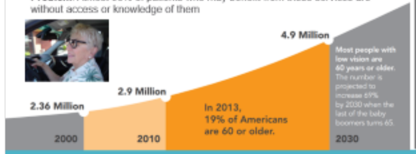
Computing Over Ophthalmology Clinical Text From Electronic Health Records: Case Study in Identifying Candidates for Low Vision Rehabilitation Using Neural Word Embeddings

Sophia Y Wang, MD; Benjamin Tseng
Byers Eye Institute, Department of Ophthalmology, Stanford University

Background

Low Vision and Rehabilitation Services

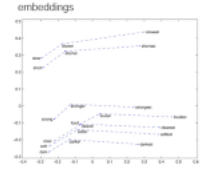
- **Low Vision**: Visual impairment that cannot be corrected with glasses, contacts, or other treatments. Prevalence increasing
- **Severely impacts activities of daily living**: 3x unemployment, 3x MVC, if they can even drive, 3x depression and anxiety, 2x falls
- **Low vision rehabilitation services**: Palliative care for the eye, maximizing use of remaining vision
- **Problem**: Almost 90% of patients who may benefit from these services are without access or knowledge of them



Ophthalmology Progress Note
HPI: 56 yo F presents with decreased VA...
OcHx:
- Cataracts OU
- Macular Degeneration OU
Gtt:
AREDS2 vitamins
Exam:
Vacc OD: 20/40-2 ph 20/25-2
OS: 20/200+1 ph 20/150-2
Tapp OD: 16
OS: 17
SLE:
L/L: wnl OU
C/S: w/q OU
K: clear OU
AC: d/q OU
Iris: r/r OU
Lens: 2+ NS with 1+CC OU
AntVx: wnl OU
DFE: CDR 0.5 /0.5, macular with drusen OD and GA OS, normal vessels/periphery OU
A/P
1) Dry macular degeneration OU...

Electronic Health Records (EHR)

- So much data in EHR
- How to make use of free text?
- Turn the text into numbers via word embeddings



Problem: General word embeddings don't cover highly domain-specific text

Purpose

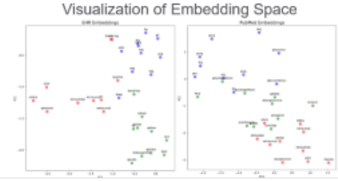
- **To train and evaluate ophthalmology domain-specific word vectors** using corpora from published ophthalmology literature and from electronic health records (EHR), comparing them to pre-trained general embeddings.
- **To predict the prognosis of low vision patients** using clinical free-text from EHR, in order to identify candidates for low vision rehabilitation services

Methods: Ophthalmology Embeddings

Text Corpora Construction

Corpora	Clean	Train	Embeddings
PubMed: MeSH: Eye Diseases, Ocular Physiological Phenomena, Ophthalmology, Ophthalmologic Surgical Procedures N=121,740 EHR: Random ophthalmology progress note from every ophthalmology patient N=89,282 notes	Text Processing: Lowercase; Tokenize (split into words); Remove stopwords (the, a, an, etc.); Remove words that occur less than 5 times in the corpus; PubMed Tokens: N=55,937; EHR Tokens: N=41,630	Word2Vec: Continuous-bag-of-words architecture: Predict the word given the context; Context window size +/-5 words	Output: 300-dimension word vectors where words used in similar contexts (assumed to have similar meaning) appear closely in vector space
Two different ophthalmology domain-specific corpora	Text pre-processed the same way for both corpora	Trained in separate models with the same specifications	Compared to GloVe : 300-d 42B vectors trained on Common Crawl

Visualization of Embedding Space



Intrinsic Evaluation: Creation of 200 Novel Ophthalmology Domain-Specific Analogies

Examples of ophthalmology domain-specific analogies			
dorzolamide	brinzolamide	bromfenac	ketorolac
limbus	limbal	canthus	canthal
ocp	conjunctiva	pseudoexfoliation	lens
left	right	os	od
brinzolamide	glaucoma	bevacizumab	amd
sclera	sclerotomy	iris	iridotomy

Methods: Predicting Low Vision Prognosis

Cohort Construction and Features

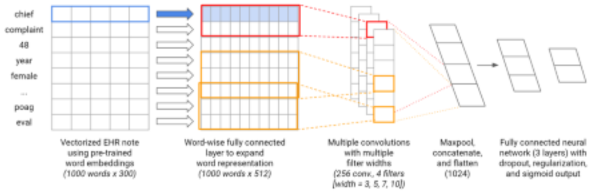
01 STARR with >4 million patients
- Free text notes
- Smartforms (where much of the eye examination information resides)
- Demographics, billing codes, meds, etc.

02 N=553,184 documented VA measurements, belonging to N=88,692 unique patients

03 ~14% of VA measurements were worse than 20/40, belonging to N=13847 unique patients. The first date of low vision was determined for each patient.

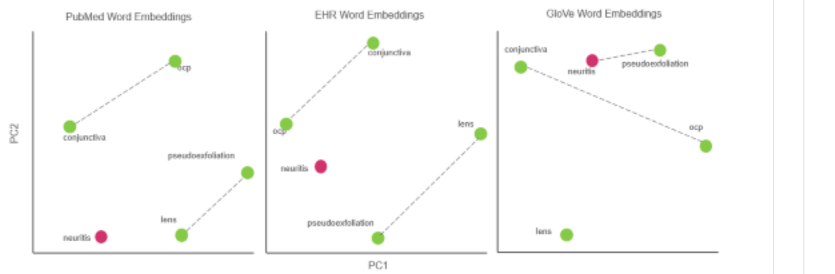
04 N=5612 patients with low vision had at least one year of follow-up, defined as ≥ one visit with documented VA ≥ 365 days from the initial low vision date

Convolutional Neural Network Model Architecture



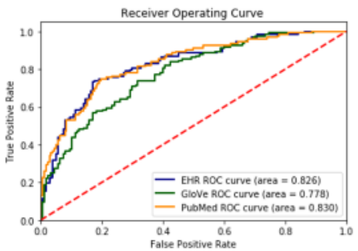
Classification Target: Patients who did not improve to better than 20/40 in 1 year (40.5% of population)

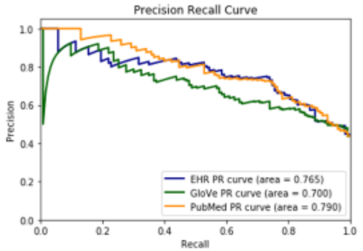
Results: Evaluation Using Analogies



Example word analogy depicted: ocp:conjunctiva::pseudoexfoliation: ???
Correct answer choice: "lens". **Incorrect answer choice**: "neuritis".
Cosine similarity between (pseudoexfoliation – ocp + conjunctiva) and either lens or neuritis.
PubMed and EHR word embeddings identifies the correct answer while GloVe chooses the wrong answer.
Overall Test Scores: PubMed: 95.0%; EHR: 86.0%; GloVe: 91.0 %

Results: Predicting Low Vision Prognosis





	F1	Sensitivity (Recall)	Specificity	PPV (Precision)	NPV	Accuracy
PubMed	0.73	0.76	0.78	0.7	0.82	0.77
EHR	0.72	0.77	0.76	0.68	0.82	0.76
GloVe	0.65	0.74	0.63	0.58	0.78	0.68

Conclusions

- For predictive tasks using highly domain-specific text, using domain-specific neural word embeddings may yield better performance than general word embeddings.
- **Training word embeddings using domain-specific published literature** abstracts is relatively easy and has good coverage even of text from electronic health records.
- We found that analyzing ophthalmology domain-specific word embeddings using analogies required **creation of ophthalmology domain-specific analogies.**
- Using ophthalmology domain-specific word embeddings, we were able to **predict the prognosis of low vision patients using clinical free text** with good performance.

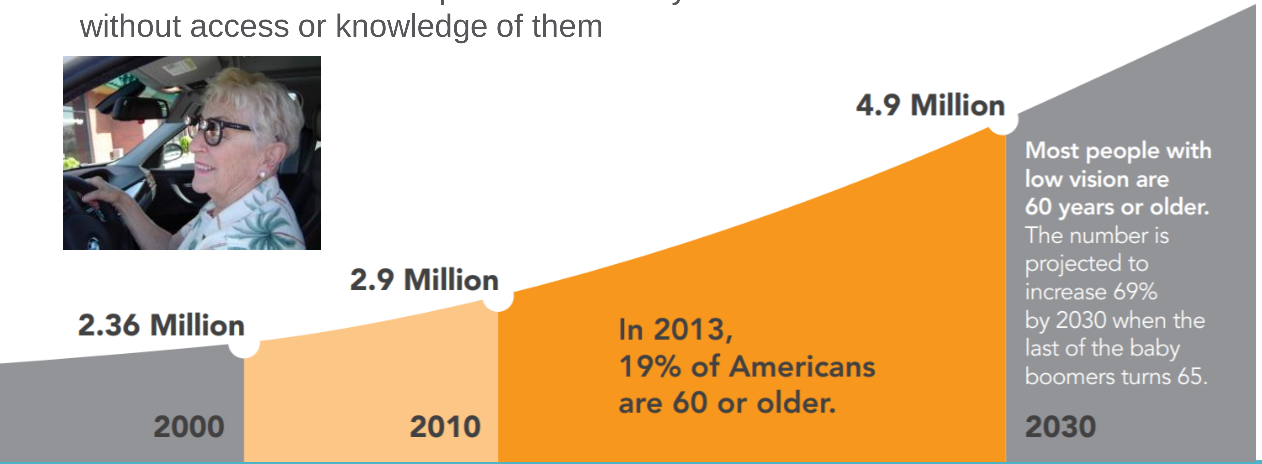
Email me at **sywang@stanford.edu**

Supported By Research to Prevent Blindness

Background

Low Vision and Rehabilitation Services

- **Low Vision**: Visual impairment that cannot be corrected with glasses, contacts, or other treatments. Prevalence increasing.
- **Severely impacts activities of daily living**: 3x unemployment, 3x MVC, if they can even drive, 3x depression and anxiety, 2x falls
- **Low vision rehabilitation services**: Palliative care for the eye, maximizing use of remaining vision
- **Problem**: Almost 90% of patients who may benefit from these services are without access or knowledge of them



Ophthalmology Progress Note
HPI: 56 yo F presents with decreased VA…
OcHx:
- Cataracts OU
- Macular Degeneration OU

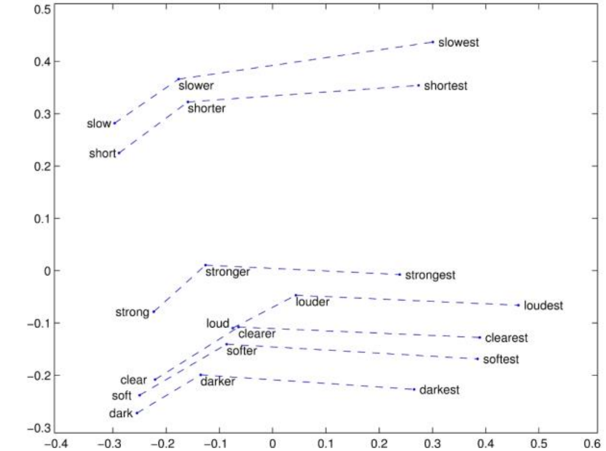
Gtt:
AREDS2 vitamins
Exam:

Vacc	OD: 20/40-2 ph 20/25-2
	OS: 20/200+1 ph 20/150-2
Tapp	OD: 16
	OS: 17

SLE:
L/L: wnl OU
C/S: w/q OU
K: clear OU
AC: d/q OU
Iris: r/r OU
Lens: 2+ NS with 1+CC OU
AntVx: wnl OU
DFE: CDR 0.5./0.5, macular with drusen OD and GA OS, normal vessels/periphery OU
A/P
1) Dry macular degeneration OU…

Electronic Health Records (EHR)

- So much data in EHR
- How to make use of free text?
- Turn the text into numbers via word embeddings



Problem: General word embeddings don't cover highly domain-specific text

Purpose

- **To train and evaluate ophthalmology domain-specific word vectors** using corpora from published ophthalmology literature and from electronic health records (EHR), comparing them to pre-trained general embeddings.
- **To predict the prognosis of low vision patients** using clinical free-text from EHR, in order to identify candidates for low vision rehabilitation services

Methods: Ophthalmology Embeddings

Text Corpora Construction

PubMed
MeSH: Eye Diseases, Ocular Physiological Phenomena, Ophthalmology, Ophthalmologic Surgical Procedures

N=121,740

EHR
Random ophthalmology progress note from every ophthalmology patient

N=89,282 notes

Corpora
Two different ophthalmology domain-specific corpora

Text Processing
Lowercase

Tokenize (split into words)

Remove stopwords (the, a, an, etc.)

Remove words that occur less than 5 times in the corpus

PubMed Tokens: N=55,937

EHR Tokens: N=41,630

Clean
Text pre-processed the same way for both corpora

Word2Vec
Continuous-bag-of-words architecture:

Predict the word given the context

Context window size +/-5 words

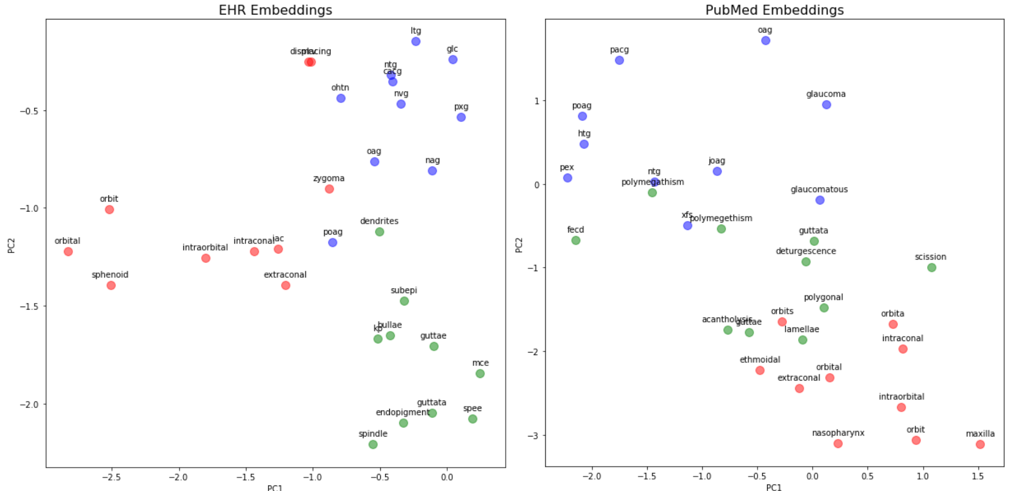
Train
Trained in separate models with the same specifications

Output
300-dimension word vectors where words used in similar contexts (assumed to have similar meaning) appear closely in vector space

Embeddings
Compared to GloVe :

300-d 42B vectors trained on Common Crawl

Visualization of Embedding Space



Intrinsic Evaluation: Creation of 200 Novel Ophthalmology Domain-Specific Analogies

Examples of ophthalmology domain-specific analogies

dorzolamide	brinzolamide	bromfenac	ketorolac
limbus	limbal	canthus	canthal
ocp	conjunctiva	pseudoexfoliation	lens
left	right	os	od
brinzolamide	glaucoma	bevacizumab	amd
sclera	sclerotomy	iris	iridotomy

Methods: Predicting Low Vision Prognosis

Cohort Construction and Features

01

STARR with >4 million patients
- Free text notes
- Smartforms (where much of the eye examination information resides)
- Demographics, billing codes, meds, etc.

02

N=553,184 documented VA measurements, belonging to N=88,692 unique patients

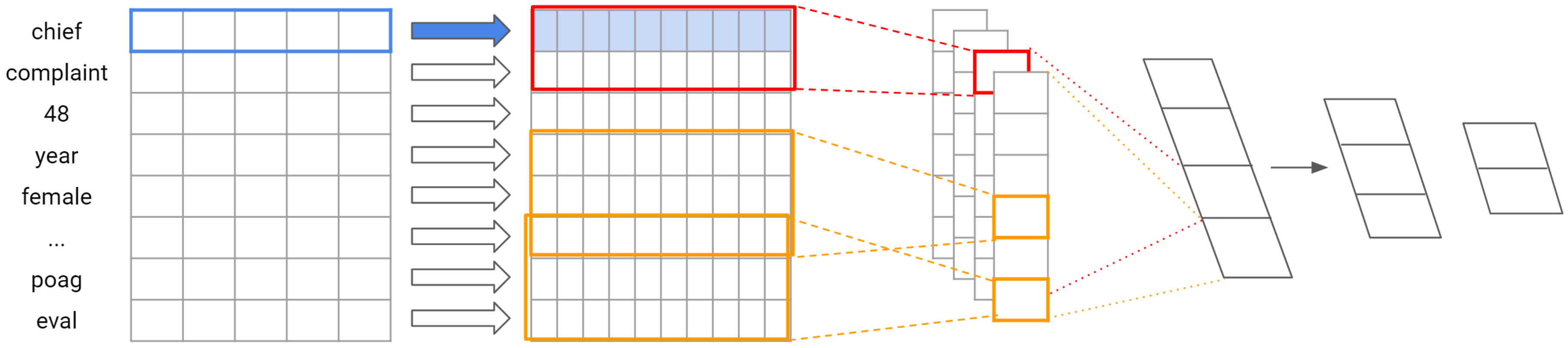
03

~14% of VA measurements were worse than 20/40, belonging to N=13847 unique patients.
The first date of low vision was determined for each patient.

04

N=5612 patients with low vision had at least one year of follow-up, defined as ≥ one visit with documented VA ≥ 365 days from the initial low vision date

Convolutional Neural Network Model Architecture



Vectorized EHR note using pre-trained word embeddings *(1000 words x 300)*

Word-wise fully connected layer to expand word representation *(1000 words x 512)*

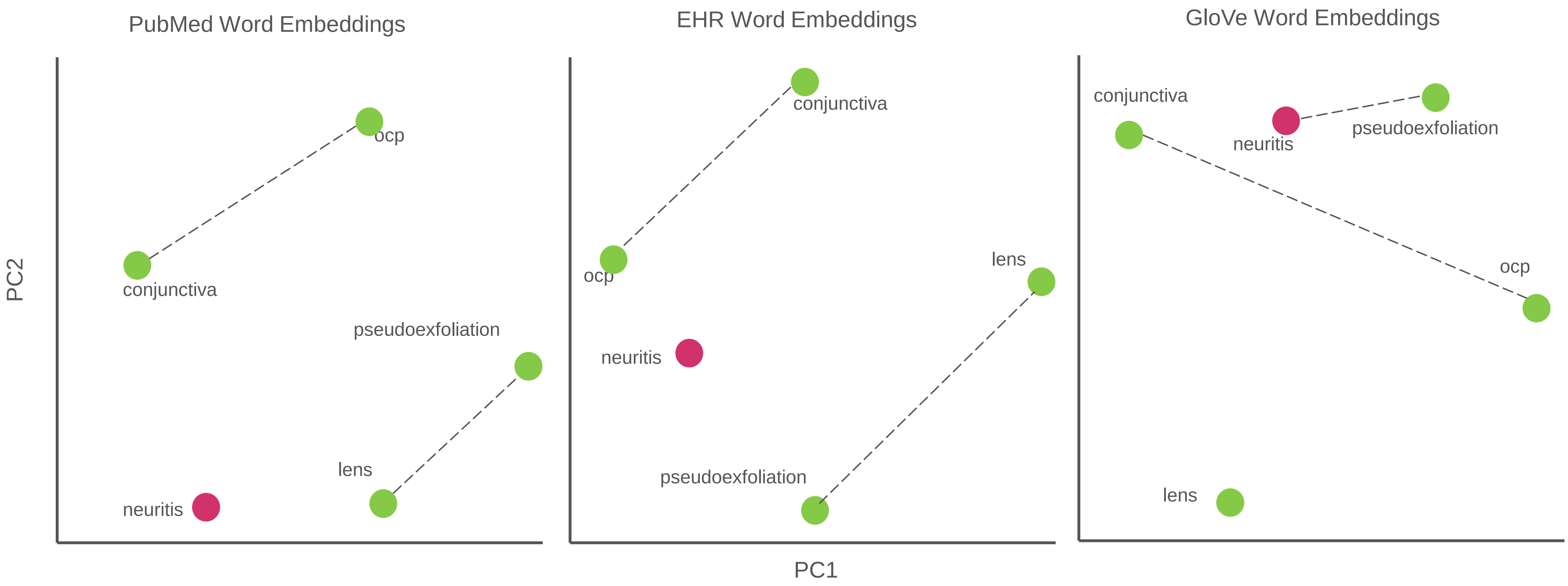
Multiple convolutions with multiple filter widths *(256 conv., 4 filters [width = 3, 5, 7, 10])*

Maxpool, concatenate, and flatten (1024)

Fully connected neural network (3 layers) with dropout, regularization, and sigmoid output

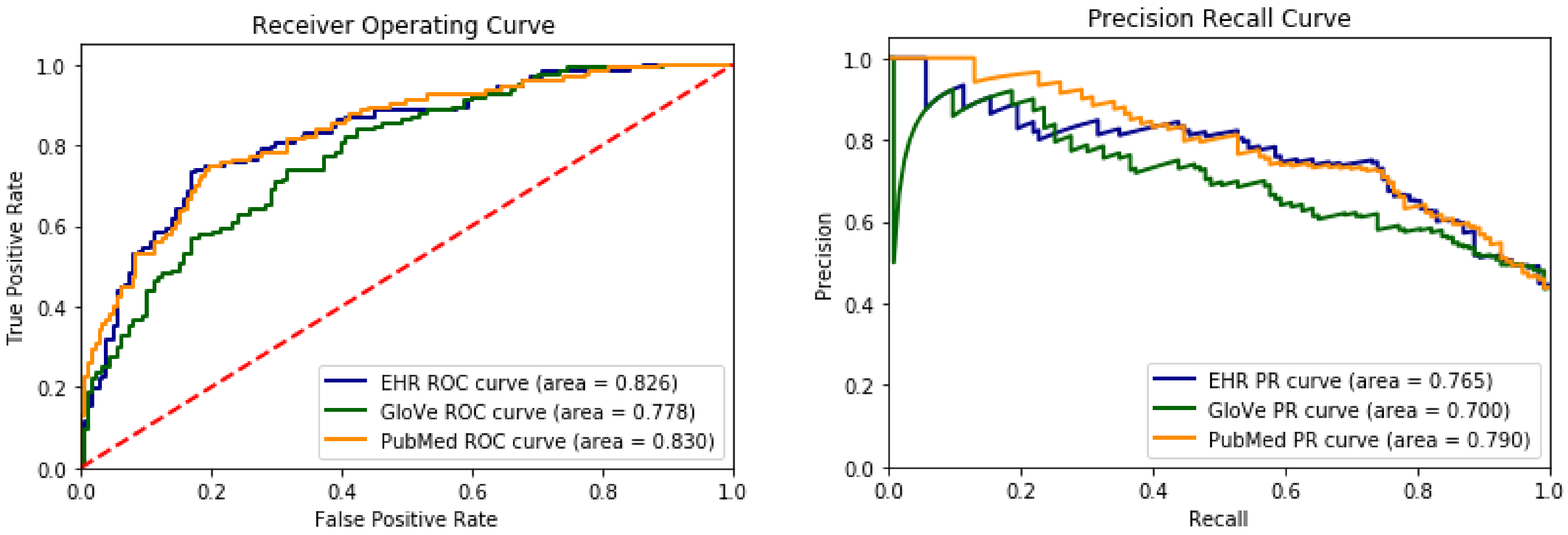
Classification Target: Patients who did not improve to better than 20/40 in 1 year
(40.5% of population)

Results: Evaluation Using Analogies



Example word analogy depicted: ocp:conjunctiva::pseudoexfoliation: ???
Correct answer choice: “lens”. **Incorrect answer choice**: “neuritis”.
Cosine similarity between (pseudoexfoliation – ocp + conjunctiva) and either lens or neuritis.
PubMed and EHR word embeddings identifies the correct answer while GloVe chooses the wrong answer.
Overall Test Scores: PubMed: 95.0%; EHR: 86.0%; GloVe: 91.0 %

Results: Predicting Low Vision Prognosis



	F1	Sensitivity (Recall)	Specificity	PPV (Precision)	NPV	Accuracy
PubMed	0.73	0.76	0.78	0.7	0.82	0.77
EHR	0.72	0.77	0.76	0.68	0.82	0.76
GloVe	0.65	0.74	0.63	0.58	0.78	0.68

Conclusions

- For predictive tasks using highly domain-specific text, using domain-specific neural word embeddings may yield better performance than general word embeddings.
- **Training word embeddings using domain-specific published literature** abstracts is relatively easy and has good coverage even of text from electronic health records.
- We found that analyzing ophthalmology domain-specific word embeddings using analogies required **creation of ophthalmology domain-specific analogies**.
- Using ophthalmology domain-specific word embeddings, we were able to **predict the prognosis of low vision patients using clinical free text** with good performance.