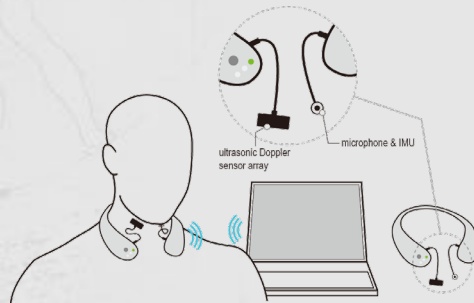




초음파 도플러 센서 기반 삼킴 장애 판별 모델 개발



김민재¹, 박종배¹, 최승현², 유희천¹

¹포항공과대학교 산업경영공학과 인간공학설계기술 연구실

²포항공과대학교 산업경영공학과 품질시스템 연구실

Contents

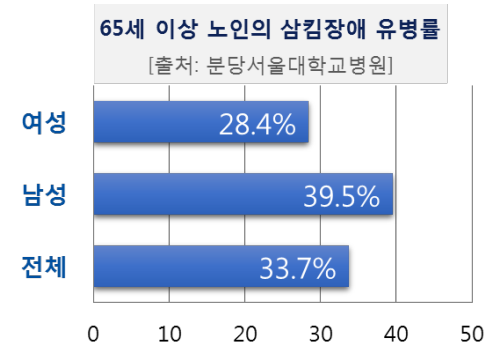
- **서론**
 - 연구 배경 및 필요성
 - 연구 목적
- **삼킴 시 초음파 도플러 센서 신호**
 - 정상인, 환자 삼킴 신호 측정
 - 삼킴 신호 measure 추출
- **삼킴 장애 판별 모델**
 - 정상인, 환자 신호 경향 파악
 - 판별 모델 개발
- **토의**

노인성 삼킴장애 (Presbyphagia)

- 노화에 따른 삼킴 관련 신체 구조 및 기능의 퇴화에 기인한 삼킴 기능의 감퇴
- 구강 단계, 인두 단계, 식도 단계에 걸친 광범위한 현상



노인성 삼킴장애 유병률
11%~35%



정상

삼킴장애



연구자	국가	n	연령대	연구방법	유병률
Bloem et al. (1990)	네덜란드	130	87세 이상	설문지 평가	16%
Sven et al. (1991)	스웨덴	476	50세 이상	설문지 평가	35%
Kawashima et al. (2004)	일본	1313	65세 이상	설문지 평가	13.8%
Roy et al. (2007)	미국	117	65세 이상	설문지 평가	33%
Holland et al. (2011)	영국	634	69세 이상	설문지 평가	11.4%
Serr-Prat et al. (2011)	스페인	254	70세 이상	V-VST	27.2%
Yang et al. (2013)	한국	415	65세 이상	설문지 평가	33.7%

* V-VST: Volume-Viscosity Swallow Test

삼킴 기능에 대한 비침습적 & 객관적 측정 필요

- 현재 수행되고 있는 **비디오 투시 조영 삼킴 검사(VFSS, Videofluoroscopic Swallowing Study)**는 방사능에 노출되는 침습적 방법으로 지속적인 모니터링을 위한 평가 도구로 사용할 경우 **피측정자의 건강을 저해할 수 있음**
- **삼킴 기능 설문 평가**의 경우 직접적 관찰에 의해 평가되지 않아 **객관성 부족**
⇒ **비침습적이고 객관적인 측정 방법**의 활용이 필요함

비디오 투시 조영 삼킴 검사



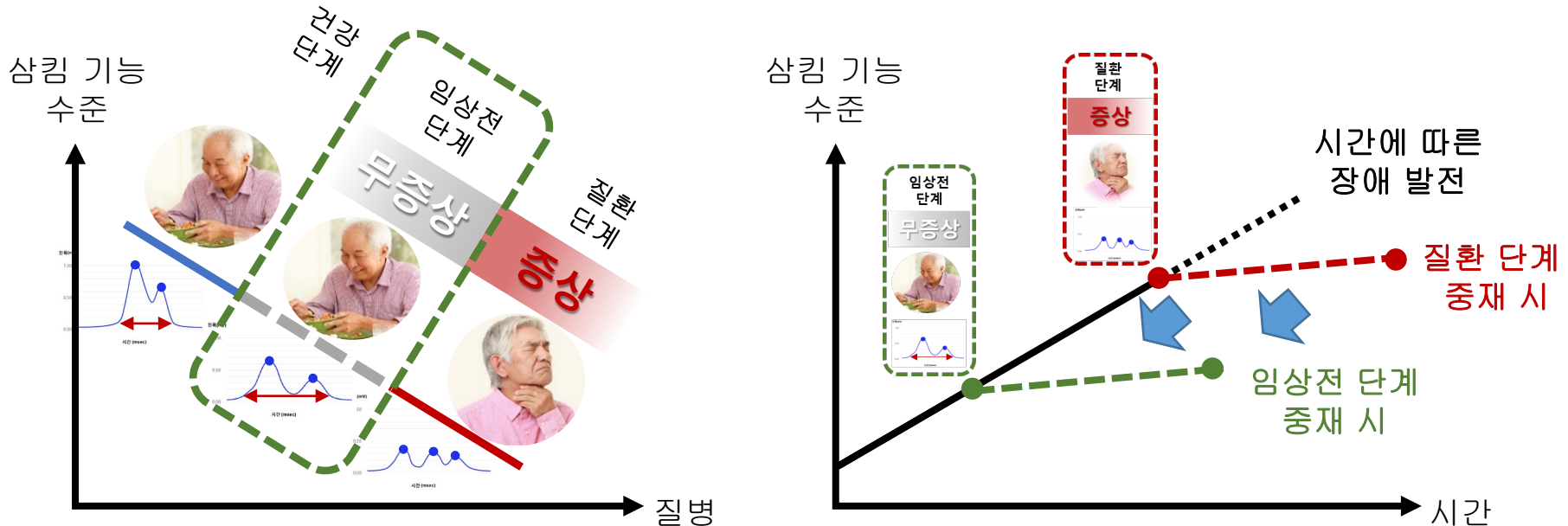
삼킴 기능 설문 평가



삼킴장애 조기 탐지 및 중재 필요

□ 증상이 드러나지 않는 **임상전 단계**에서 삼킴 문제를 조기 탐지 및 중재를 통해 **삼킴 장애로의 발전을 방지하거나 장애 정도를 낮출 수 있음**

- ✓ 진행 단계의 시기 및 속도에 대한 지표를 마련하여 **조기 선별 근거 구축** 필요
- ✓ 노화 및 퇴행성 뇌질환의 삼킴 특성을 파악하여 **삼킴 기능 저하의 원인 규명** 필요



- 조기 선별 근거 구축
- 삼킴 기능 저하의 원인 규명

삼킴 장애 판별 모델 개발 필요성

- 선행 연구(Choi et al., 2020)는 삼킴 측정에 특화된 장비로서 **ultrasonic Doppler**를 사용하여 **삼킴 관련 기관의 움직임을 측정**하는 장비 개발
- ⇒ 측정된 **삼킴 signal**을 분석하여 의미를 파악하고 **삼킴을 정량화**하여 **삼킴 장애 평가에 특화**하는 연구 필요

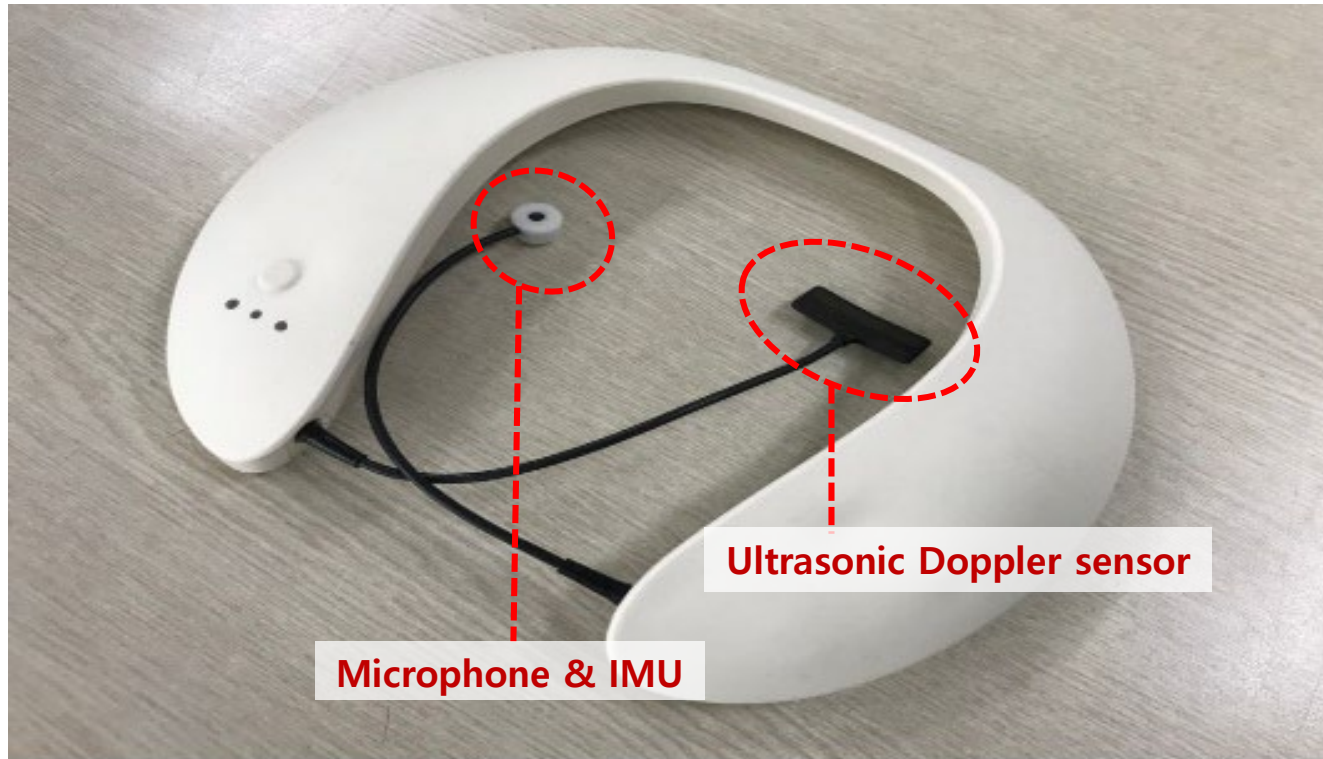


초음파 도플러 센서 기반 삼킴 장애 판별 모델 개발

1. 초음파 도플러 센서 기반 삼킴 신호 **측정 장비**
2. 정상인 및 환자 대상 **삼킴 data 측정**
3. 초음파 도플러 센서 기반 **삼킴 장애 판별 모델 구축**

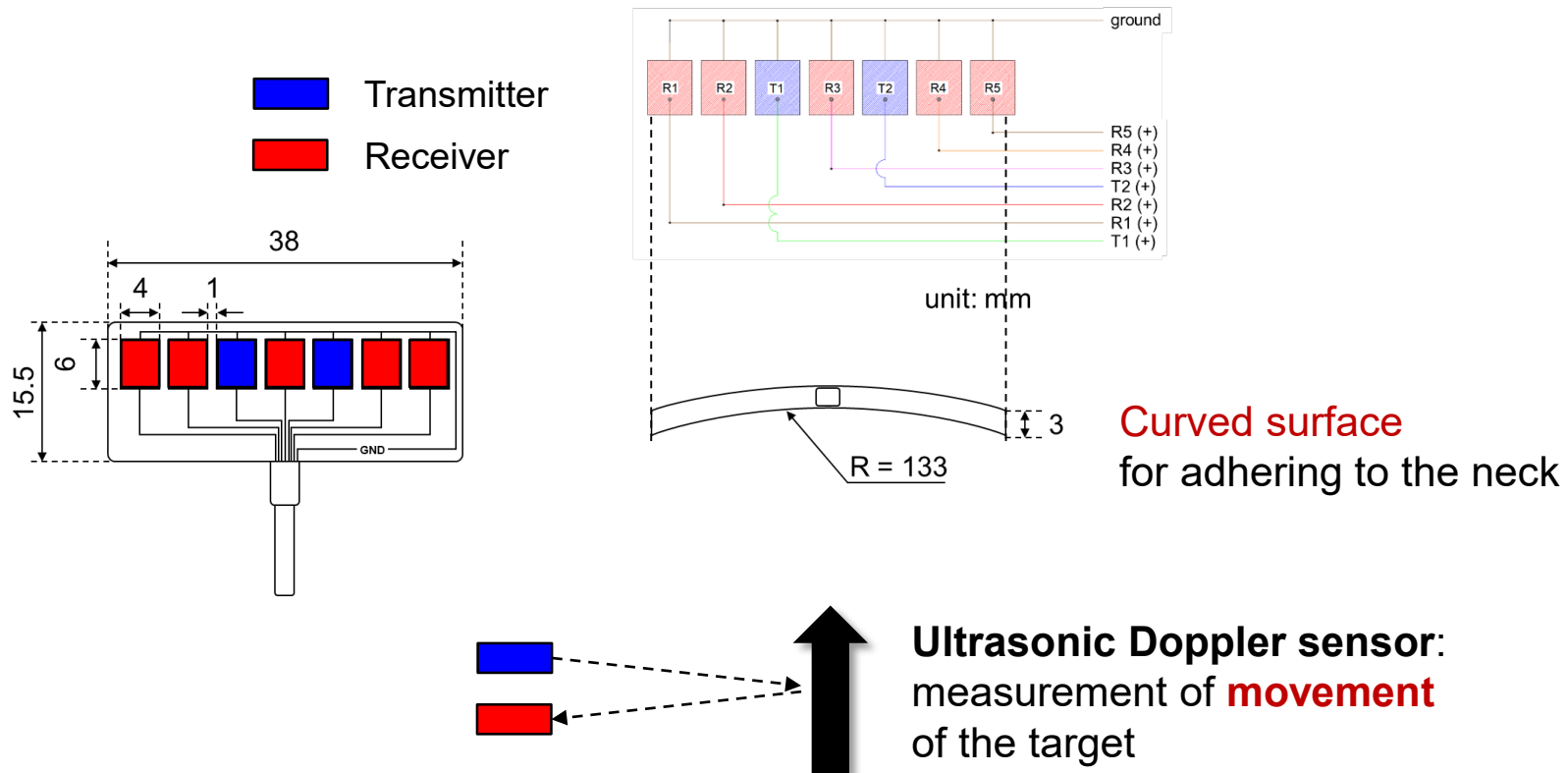
Ultrasonic Doppler sensor기반 삼킴 측정 장비

- ❑ **Ultrasonic Doppler sensor**, microphone 및 IMU 탑재
- ❑ 적절한 너비로 벌려 쉽게 착용할 수 있는 유연한 재질의 **어깨 거치형 기구 설계**



Ultrasonic Doppler sensor기반 삼킴 측정 장비

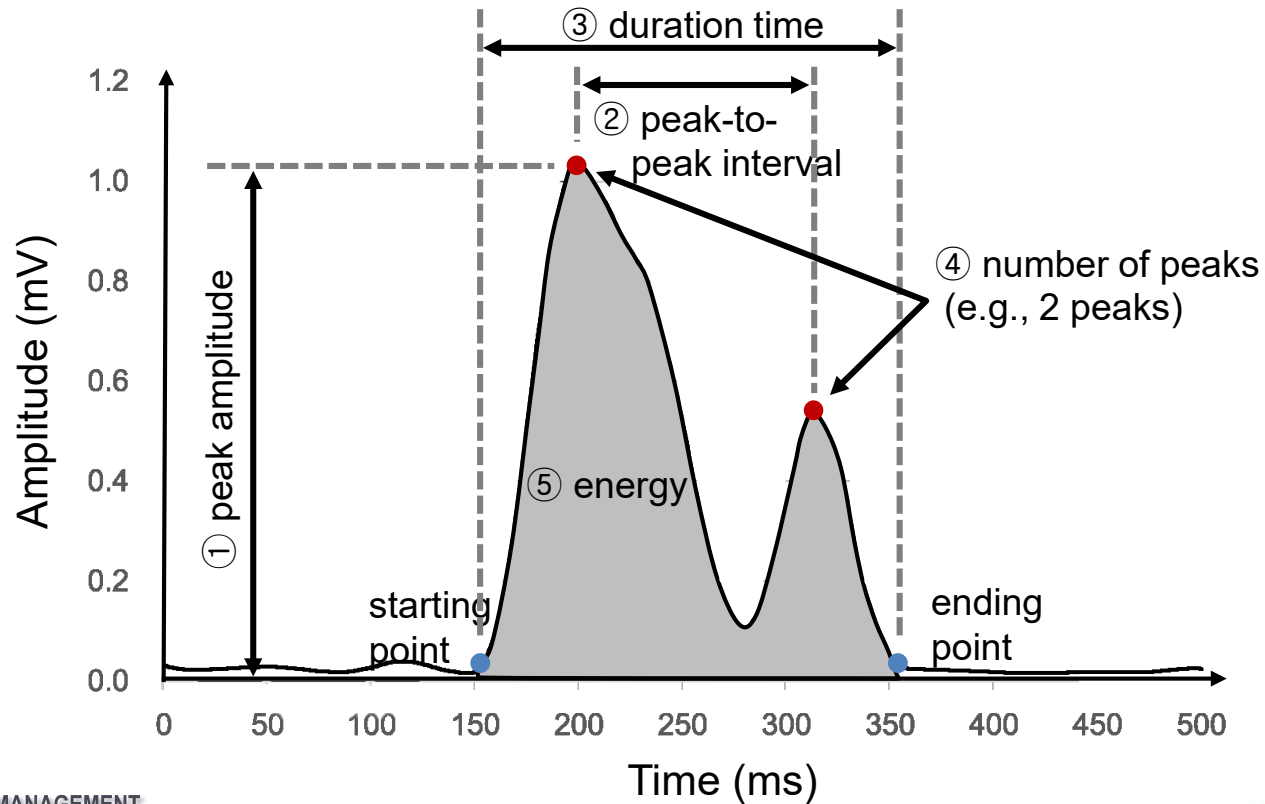
- Ultrasonic Doppler sensor를 통해 대상의 움직임 측정



SMAS 신호의 정량화 척도

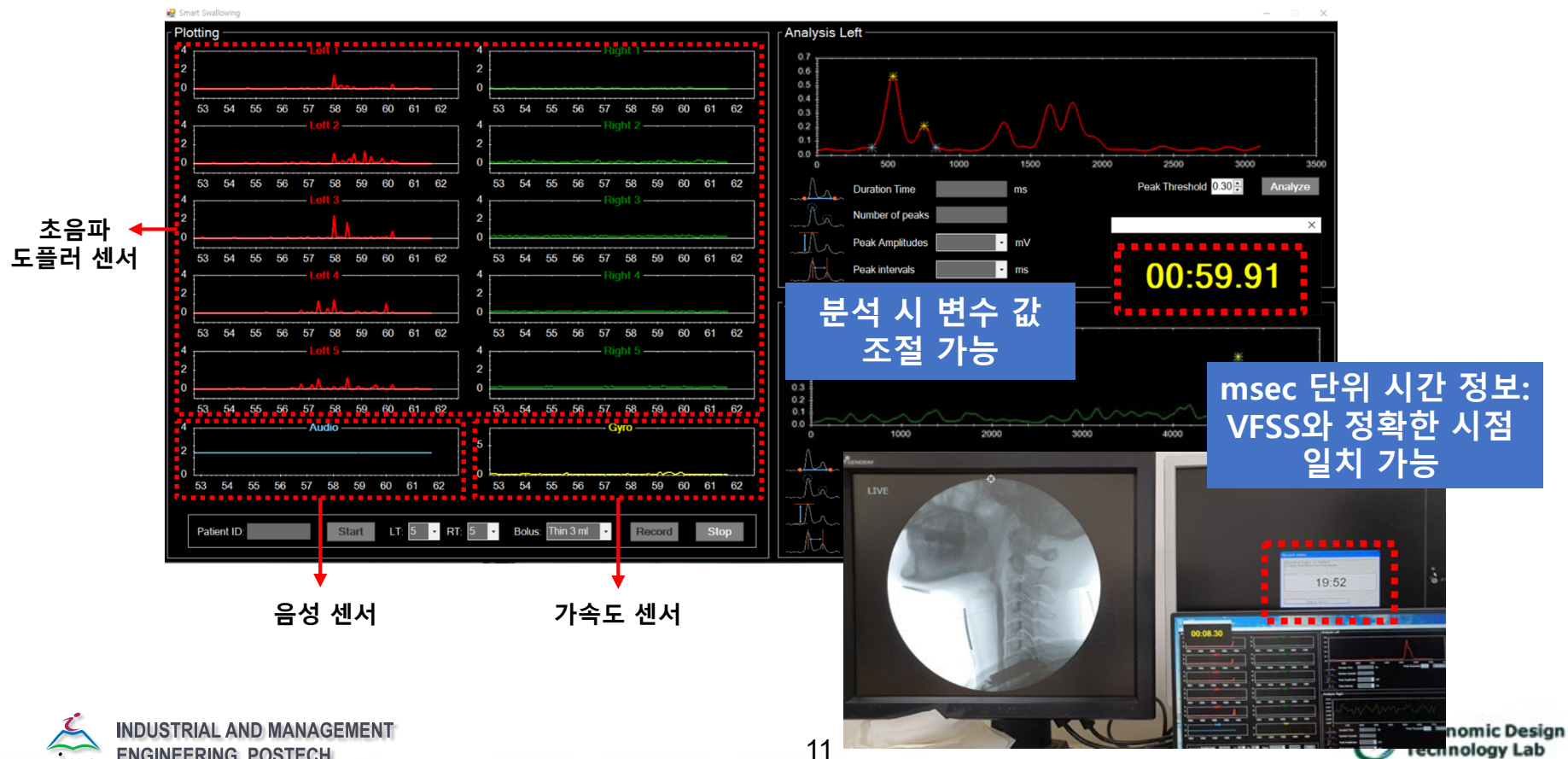
□ 추출된 peaks, starting/ending points를 참조하여 **5가지 삼킴 정량화 척도 정립**

1. Peak amplitude
2. Number of peaks
3. Peak-to-peak interval
4. Energy
4. Duration time



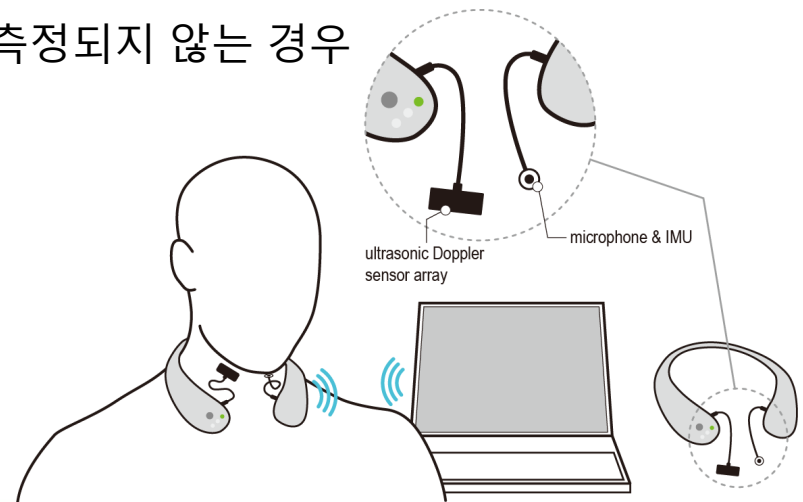
삼킴 측정 Software

- 초음파 도플러 센서, 음성 센서, 가속도 센서의 **실시간 데이터 시각화**
- 분석 범위 및 starting/ending points를 직접 click하여 지정 가능
- 실시간 측정 시 msec 단위 시간 정보를 제공하여 VFSS와 연계 분석이 용이



정상인 대상 실험

- 실험 참여자: $n = 300$ (정상인)
- 측정 data: SMAS data, 나이, 성별
- 평가 내용(3회 반복)
 - 물 3 ml, 9 ml
 - 요플레 3 ml, 9 ml
- 실험 평가 제외 기준
 - CV(Coefficient of variation) 확인을 통해 반복성이 낮게 확인되는 경우 (0.3 기준)
 - 삼킴 능력의 현저한 저하로 인해 신호가 거의 측정되지 않는 경우



Experimental Design

- 삼킴 종류 및 용량: 4가지
 - 1) Thin liquid (생수): 3, 9 ml
 - 2) Thick liquid (요플레): 3, 9 ml
- 반복: 삼킴 종류 및 용량 별 3회
- 총 실험 회수: 12회 (= 4가지 x 3회)
- 장비: Ultrasonic Doppler sensor-based system (SMAS-X05)

SMAS-X05



Experimental Procedure

※ 총 실험시간: 15 min

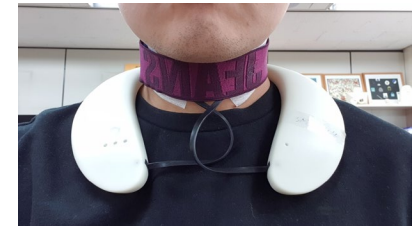
S1. Informed consent
(3 min)

- 실험 목적 설명
- 실험 참여 동의서 작성



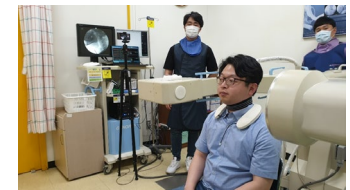
S2. Exercise
(5 min)

- SMAS 착용 및 신호 확인



S3. Swallowing session
(5 min)

- 각 삼킴에 SMAS 3회씩 측정
- 실험 진행 중 SMAS 신호 발생 여부 확인



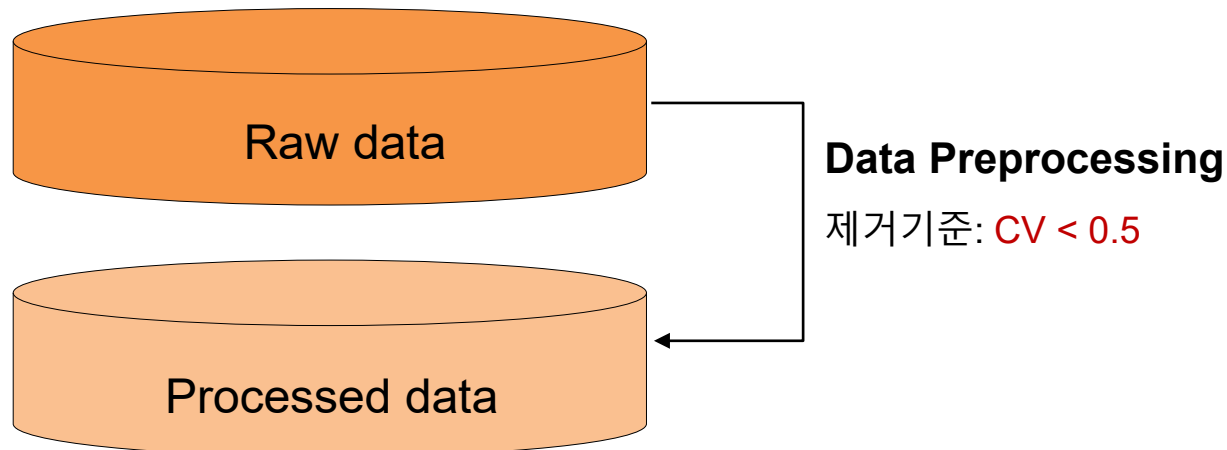
S4. Debriefing
(2 min)

- 삼킴 신호 확인
- 실험 진행 중 이상 여부 확인



Data Preprocessing

- 3회 반복 삼킴에 대한 **CV**(coefficient variance) **확인**
- 실험 평가 제외 기준
 - CV 값 확인을 통해 반복성이 낮게 확인되는 경우 (0.5 기준)
 - 삼킴 능력의 현저한 저하로 인해 신호가 거의 측정되지 않는 경우



SMAS Data 분석: Viscosity, Volume

□ 선행 연구, 정상인, 환자의 증감 경향 일치 확인

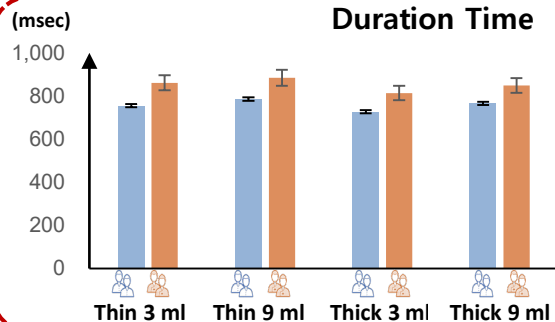
- Viscosity ↑ ⇒ peak amplitude, duration time ↓
- Volume ↑ ⇒ peak amplitude, duration time ↑

Measure	Factor	경향				비고
		이백희 선행 연구 (n = 120)	최영근 선행 연구 (n = 24)	현재 연구		
				정상인 (n = 300)	환자 (n = 60)	
Peak amplitude	Viscosity 증가	감소(약 20%)	감소(26.2%)	감소(9.4%)	감소(7.2%)	일치
	Volume 증가	증가(약 20%)	증가(12.3%)	증가(2.7%)	증가(7.6%)	일치
Duration time	Viscosity 증가	-	감소(1.4%)	감소(3.1%)	감소(4.8%)	일치
	Volume 증가	-	증가(5.6%)	증가(4.6%)	증가(3.5%)	일치

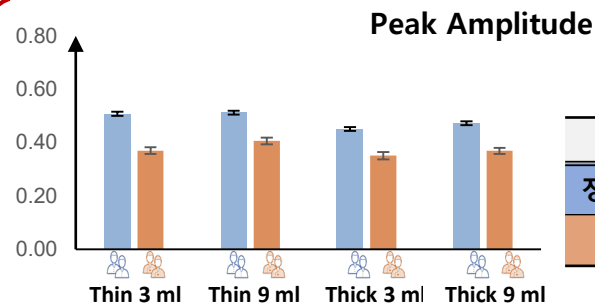
SMAS Data 분석: 정상인 vs. 삼킴 장애 환자

□ Duration time, peak amplitude, energy에서 정상인과 환자의 차이 확인

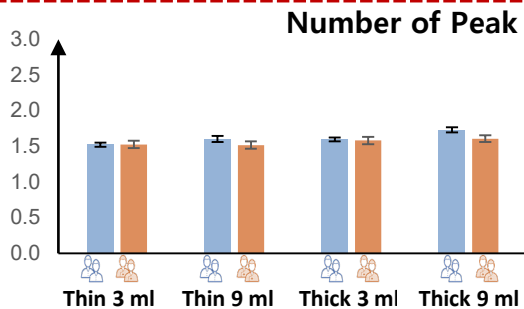
□ Peak amplitude와 energy의 경향 유사



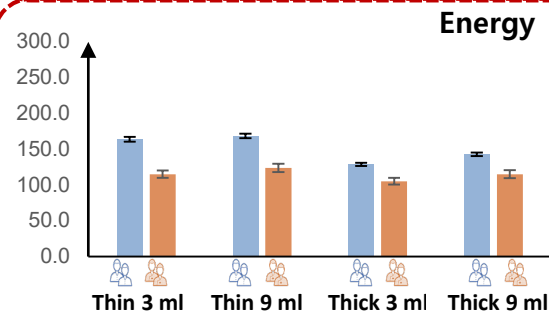
	점도 ↑	양 ↑
정상인	↓ -3.1%	↑ 4.6%
환자	↓ -4.8%	↑ 3.5%



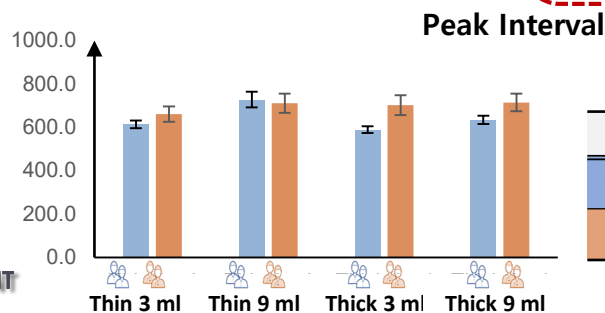
	점도 ↑	양 ↑
정상인	↓ -9.4%	↑ 2.7%
환자	↓ -7.2%	↑ 7.6%



	점도 ↑	양 ↑
정상인	↑ 6.4%	↑ 6.8%
환자	↑ 4.8%	↑ 0.6%



	점도 ↑	양 ↑
정상인	↓ -18.3%	↑ 6.5%
환자	↓ -7.8%	↑ 8.5%



	점도 ↑	양 ↑
정상인	↓ -8.8%	↑ 13.3%
환자	↑ 3.3%	↑ 4.6%

삼킴 장애 판별 모델

□ Machine learning 기반 삼킴 장애 여부 판별 모델 구축 및 성능 비교

➤ Data

- ✓ 정상인 300명, 삼킴 장애 환자 60명의 SMAS data

➤ Variables

- ✓ Independent: SMAS 척도 5가지(duration time, peak amplitude, number of peaks, impulse, peak interval)
- ✓ Dependent: 삼킴 장애 환자 여부

➤ 적용 model

- ✓ Random forests
- ✓ XGBoost
- ✓ Logistic regression
- ✓ Support vector machines (SVM)
- ✓ CNN

판별 모델 성능 지표

□ Machine learning 기반 삼킴 장애 판별 모델 6가지 성능 비교 지표

- **정확도 (Accuracy):** 전체 판별 중 모델이 정확한 판별을 수행하는 비율

$$Accuracy = \frac{True\ Positive + True\ Negative}{Total\ Predictions}$$

- **민감도 (Sensitivity):** 실제로 환자인 경우 중에 환자로 판별된 비율

$$Sensitivity = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

- **특이도 (Specificity):** 실제로 정상인인 경우 중에 정상인으로 판별된 비율

$$Specificity = \frac{True\ Negative}{True\ Negative + False\ Positive}$$

- **정밀도 (Precision):** 모델이 환자로 판별한 경우 중 실제 환자의 비율

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

- **F1-Score:** precision와 sensitivity의 조화 평균

$$F1 = 2 \times \frac{Precision \times Sensitivity}{Precision + Sensitivity}$$

- **ROC-AUC:** 실제 환자와 정상인을 구분하는지 나타내는 척도

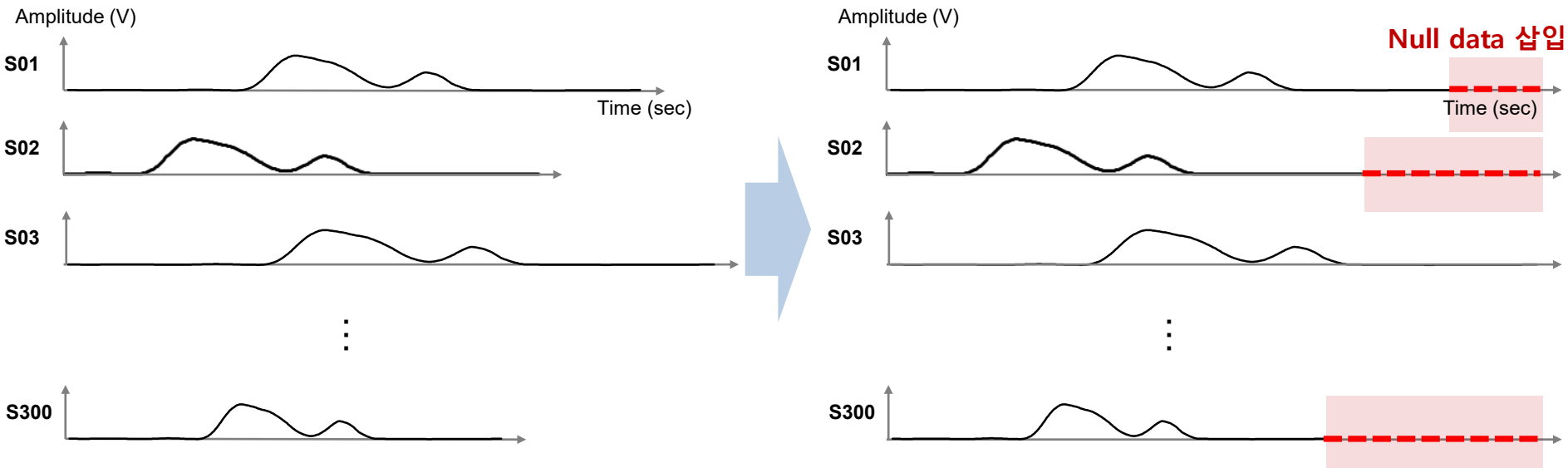
- 삼킴 장애 여부 판별 모델 구축을 위한 **data set**
 - SMAS data
 - ✓ 정상인 삼킴 3600개 (300명 × 12번), 환자 삼킴 720개 (60명 × 12번)
 - ✓ 1명당 12개의 삼킴 data (2 viscosity (thin, thick) × 2 volume (3ml, 9ml) × 3회)
 - ✓ 정상인 data는 60, 120, 180, 300의 4가지 조합으로 구성
 - Data preprocessing
 - ✓ 동일 조건 3회 반복 삼킴에 대해 CV(coefficient of variation) 0.5 이상일 경우 이상치로 제거
 - ✓ SMAS 척도에 대해 min-max normalization 수행
 - Random data split
 - ✓ Training data: 70%
 - ✓ Test data: 30%

판별 모델 Data Preprocessing

□ CNN(Convolutional neural network) model 적용

- Input data: 삼킴 시 나타나는 파형 image
- Data preprocessing - padding
 - ✓ **신호 길이 통일**: 측정된 삼킴 신호마다 신호의 길이가 달라 **가장 긴 신호를 기준으로 통일**
 - ✓ **Null data 삽입**: 모든 data의 길이를 기준에 맞추기 위해 **null data 삽입**

Padding 수행 예시



판별 모델 결과 (1/4)

- 전체 model 중 XGBoost와 Logistic regression이 가장 우수한 성능을 나타냄

Model	Data sets		Accuracy	Sensitivity	Specificity	F1-Score	ROC-AUC
	삼킴 장애 환자	정상인					
Random Forest	60	60	0.82	0.74	0.86	0.73	0.86
XGBoost			0.85	0.74	0.89	0.75	0.85
Logistic regression			0.85	0.74	0.89	0.84	0.75
SVM			0.71	0.66	0.74	0.77	0.60
CNN			0.70	0.47	0.90	0.79	0.60
Random Forest	60	120	0.72	0.21	0.94	0.85	0.31
XGBoost			0.74	0.26	0.94	0.85	0.38
Logistic regression			0.74	0.26	0.94	0.84	0.38
SVM			0.75	0.17	0.99	0.77	0.30
CNN			0.68	0.64	0.70	0.74	0.57
Random Forest	60	180	0.80	0.13	0.99	0.85	0.22
XGBoost			0.81	0.18	1.00	0.87	0.30
Logistic regression			0.81	0.18	1.00	0.84	0.30
SVM			0.79	0.08	1.00	0.76	0.14
CNN			0.74	0.20	0.91	0.73	0.27
Random Forest	60	300	0.86	0.00	1.00	0.84	0.00
XGBoost			0.86	0.03	0.99	0.81	0.05
Logistic regression			0.86	0.03	0.99	0.78	0.05
SVM			0.86	0.00	1.00	0.77	0.00
CNN			0.78	0.25	0.88	0.67	0.27

판별 모델 결과 (2/4)

- 환자 60명, 정상인 60명 data set에서 가장 좋은 성능이 나타남
 - 다른 data set에서 accuracy는 높은 편이지만 sensitivity가 낮음
 - 정상인 수가 늘어날 수록 sensitivity가 낮아짐

Model	Data sets		Accuracy	Sensitivity	Specificity	F1-Score	ROC-AUC
	삼킴 장애 환자	정상인					
Random Forest	60	60	0.82	0.74	0.86	0.73	0.86
XGBoost			0.85	0.74	0.89	0.75	0.85
Logistic regression			0.85	0.74	0.89	0.84	0.75
SVM			0.71	0.66	0.74	0.77	0.60
CNN			0.70	0.47	0.90	0.79	0.60
Random Forest	60	300	0.86	0.00	1.00	0.84	0.00
XGBoost			0.86	0.03	0.99	0.81	0.05
Logistic regression			0.86	0.03	0.99	0.78	0.05
SVM			0.86	0.00	1.00	0.77	0.00
CNN			0.78	0.25	0.88	0.67	0.27

판별 모델 결과 (3/4)

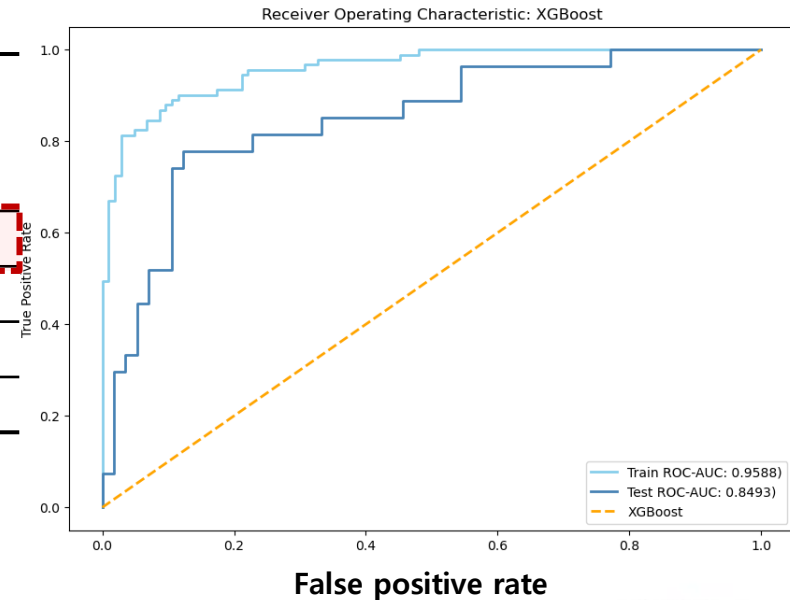
□ 환자 60명, 정상인 60명 data set의 XGBoost에서 가장 좋은 성능이 나타남

- Accuracy: 0.85
- Sensitivity: 0.74
- F1-Score: 0.75
- Precision: 0.77
- Specificity: 0.89

XGBoost 모델 성능

True positive rate

Model	Data sets		Accuracy	Sensitivity	Specificity	F1-Score	ROC-AUC
	환자	정상인					
XGBoost	60	60	0.85	0.74	0.89	0.75	0.85
		120	0.74	0.26	0.94	0.85	0.38
		180	0.81	0.18	1.00	0.87	0.30
		300	0.86	0.03	0.99	0.81	0.05



판별 모델 결과 (4/4)

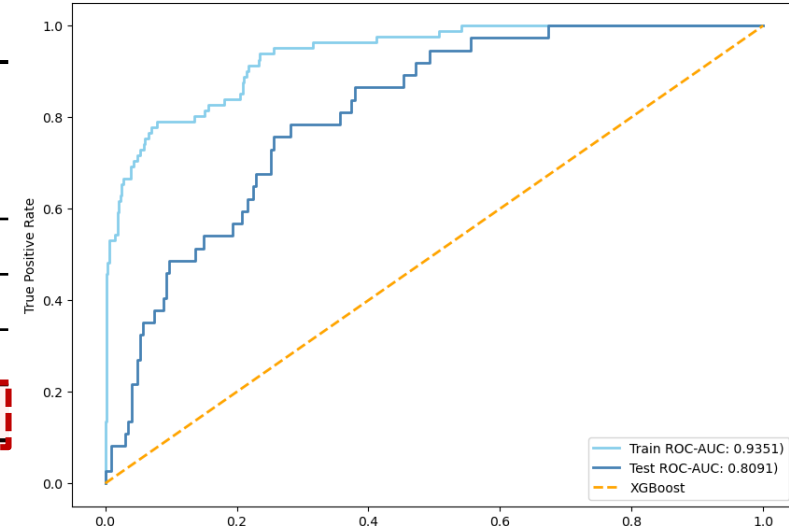
□ 환자 60명, 정상인 300명 data set의 경우 전반적인 모델 성능이 낮게 나타남

- Accuracy: 0.86
- Sensitivity: 0.03
- F1-Score: 0.05
- Precision: 0.33
- Specificity: 0.99

XGBoost 모델 성능

True positive rate

Receiver Operating Characteristic: XGBoost



False positive rate

Model	Data sets		Accuracy	Sensitivity	Specificity	F1-Score	ROC-AUC
	환자	정상인					
XGBoost	60	60	0.85	0.74	0.89	0.75	0.85
		120	0.74	0.26	0.94	0.85	0.38
		180	0.81	0.18	1.00	0.87	0.30
		300	0.86	0.03	0.99	0.81	0.05

Discussion (1/2)

□ 3가지 measure(peak amplitude, energy, duration time)에서 삼킴 장애 환자와 정상인 간의 유의한 차이 파악

- ✓ 삼킴 장애 환자가 정상인보다 duration time이 길게 나타남
- ✓ 삼킴 장애 환자의 peak amplitude와 energy가 정상인보다 낮게 나타남

□ 5가지 machine learning 모델을 통해 삼킴 판별 모델 성능 확인 및 비교

- ✓ Random forest, logistic regression, SVM, XGBoost, CNN을 통한 **삼킴 판별 모델 제안**
- ✓ **정상인과 환자의 수**를 60명으로 **동일**하게 조정하였을 때 **높은 성능** 확인
- ✓ Data set에 관계없이 **XGBoost에서 높은 판별 성능**이 나타남
- ✓ 정상인이 많은 경우 **data 불균형**으로 **sensitivity가 낮게** 나타남

Discussion (2/2)

□ 한계점

- ✓ 정상인 수(300)에 비해 환자 수(60)가 상대적으로 적어 **data 불균형 발생**하여 sensitivity, precision 등이 낮게 나타남
- ✓ 환자 data는 삼킴 장애의 중증도 또는 원인에 관계없이 **삼킴 장애 발생 여부에 대해서만 확인**하고 삼킴 장애 환자로 분류하여 수집됨

□ 추후 연구

- ✓ **삼킴 장애 증상** 및 **원인**별로 발생할 수 있는 **SMAS 신호 특성**에 대한 연구 필요
- ✓ SMAS를 활용하여 삼킴 장애를 판별하기 위해 **다수의 정상인과 환자**를 대상으로 측정된 **normative data** 구축 수행

Q&A

경청해 주셔서 감사합니다.

