS)
  - Antihypertensive drugs from table course3_data.D_ANTIHYPERTENSIVES

Clinical criteria alone

BP events	Manual review hypertension: +	Manual review hypertension: -	*n*
+	14	8	22
-	49	28	77
n	63	36	99

Sensitivity: 22.2%

Specificity: 77.8%

PPV: 63.7 %

NPV: 36.3%

ICD-9 codes alone: Essential hypertension (EH) or all codes for “hypertension”

EH: 401.0 -OR- 401.1 -OR- 401.9

Manual review hypertension

EH codes	+	-	*n*
+	35	3	38
-	28	33	61
n	63	36	99

Sensitivity: 55.6%; Specificity: 91.7%
PPV: 92.1%; NPV: 54.1%

all: ICD codes for “hypertension”

Manual review hypertension

all codes	+	-	*n*
+	37	4	41
-	26	32	58
n	63	36	99

Sensitivity: 58.7%; Specificity: 88.9%
PPV: 90.2%; NPV: 55.2%

Prescriptions alone

prescriptions \ Manual review hypertension	+	-	*n*
+	52	24	76
-	11	12	23
n	63	36	99

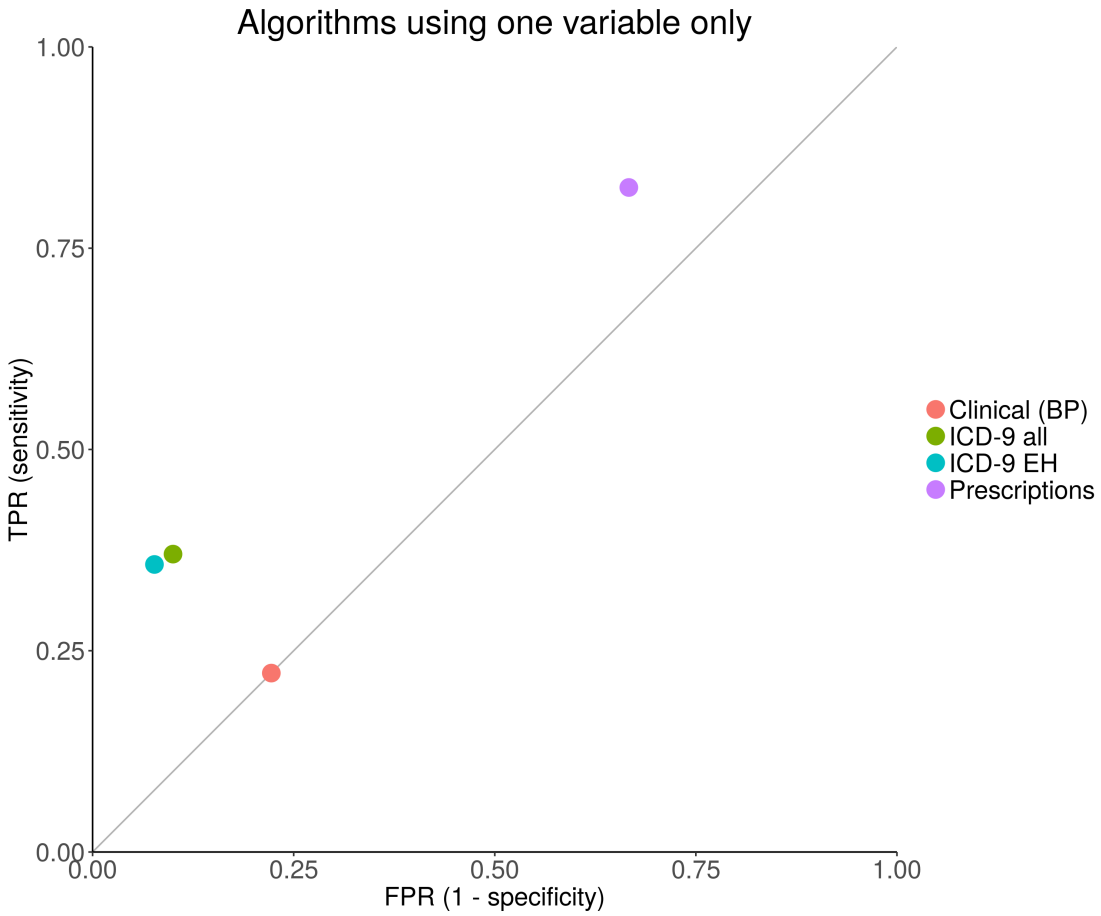
Sensitivity: 82.5%

Specificity: 33.3%

PPV: 68.4 %

NPV: 52.2%

Summary of individual data types



- Clinical criteria don't perform better than random.
- ICD-9 codes have low sensitivity; using all codes doesn't seem to improve the performance dramatically. Therefore, one could debate on the small trade-off between FPs and TPs in the classification.
- Prescriptions seem to be the best data type.

Combining clinical criteria with ICD-9 codes

BP -OR- ICD-9 (EH)

Manual review hypertension

criteria	+	-	*n*
+	39	9	48
-	24	27	51
n	63	36	99

Sensitivity: 61.9%; Specificity: 75%
PPV: 81.3%; NPV: 52.9%

BP -OR- ICD-9 (all)

Manual review hypertension

all codes	+	-	*n*
+	41	9	50
-	22	27	49
n	63	36	99

Sensitivity: 65.1%; Specificity: 75%
PPV: 82.0%; NPV: 55.1%

Combining clinical OR prescription

		Manual review hypertension		
		+	-	*n*
criteria	+	54	24	78
	-	9	12	21
	n	63	36	99

Sensitivity: 85.7%

Specificity: 33.3%

PPV: 69.2%

NPV: 57.1%

Combining all criteria
(clinical AND prescription) OR ICD-9 code

		Manual review hypertension: +	Manual review hypertension: -	Manual review hypertension: *n*
criteria	+	40	8	48
	-	23	28	51
	n	63	36	99

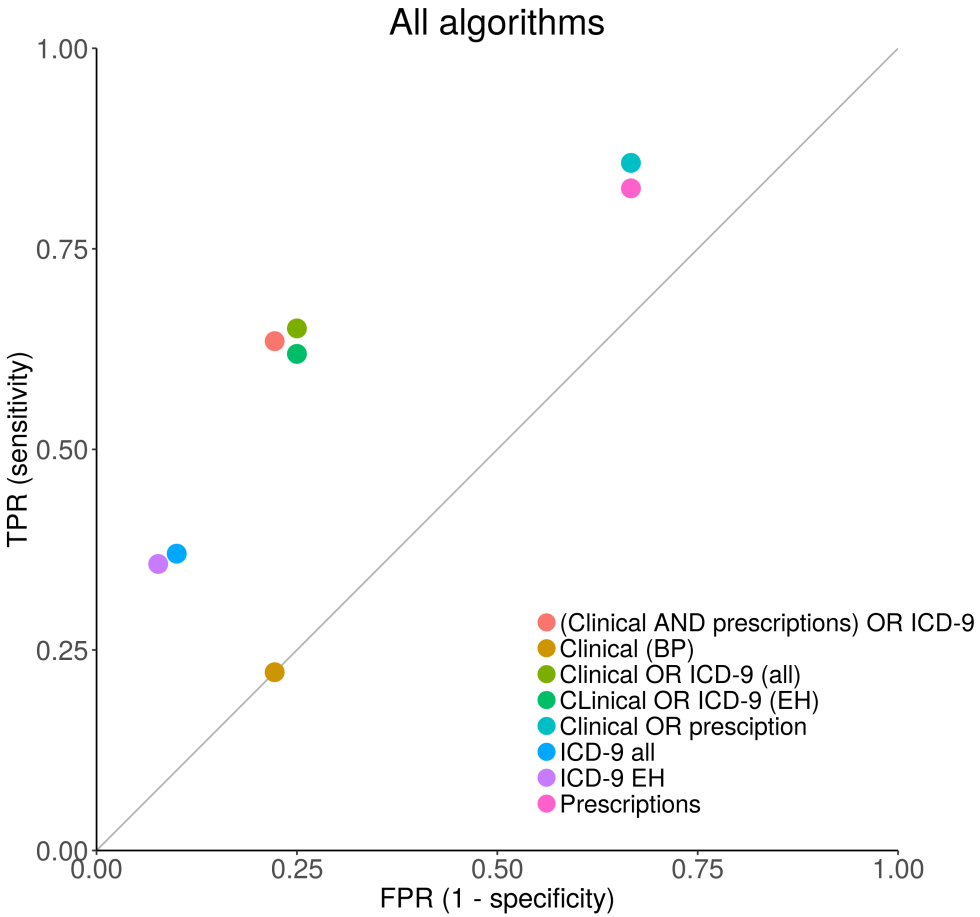
Sensitivity: 63.5%

Specificity: 77.8%

PPV: 83.3%

NPV: 54.9%

Summary of all algorithms data types



- Using only ICD-9 codes or clinical criteria leads to poor performance.
- Best combination is achieved by combining clinical criteria OR prescriptions.
- Adding ICD-9 codes leads to improved sensitivity of the algorithm.

The best algorithm is

- “(clinical AND prescription) OR ICD-9 code”, because:
- The algorithm has a good performance in terms of balanced of sensitivity (63.5%) and specificity (77.8%).
- The algorithm is however rather complex in terms of combining clinical data with billing codes.
- The algorithm needs several pieces of information to be assembled, so it may not be portable to other providers not having access to all these data.