LEARNING INFLAMMATORY BIOMARKERS FROM NOCTURNAL BREATHING, BMI AND DEMOGRAPHICS

Hao He, Dina Katabi
Computer Science and Artificial Intelligence Lab
Massachusetts Institute of Technology
{haohe,dk}@mit.edu

ABSTRACT

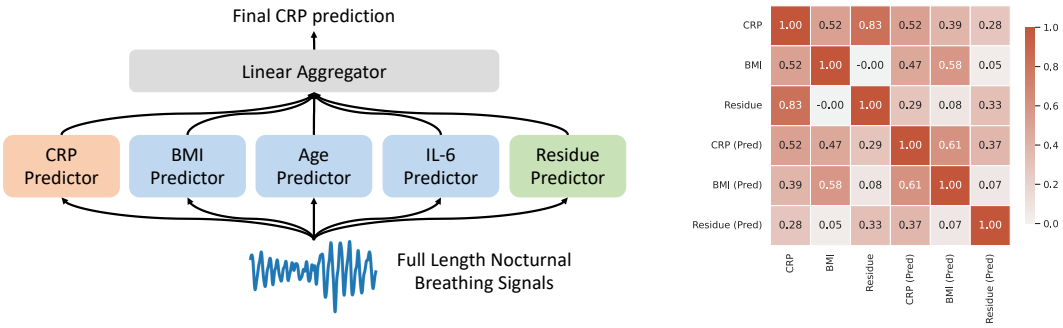
Monitoring inflammation's presence is crucial in both preventive medicine and disease management. However, conventional inflammatory biomarkers, levels of C-reactive protein (CRP), require invasive blood tests for assessment, which can only be conducted in hospitals. In this study, we develop machine learning models to detect inflammation and track its progression from nocturnal breathing signals. The performance was evaluated across multiple datasets comprising 1,174 nights of recordings from 950 individuals. Results demonstrate that our model offers promising accuracy in predicting CRP levels: achieving a Pearson's correlation of $r = 0.54$ using breathing signals alone, and $r = 0.63$ when incorporating additional data on Body Mass Index (BMI) and demographics. It is the first time that decent accuracy in predicting CRP levels from breathing signals has been achieved. Notably, the model exhibits proficiency in identifying high inflammatory states (CRP $>$ 10 mg/l) with an impressive area-under-the-curve value of 0.84. The results demonstrates the potential for non-invasive, longitudinal, in-home inflammation monitoring through breathing-based biomarkers.

1 INTRODUCTION

Inflammation is a complex biological response of the body's immune system to harmful stimuli. It serves as a frontline defense mechanism, aiming to eliminate the cause of cell injury, clear out necrotic cells and tissues, and initiate tissue repair. However, when dysregulated, inflammation can become chronic and contribute to the pathogenesis of various diseases, including cardiovascular diseases (Libby, 2006; Golia et al., 2014), autoimmune disorders (Abou-Raya & Abou-Raya, 2006), metabolic syndrome (Sutherland et al., 2004), depression (Miller & Raison, 2016) and cancer (Coussens & Werb, 2002). Consequently, understanding and monitoring inflammation levels have become paramount in both preventive medicine and disease management.

C-reactive protein (CRP) is well-known as a sensitive albeit non-specific biomarker of inflammation. It is synthesized by the liver in response to pro-inflammatory cytokines, particularly interleukin-6 (IL-6). Although, CRP has been widely used as a predictor of the occurrence and prognosis of various diseases (Dahl et al., 2007), measuring CRP can be burdensome, as it requires invasive procedures of getting blood samples and conducting laboratory tests. This limitation has motivated researchers to explore alternative approaches for monitoring inflammation. For example, Jiang et al. (2022) has shown that a deep learning model can detect high CRP level ($>$ 5 mg/l) with an AUC of 0.85 using electrocardiogram (ECG) signals.

In this work, we propose predicting CRP levels from nocturnal breathing, motivated by (1) the known association between breathing and inflammation: respiratory diseases correlate with increased inflammation (Vassilakopoulos et al., 2004), and controlled breathing reduces inflammation (Twal et al., 2016); and (2) the potential for effortless, in-home inflammation monitoring since breathing signals can be remotely monitored via wireless devices (Yue et al., 2018). Our model assembles multiple neural networks that are trained to predict CRP, other related variables (like BMI and age), and CRP's residue orthogonal to those variables. This design allows our model to independently learn distinct feature sets relevant to CRP, enhancing its predictive capability. With only nocturnal breathing signals, our model demonstrates a Pearson correlation of 0.54 with actual



(a) Breathing-Based CRP Model Structure (b) Correlation Matrix

Figure 1: (a) Structure of our breathing-based CRP model which is an ensemble of multiple neural network predictors trained with different labels. (b) Correlations between the ground truth CRP, BMI, CRP residue and the outputs of CRP/BMI/residue predictors.

CRP levels, which increases to 0.63 when BMI and demographic data are included. This accuracy enables the model to detect acute inflammation (defined as CRP $>$ 10 mg/l) with an AUC of 0.84. Moreover, our model effectively monitors CRP progression, accurately forecasting the variations in CRP levels as patients experience or recover from acute inflammation, evidenced by a correlation of 0.55. In summary, our contributions are:

- We present the first model that accurately predicts CRP levels from breathing signals.
- We propose a novel ensemble strategy to leverage auxiliary variables that have high correlations with the target.

2 Predicting CRP from BMI and Demographics

In this work, our objective is to predict CRP on a **logarithm** scale. As CRP distribution is highly skewed, it is common to apply log transform to facilitate analysis (Choi EunYoung et al., 2006). Throughout the remainder of the paper, when we refer to CRP in the context of predicting, it pertains to log-transformed CRP values.

We use linear regression model to predict CRP from Body Mass Index (BMI) and three relevant demographics: age, sex, race. These variables establish a baseline CRP level for healthy subjects and illuminate chronic inflammation states. The predictive capabilities of these factors on CRP and their combined predictive power are reported in Table 4. In the following, we summarize the relationship between CRP and these factors.

BMI. BMI exhibits a strong correlation with CRP levels, as proinflammatory cytokines from adipose tissue stimulate the secretion of CRP in the liver. Popko et al. (2010) reports a correlation of 0.4 between BMI values and CRP levels.

Age. CRP levels increase with age. In the USA, the median CRP level is 1.4 mg/l among those aged 20 to 29 years and 2.7 mg/l among those aged 80 years or older (Woloshin & Schwartz, 2005).

Sex & Race. CRP levels vary across sex and racial groups. Firstly, CRP levels are higher among women than among men (median, 2.7 vs. 1.6 mg/l), as reported by Woloshin & Schwartz (2005). According to Khera et al. (2005), black subjects have higher CRP levels than white subjects (median, 3.0 vs. 2.3 mg/l). On the other hand, the level of CRP in Asian populations is approximately one-third of the median value in Caucasians (Saito et al., 2014).

3 Predicting CRP from Nocturnal Breathing

Design of ensemble. Figure 1(a) illustrates the design of our breathing-based model, an ensemble of multiple neural network (NN) predictors, consisting of three parts: (1) a vanilla CRP predictor learned with CRP values as labels, (2) relevant variable predictors that learn to predict other variables having high correlations with CRP. Here, we have three predictors for BMI, age, and IL-6, respectively, and (3) a CRP residue predictor. The CRP residue is defined as its component orthogonal to the relevant variables. Mathematically, we denote CRP as $y \in \mathbb{R}^n$ and BMI,

age, and IL-6 as $v_1, v_2, v_3 \in \mathbb{R}^n$ (where n is the number of data samples). The CRP residue, $y_\perp := y - V(V^\top V)^{-1}V^\top y$ (where $V = [v_1, v_2, v_3]$), is CRP subtracted by its projection onto the space of relevant variables.

Motivation of ensemble. To facilitate the discussion, we focus on CRP and BMI. Firstly, the inherent correlation between CRP and BMI implies that a model that learns to predict one naturally serve as a weak predictor for the other. This is substantiated by the results shown in Figure 1(b), where the BMI predictor's output exhibits a correlation of 0.39 with CRP. Therefore, incorporating the BMI predictor into our model ensemble enhances our ability to predict CRP. Secondly, the significant correlation between CRP and BMI may cause the CRP predictor to disproportionately focus on BMI, neglecting other pertinent information. Evidence of this is seen in Figure 1(b), where the CRP predictor's output correlates more strongly with BMI (0.47) than with CRP residue (0.29). In contrast, the actual CRP levels show a lower correlation with BMI (0.52) compared to CRP residuals (0.83). This discrepancy highlights CRP predictor's deficiency in learning CRP residue. Thus, to mitigate this issue, we add a dedicated CRP residue predictor into our model ensemble.

Finally, the multiple predictors are assembled via linear regression. Table 5 in the appendix summarizes the performance of using each single predictor and the final ensemble. The ensemble strategy is demonstrated to be effective, increasing the R^2 by 7% to 20% compared to the vanilla CRP predictor. Please refer to appendix D for the details of NN architecture and training procedure.

4 Experiment

4.1 Datasets

We use two public medical datasets from NSRR (Zhang et al., 2018), Cleveland Family Study (*CFS*) (Redline et al., 1995) and Heart Biomarker Evaluation in Apnea Treatment (*HeartBEAT*) (Gottlieb et al., 2014). For each subject, the dataset contain basic demographics, anthropometric measurement, sleep studies which include the respiration signals measured by a breathing belt, and clinical laboratory tests which include levels of CRP and IL-6. Please refer to appendix C for more details such as dataset statistics and inclusion-exclusion criteria.

4.2 Models

We evaluate three models: **(A)** *The BMI and demographic-based model (BadCRP)*, which is a linear model taking BMI, age, sex, and race as input. **(B)** *The respiration-derived CRP model (RedCRP)*, an ensemble of multiple predictors of CRP, IL-6, BMI, age, and CRP residue. These predictors are aggregated via linear combination. **(C)** *The compositional model (RedCRP+), which takes both variables and breathing as inputs*. It assembles BMI, age, sex, race, and CRP predictor, IL-6 predictor, and CRP residue predictor. Note that RedCRP+ does not include predictors for BMI and age, as the ground truth BMI and age are already provided as input.

4.3 Tasks & Evaluating Metrics

CRP regression. We employ three evaluation metrics to compare the predicted CRP against the ground truth: R squared (R^2), Pearson correlation (r), and mean absolute error (MAE).

High CRP detection. A CRP level greater than 10 mg/l indicates acute inflammation typically caused by infections (Black et al., 2004). Thus, we assess our model's ability to detect acute inflammation, by the following metrics: area-under-the-curve (AUC), sensitivity (SE), specificity (SP).

CRP tracking. The HeartBEAT dataset comprises two visits with a two to three-week interval between them. We evaluate our model's ability to capture the change in CRP levels between the two visits. The evaluation is based on the Pearson correlation between the ground truth CRP delta and the predicted CRP delta.

4.4 Results

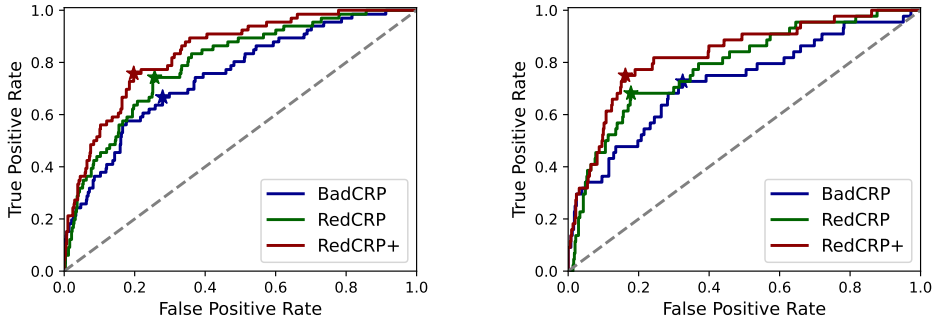
Table 1 summarizes the performance of the model in predicting CRP levels. The key insights include: (1) Models that rely solely on breathing signals either match or surpass the performance

Table 1: Performance on CRP regression.

Model	CFS			HeartBEAT		
	R^2	r	MAE	R^2	r	MAE
BadCRP	0.29	0.54	0.83	0.13	0.36	0.83
RedCRP	0.28	0.53	0.83	0.19	0.44	0.82
RedCRP+	**0.40**	**0.63**	**0.76**	**0.26**	**0.51**	**0.78**

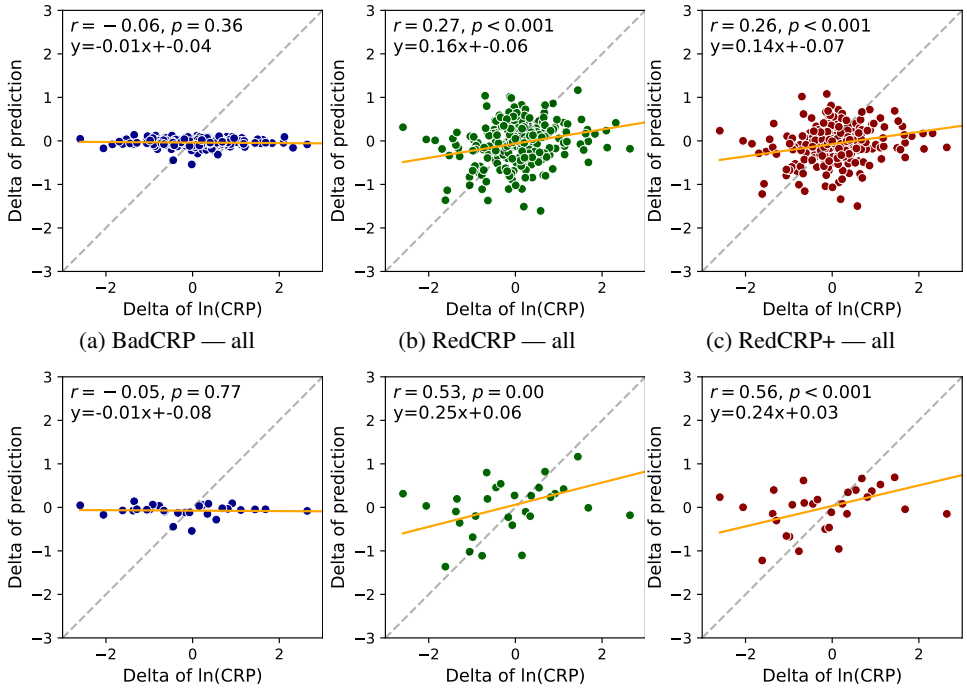
Table 2: Performance on high CRP detection.

Model	CFS			HeartBEAT		
	AUC	SE	SP	AUC	SE	SP
BadCRP	0.75	0.67	0.72	0.73	0.73	0.68
RedCRP	0.80	0.74	0.74	0.79	0.68	0.82
RedCRP+	**0.84**	**0.76**	**0.80**	**0.83**	**0.75**	**0.84**



(a) CFS (b) HeartBEAT

Figure 2: ROC curves of detecting high CRP levels ($>$ 10 ug/ml). Asterisks indicates best trade-offs between sensitivity and specificity.



(a) BadCRP — all (b) RedCRP — all (c) RedCRP+ — all

(d) BadCRP — inflamed (e) RedCRP — inflamed (f) RedCRP+ — inflamed

Figure 3: Performance in capturing the changes of CRP levels between two visits. (a-c) are results for *all* subjects. (d-f) are results for *inflamed* subjects who had a CRP level $>$ 10 ug/ml in at least one visit.

of models that use BMI and demographics. Specifically, RedCRP achieves comparable results to BadCRP in the CFS dataset and demonstrates superior R^2 and Pearson correlation values in the HeartBEAT dataset. (2) The inclusion of breathing signals considerably enhances the model's ability to predict CRP levels beyond what is possible with just BMI and demographic data. This is highlighted by RedCRP+'s significant improvement over BadCRP, with an increase of 0.11 in the R^2 value for CFS and 0.13 for HeartBEAT. It suggest that breathing signals account for an additional 11%-13% of the variance in CRP levels compared to BMI and demographics, indicating the substantial predictive power of respiratory patterns in assessing inflammation.

Table 2 summarizes the model performance in detecting acute inflammation, with the corresponding ROC curves displayed in Figure 2. The results reveal that: (1) RedCRP outperforms BadCRP, indicating that breathing signals are more predictive of acute inflammation than BMI and demographics. This outcome aligns with expectations, as BMI, age, and sex do not undergo significant changes in a short timeframe, and thus, are more indicative of chronic rather than acute inflammation. In contrast, acute inflammation has an immediate impact on breathing patterns, with severe inflammation linked to symptoms of acute respiratory distress syndrome, such as rapid breathing (tachypnea) and difficulty breathing (dyspnea) (Bhatia & Moochhala, 2004). (2) RedCRP+ outperforms both BadCRP and RedCRP with a sensitivity above 0.75 and specificity above 0.80 across both datasets. This indicates that by integrating breathing signals with BMI and demographics, we can accurately detect elevated CRP levels.

Figure 3 presents the efficacy of the model in monitoring fluctuations in subjects' CRP levels. The findings reveal that (1) BMI and demographic factors lack the capacity to detect changes in inflammation levels since these parameters either not change over a short period (BMI, age) or remain constant (sex, race). Consequently, BadCRP demonstrates negligible correlations. (2) Conversely, RedCRP and RedCRP+ exhibit a substantial ability to identify variations in CRP levels, as evidenced in Figure 3(b,c,e,f), which shows positive correlations between 0.26 and 0.56 and p-values below 0.001. (3) When compared to patients lacking acute inflammation, breathing based model achieves better performance in individuals with acute inflammation. Specifically, Figure 3(b,c,e,f) indicates that in patients with severe inflammation, RedCRP+ achieves correlations of 0.56 and slopes of 0.24, markedly superior to the correlations of 0.26 and slopes of 0.14 observed across all patients. This enhanced performance is crucial as it aids in the more effective monitoring of CRP progression in patients with acute inflammation, potentially influencing their treatment decisions.

5 Discussion

Impact. (1) Our work pioneers in accurately predicting CRP levels from breathing signals, moving beyond merely detecting CRP above a threshold as previous studies have (Jiang et al., 2022). We achieve a notable correlation of 0.5-0.6 on a logarithmic scale, at the same time, highlighting the capability of breathing signals to precisely reflect acute inflammation, offering new insights into the relationship between the respiratory system and inflammation. (2) Leveraging advancements in remote breathing monitoring via low-power RF signal analysis (Yue et al., 2018), our model facilitates at-home, continuous inflammation tracking. This enables effective disease progression monitoring, medication impact assessment, and early infection detection. (3) Our technique of using an ensemble to leverage correlated variables to improve the prediction of the target variable is generally applicable to other applications.

Limitations. (1) Our work shows that breathing signals (partially) predict CRP, but does not explain the underlying mechanism. It remains unknown whether it is related to breathing rate, depth of breathing, or changes in breathing patterns throughout the night. Interpretable analysis would be desired and considered as our future work. (2) Our work focuses on predicting CRP while not considering other inflammatory biomarkers such as erythrocyte sedimentation rate (which measures how quickly red blood cells settle at the bottom of a tube of blood), procalcitonin, and cytokines that regulate immune responses and inflammation (such as interleukins and tumor necrosis factor-alpha).

Acknowledgments

This work is funded by the National Science Foundation Award No. IIS-2014391.

REFERENCES

Anna Abou-Raya and Suzan Abou-Raya. Inflammation: a pivotal link between autoimmune diseases and atherosclerosis. *Autoimmunity reviews*, 5(5):331–337, 2006.

Madhav Bhatia and Shabbir Moochhala. Role of inflammatory mediators in the pathophysiology of acute respiratory distress syndrome. *The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland*, 202(2):145–156, 2004.

Steven Black, Irving Kushner, and David Samols. C-reactive protein. *Journal of Biological Chemistry*, 279(47):48487–48490, 2004.

Andreas Brink-Kjaer, Eileen B Leary, Haoqi Sun, M Brandon Westover, Katie L Stone, Paul E Peppard, Nancy E Lane, Peggy M Cawthon, Susan Redline, Poul Jennum, et al. Age estimation from sleep studies using deep learning predicts life expectancy. *NPJ digital medicine*, 5(1):103, 2022.

Youngjun Cho, Nadia Bianchi-Berthouze, and Simon J Julier. Deepbreath: Deep learning of breathing patterns for automatic stress recognition using low-cost thermal imaging in unconstrained settings. In *2017 Seventh international conference on affective computing and intelligent interaction (acii)*, pp. 456–463. IEEE, 2017.

Choi EunYoung Choi EunYoung, Park EalHwan Park EalHwan, Cheong YooSeock Cheong YooSeock, Rheem InSoo Rheem InSoo, Park SeoukGun Park SeoukGun, and Yoo SunMi Yoo SunMi. Association of c-reactive protein with the metabolic risk factors among young and middle-aged koreans. 2006.

Lisa M Coussens and Zena Werb. Inflammation and cancer. *Nature*, 420(6917):860–867, 2002.

Morten Dahl, Jørgen Vestbo, Peter Lange, Stig E Bojesen, Anne Tybjærg-Hansen, and Børge G Nordestgaard. C-reactive protein as a predictor of prognosis in chronic obstructive pulmonary disease. *American journal of respiratory and critical care medicine*, 175(3):250–255, 2007.

A Magon de la Villehuchet, Michel Brack, Gerard Dreyfus, Yacine Oussar, Dominique Bonnefont-Rousselot, MJ Chapman, and A Kontush. A machine-learning approach to the prediction of oxidative stress in chronic inflammatory disease. *Redox Report*, 14(1):23–33, 2009.

Enrica Golia, Giuseppe Limongelli, Francesco Natale, Fabio Fimiani, Valeria Maddaloni, Ivana Pariggiano, Renatomaria Bianchi, Mario Crisci, Ludovica D'Acierno, Roberto Giordano, et al. Inflammation and cardiovascular disease: from pathogenesis to therapeutic target. *Current atherosclerosis reports*, 16:1–7, 2014.

Daniel J Gottlieb, Naresh M Punjabi, Reena Mehra, Sanjay R Patel, Stuart F Quan, Denise C Babineau, Russell P Tracy, Michael Rueschman, Roger S Blumenthal, Eldrin F Lewis, et al. Cpap versus oxygen in obstructive sleep apnea. *New England Journal of Medicine*, 370(24): 2276–2285, 2014.

Hao He, Yuan Yuan, Ying-Cong Chen, Peng Cao, and Dina Katabi. Contactless oxygen monitoring with radio waves and gated transformer. In *Machine Learning for Healthcare Conference*, pp. 248–265. PMLR, 2023.

Junrong Jiang, Hai Deng, Hongtao Liao, Xianhong Fang, Xianzhang Zhan, Shulin Wu, and Yumei Xue. Development and validation of a deep-learning model to detect crp level from the electrocardiogram. *Frontiers in Physiology*, 13:864747, 2022.

Amit Khera, Darren K McGuire, Sabina A Murphy, Harold G Stanek, Sandeep R Das, Wanpen Vongpatanasin, Frank H Wians, Scott M Grundy, and James A de Lemos. Race and gender differences in c-reactive protein levels. *Journal of the American college of cardiology*, 46(3): 464–469, 2005.

Jiseon Kim and Jooyong Kim. Classification of breathing signals according to human motions by combining 1d convolutional neural network and embroidered textile sensor. *Sensors*, 23(12): 5736, 2023.

Peter Libby. Inflammation and cardiovascular disease mechanisms. *The American journal of clinical nutrition*, 83(2):456S–460S, 2006.

Andrew H Miller and Charles L Raison. The role of inflammation in depression: from evolutionary imperative to modern treatment target. *Nature reviews immunology*, 16(1):22–34, 2016.

Nghia H Nguyen, Dominic Picetti, Parambir S Dulai, Vipul Jairath, William J Sandborn, Lucila Ohno-Machado, Peter L Chen, and Siddharth Singh. Machine learning-based prediction models for diagnosis and prognosis in inflammatory bowel diseases: a systematic review. *Journal of Crohn's and Colitis*, 16(3):398–413, 2022.

Katarzyna Popko, E Gorska, A Stelmaszczyk-Emmel, R Plywaczewski, A Stoklosa, D Gorecka, B Pyrzak, and U Demkow. Proinflammatory cytokines il-6 and tnf-α and the development of inflammation in obese subjects. *European journal of medical research*, 15:1–3, 2010.

Susan Redline, Peter V Tishler, Tor D Tosteson, John Williamson, Kenneth Kump, Ilene Browner, Veronica Ferrette, and Patrick Krejci. The familial aggregation of obstructive sleep apnea. *American journal of respiratory and critical care medicine*, 151(3):682–687, 1995.

Isao Saito, Koutatsu Maruyama, and Eri Eguchi. C-reactive protein and cardiovascular disease in east asians: a systematic review. *Clinical Medicine Insights: Cardiology*, 8:CMC–S17066, 2014.

Julie P Sutherland, Benjamin McKinley, and Robert H Eckel. The metabolic syndrome and inflammation. *Metabolic syndrome and related disorders*, 2(2):82–104, 2004.

Waleed O Twal, Amy E Wahlquist, and Sundaravadivel Balasubramanian. Yogic breathing when compared to attention control reduces the levels of pro-inflammatory biomarkers in saliva: a pilot randomized controlled trial. *BMC complementary and alternative medicine*, 16(1):1–10, 2016.

T Vassilakopoulos, C Roussos, and S Zakynthinos. The immune response to resistive breathing. *European Respiratory Journal*, 24(6):1033–1043, 2004.

Steven Woloshin and Lisa M Schwartz. Distribution of c-reactive protein values in the united states. *New England Journal of Medicine*, 352(15):1611–1613, 2005.

Yuzhe Yang, Yuan Yuan, Guo Zhang, Hao Wang, Ying-Cong Chen, Yingcheng Liu, Christopher G Tarolli, Daniel Crepeau, Jan Bukartyk, Mithri R Junna, et al. Artificial intelligence-enabled detection and assessment of parkinson's disease using nocturnal breathing signals. *Nature medicine*, 28(10):2207–2215, 2022.

Shichao Yue, Hao He, Hao Wang, Hariharan Rahul, and Dina Katabi. Extracting multi-person respiration from entangled rf signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(2):1–22, 2018.

Guo-Qiang Zhang, Licong Cui, Remo Mueller, Shiqiang Tao, Matthew Kim, Michael Rueschman, Sara Mariani, Daniel Mobley, and Susan Redline. The national sleep research resource: towards a sleep data commons. *Journal of the American Medical Informatics Association*, 25(10):1351–1358, 2018.

Qiang Zhang, Xianxiang Chen, Qingyuan Zhan, Ting Yang, and Shanhong Xia. Respiration-based emotion recognition with deep learning. *Computers in Industry*, 92:84–90, 2017.

Mingmin Zhao, Shichao Yue, Dina Katabi, Tommi S Jaakkola, and Matt T Bianchi. Learning sleep stages from radio signals: A conditional adversarial architecture. In *International Conference on Machine Learning*, pp. 4100–4109. PMLR, 2017.

A Extra Tables

Table 3: Statistics of datasets.

Dataset	CFS	HeartBEAT
No. of individuals	639	311
No. of nights	639	555
Age	[6.9, 88.5], 43.0 (18.9)	[45.1, 75.8], 63.2 (7.2)
Sex (Male)	44.6%	73.7%
Race (White, Black)	42.1% , 54.8%	81.3% , 12.6%
BMI	[14.0, 84.8], 33.3 (9.2)	[21.2, 64.4], 34.2 (6.2)
AHI	[0.0, 125.1], 14.0 (20.7)	[0.3, 75.7], 22.1 (11.9)
ln(CRP)	[-1.8, 4.8], 0.8 (1.2)	[-1.9, 4.5], 0.7 (1.1)

*For age, bmi, ahi, ln(CRP), each cell reports [min, max], mean (std).

Table 4: Performance of predicting CRP using BMI and demographics by a linear model.

Feature				CFS			HeartBEAT		
BMI	Age	Sex	Race	R^2	r	MAE	R^2	r	MAE
✓				0.265	0.514	0.847	0.103	0.321	0.848
	✓			0.057	0.240	0.990	0.013	0.116	1.106
		✓		0.010	0.100	1.017	0.061	0.248	0.873
			✓	0.014	0.122	1.011	0.023	0.159	0.896
✓	✓	✓	✓	0.295	0.543	0.828	0.168	0.361	0.829

Table 5: Performance of breathing-based predictors and their ensemble on predicting CRP.

Predictors in ensemble				CFS			HeartBEAT		
CRP	BMI	IL-6	CRP residue	R^2	r	MAE	R^2	r	MAE
✓				0.263	0.513	0.843	0.164	0.406	0.819
	✓			0.151	0.389	0.932	0.060	0.246	0.876
		✓		0.161	0.401	0.913	0.093	0.306	0.865
			✓	0.074	0.272	0.974	0.124	0.352	0.856
✓	✓			0.272	0.522	0.841	0.166	0.408	0.817
✓		✓		0.264	0.514	0.843	0.165	0.407	0.819
✓			✓	0.271	0.520	0.838	0.189	0.435	0.817
✓	✓	✓	✓	0.281	0.530	0.833	0.195	0.442	0.811

B Related Work

Machine learning on predicting inflammation. Research on predicting inflammation using machine learning is relatively sparse. Some studies have concentrated on diseases related to inflammation. For instance, de la Villehuchet et al. (2009) developed a model to predict oxidative stress in chronic inflammatory diseases, while Nguyen et al. (2022) employed machine learning techniques to forecast the occurrence of inflammatory bowel diseases. To our knowledge, only one existing study directly addresses the prediction of inflammation, specifically detecting C-reactive protein (CRP) levels greater than 5 mg/L using ECG signals (Jiang et al., 2022). Our work distinguishes itself from this prior study in two significant ways. Firstly, we focus on regressing the CRP levels rather than performing binary classification. Secondly, our approach utilizes breathing signals as opposed to ECG signals.

Machine learning on breathing signals. Breathing signals are a critical aspect of health monitoring, offering rich and insightful information about an individual's well-being. Previous studies have shown that using AI models to analyze breathing patterns can predict daily activities, such as human motions (Kim & Kim, 2023), assess sleep quality, including sleep stages (Zhao et al., 2017), and estimate demographics like age (Brink-Kjaer et al., 2022). Additionally, they can monitor other physiological signals, such as blood oxygen levels (He et al., 2023), detect emotions (Zhang et al., 2017), and identify stress (Cho et al., 2017), as well as diseases, for example, Parkinson's disease (Yang et al., 2022). In our work, we demonstrate for the first time that it is possible to predict inflammation from breathing signals.

C DATASETS

Overview. In Table 3, we summarize the statistics of the datasets. The CFS dataset was primarily collected to study the familial aggregation of obstructive sleep apnea. Thus, the dataset spans a wide range of ages, from young children below 10 years old to seniors above 80 years old. As the data were collected from families, it has a balanced sex distribution (44.6% vs. 55.4%). The HeartBEAT dataset was primarily collected to evaluate the effects of supplemental nocturnal oxygen or Positive Airway Pressure (PAP) therapy in patients with cardiovascular disease (CVD). The dataset mainly comprises middle-aged males with CVD and moderate to severe obstructive sleep apnea (AHI from 15 to 50).

Study inclusion and exclusion criteria. For the CFS dataset, we included participants with polysomnography (PSG) data from sleep studies and valid CRP readings from clinical laboratory tests. Based on the PSG data, we excluded participants with recordings of less than 2 hours during sleep. In total, we included 639 participants in our study. For the HeartBEAT dataset, we included participants with polysomnography (PSG) data from sleep studies and valid CRP readings from clinical laboratory tests. Based on the breathing signals from PSG, we excluded participants with low-quality breathing signals, defined as having both thorax and abdominal breathing channels unusable for more than 5% of the total sleep time. In total, we included 311 participants with 555 nights of data in our study.

D NEURAL NETWORK

D.1 MODEL ARCHITECTURE

We utilize the same deep neural network architecture to predict CRP and other variables from nocturnal breathing. Our model is divided into three parts: (1) at the bottom layers, convolutional layers with residue links extract features from local breathing patterns; (2) at the bottleneck, recurrent layers and an attention module aggregate local features and process them in the context of the global sequence; (3) at the top layers, a multi-layer perceptron (MLP) head predicts the target variable using the max-pooled feature from the bottleneck.

Specifically, in the convolutional layers, we have a total of nine 1-D convolutional residue blocks, which together subsample the input signal by a factor of 300 and encode it into a series of 256-dimensional latent representations. Our breathing signal input has a sampling frequency of 10Hz. Thus, in the latent space, we have a representation every 30 seconds. Then, at the bottleneck, we have a bidirectional LSTM with a hidden size of 256 followed by a transformer layer with 8 heads and a hidden size of 256. At the top, we have an MLP with 2 hidden layers with a size of 512 to make the final prediction.

D.2 TRAINING PROCEDURE

Train/test splits. Given the population differences between the CFS and HeartBEAT datasets, we train separate models for each dataset. Each dataset is divided into four folds. For each model, we run four instances trained on three folds and tested on the remaining fold. Subsequently, we combine the predictions from the test sets across all four folds, representing the entire dataset, and compare them with the ground truth CRP levels to calculate evaluation metrics.

Hyper-parameters. All the deep models are trained with a learning rate of 1e-4 and weight decay of 1e-6 using the Adam optimizer for 50 epochs. The hyperparameters were determined through a primary grid search. We searched for the number of epochs from 20, 50, 100, 200, learning rates from 1e-5, 3e-5, 1e-4, 3e-4, and weight decay from 1e-5, 1e-6, 1e-7.

Loss function. Given that all models are aimed at predicting a continuous variable, we employ the mean squared error (MSE) as the loss function.