

Article

Disclosing Critical Voice Features for Discriminating between Depression and Insomnia—A Preliminary Study for Developing a Quantitative Method

Ray F. Lin ^{1,*}, Ting-Kai Leung ^{2,3,†}, Yung-Ping Liu ⁴ and Kai-Rong Hu ¹

¹ Department of Industrial Engineering and Management, Yuan Ze University, Taoyuan 32003, Taiwan; karenhu0205@gmail.com

² Department of Radiology, Taoyuan General Hospital, Ministry of Health and Welfare, No. 1492, Zhongshan Rd., Taoyuan City 33004, Taiwan; hk8648@tmu.edu.tw

³ Graduate Institute of Biomedical Materials and Tissue Engineering, College of Biomedical Engineering, Taipei Medical University, Taipei 11031, Taiwan

⁴ Department of Industrial Engineering and Management, Chaoyang University of Technology, Taichung 413310, Taiwan; ypliu@cyut.edu.tw

* Correspondence: juifeng@saturn.yzu.edu.tw

† These authors contributed equally to this work.

Abstract: **Background:** Depression and insomnia are highly related—insomnia is a common symptom among depression patients, and insomnia can result in depression. Although depression patients and insomnia patients should be treated with different approaches, the lack of practical biological markers makes it difficult to discriminate between depression and insomnia effectively. **Purpose:** This study aimed to disclose critical vocal features for discriminating between depression and insomnia. **Methods:** Four groups of patients, comprising six severe-depression patients, four moderate-depression patients, ten insomnia patients, and four patients with chronic pain disorder (CPD) participated in this preliminary study, which aimed to record their speaking voices. An open-source software, openSMILE, was applied to extract 384 voice features. Analysis of variance was used to analyze the effects of the four patient statuses on these voice features. **Results:** statistical analyses showed significant relationships between patient status and voice features. Patients with severe depression, moderate depression, insomnia, and CPD reacted differently to certain voice features. Critical voice features were reported based on these statistical relationships. **Conclusions:** This preliminary study shows the potential in developing discriminating models of depression and insomnia using voice features. Future studies should recruit an adequate number of patients to confirm these voice features and increase the number of data for developing a quantitative method.

Keywords: depression; insomnia; openSMILE; voice feature; computer-aided diagnosis



Citation: Lin, R.F.; Leung, T.-K.; Liu, Y.-P.; Hu, K.-R. Disclosing Critical Voice Features for Discriminating between Depression and Insomnia—A Preliminary Study for Developing a Quantitative Method. *Healthcare* **2022**, *10*, 935. <https://doi.org/10.3390/healthcare10050935>

Academic Editor: Daniele Giansanti

Received: 11 April 2022

Accepted: 16 May 2022

Published: 18 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Depression vs. Insomnia

Depression and insomnia are both prevalent disorders that are highly related. Lim, et al. [1] evaluated the prevalence of depression in different countries between 1994 and 2014 and found that the aggregated point, one-year, and lifetime prevalence of depression were 12.9, 7.2 and 10.8%, respectively. On the other hand, many studies reported that chronic insomnia affects, approximately, from 5 to 30% of the population [2–7]. Research has shown a strong relationship between insomnia and depression [8–12]. However, the exact nature of the relationship is complex [9]. Insomnia, defined as sleep difficulty, can occur independently or due to other problems, including psychological factors (e.g., stress), physical factors (e.g., chronic pain), and other possibilities. Many studies considered insomnia a symptom of depression, but recent studies have found insomnia is also a risk factor for the development of depression [12–14]. Depressed patients with insomnia as

a symptom could be treated with antidepressant medication or cognitive therapy [15]. However, insomnia patients should be treated differently from depression patients. Recent studies show that cognitive behavioral therapy for insomnia is highly effective compared to other psychological interventions [16,17]. Nonetheless, applying depression treatments to insomnia patients may cause a relapse of depression [13]. Hence, it is critical to have accurate diagnoses so that these two types of patients can be treated well.

1.2. Diagnoses of Depression and Insomnia

It has been pointed out that both diagnoses of depression and insomnia are difficult tasks. To diagnose and assess depression, one should rely on subjective behavior measures, such as self-reports or family reports and clinical interviews [18,19]. Commonly used assessment tools are the Hamilton Rating Scale for Depression [20] and the Beck Depression Index [21]. An expert practitioner subjectively performs these rating scales based on a patient's mental state, driven by interviews and self-report experiences [22]. However, the assessment is costly, time-consuming, and often requires patients' presence at the clinic.

On the other hand, similar difficulties were reported while diagnosing insomnia. Schramm, et al. [23] pointed out practical issues in structured clinical interviews, which are ineffective and time-consuming. Although tools (e.g., the Insomnia Severity Index, Athens Insomnia Scale, and Pittsburg Sleep Quality Index) have been developed for screening insomnia, evidence shows that insomnia is still under-recognized, underdiagnosed, and undertreated [24,25]. Consequently, it is necessary to develop a more systematic way of measuring and quantifying depression and insomnia within and across clinical sessions for a more effective and efficient diagnosis.

1.3. Speech Voice May Help Diagnoses

Human speech voice correlates with the emotional state, so using voice may aid in developing a quantitative method for distinguishing depression patients from insomnia patients [19]. It is well-known that gender influences speech differences. For example, compared to males, females have a relatively higher pitch [26,27], greater voice quality [27], and a larger pitch range [28]. Except for the changes due to gender, slight physiological and cognitive changes can also produce acoustic changes [29–31]. Research on emotional speech has shown that emotion affects prosodic and spectral speech characteristics [32]. The speech production and acoustic quality of the speech are affected by depression caused by cognitive and physiological changes [31].

Although studies have shown that depression leads to speech differences and using speech to distinguish depression is feasible [31–36], there is little evidence showing that insomnia causes differences in voice. In contrast, Heydarifard and Krasikova [37] found no association between insomnia and next-day prohibitive voice. While depression caused by cognitive and physiological changes affects speech and insomnia has no effects on speech, it is reasonable to hypothesize that voice features exist to discriminate between depression and insomnia. However, to the authors' best knowledge, few studies have directly reported speech differences between depression and insomnia.

1.4. Research Objective

While diagnosing depression and insomnia is difficult, and research gaps exist in determining their effects on voice, this preliminary study aimed to test the effects of depression and insomnia on voice features. With a limited number of participants, we expected that certain voice features would show the differences between depression and insomnia. If so, this preliminary study could confirm the possibility of executing the following research with an adequate number of patients for data collection and, hence, develop a quantitative method to discriminate between depression and insomnia.

2. Materials and Methods

2.1. Participants and Clinical Assessment

Twenty-four patients who visited the Department of Radiology, the Taoyuan General Hospital, Ministry of Health and Welfare, from July 2018 to September 2018, volunteered to participate in this study. They belonged to different outpatient departments and suffered from psychological diseases, sleep difficulty, or chronic pain disorders. Outpatient physicians recruited these participants by inviting every visiting patient who presented with the above symptoms during off-peak clinic hours. To categorize these participants, they were required to fill out the Hamilton Depression Rating Scale [38] questionnaire and other clinical assessments by our clinical research physician. As suggested by Zimmerman, et al. [39], score thresholds of 17 and 23 were first used to determine groups of moderate depression (17–23) and severe depression (23–65). The rest of the patients whose scores were less than 17 were defined as a non-depression group. The clinical research physician further categorized them according to their clinical features. As a result, there were six severe-depression patients (two females/four males), four moderate-depression patients (two females/two males), ten insomnia patients (six females/four males), and four patients (one females/three males) with chronic pain disorders (CPD; e.g., myalgia and knee pain).

2.2. Collection and Preprocessing of the Speech Voice

After the informed consent, the participants moved to a quiet room located in the hospital to record their speech voice data. The surrounding noise in the rooms was maintained under 45 dB. The participants followed the experimenter's instructions to read out the 21 questions of the Chinese version of the Beck Anxiety Inventory [40]. A microphone connected to a personal computer collected voice data while the participant was reading these questions.

The speech data were saved as a single channel, MP3 files, sampled at 44.1 kHz, 8-bit. The recorded MP3 files underwent several preprocesses using a voice processing software, Audacity™. First, the MP3 files were converted to WAV files. Second, irrelevant proportions of sound clips (e.g., cough, sneeze, chair moving, etc.) were cut out. Third, background noises were eliminated. Last, the files were cut into clips as the participants read a single question of the CBAI.

2.3. Voice Features Extraction and Statistical Analysis

After data preprocessing, an open-source software, openSMILE, IS09_eotion [41], was applied to extract 384 voice features. These features were calculated based on 16 descriptors, comprising root mean square (RMS), zero-crossing rate (ZCR), pitch frequency (F0), harmonics-to-noise ratio (HNR), and Mel-frequency cepstral coefficients (MFCC) 1–12. These original 16 descriptors (d) were then used to capture de-differentiated (partial differential) descriptors (d'), representing non-personalized features [42]. Next, the 12 statistic properties of the mean, standard deviation (SD), kurtosis, skewness, maximum and minimum values, maximum and minimum positions, and range, as well as two linear regression coefficients (offset and slope) with their mean square error (MSE), were computed for these 32 descriptors. The effects of participant status (i.e., severe-depression, moderate-depression, insomnia, and CPD) on these 384 voice features ($16 \times 2 \times 12$) were then assessed using analysis of variance (ANOVA). In addition to patient status, gender was also treated as an independent variable in analyses because gender is a critical factor that affects speech voice.

3. Results

ANOVA was performed on all the 384 voice features, using a model with status and gender as fixed effects and participant as a random effect nested within status and gender. As shown in Tables 1–3, there were many significant main effects and interaction effects of status and gender on voice features. While the focus is on patient status, all the main effects of status are depicted and detailed. For the main effects of gender and interaction effects

of gender and status, the distributions of statistical effects were presented to provide the recommendations for selecting voice features due to the limited space.

3.1. Effects of Patient Status

Regarding RMS, ZCR, F0, and HNR, status had significant effects (p -value < 0.05) on six d features but zero d' features. Status had significant effects on skewness and kurtosis of original descriptors of RMS and F0. As shown in Figure 1, the severe-depression group had the greatest RMS-skewness and RMS-kurtosis values compared to the other three groups. This phenomenon was even more apparent for F0-skewness and F0-kurtosis values. As shown in Figure 2, depression patients had significantly greater F0-skewness and F0-kurtosis values than patients with insomnia and CPD. Moreover, the severe-depression group had significantly greater values than the moderate-depression group. Furthermore, status had significant effects on slopes of F0 and HNR. As shown in Figure 3, the insomnia group had the greatest slope values of F0 and HNR compared to the other three groups.

Table 1. Effects of patient status on 384 voice features.

Effects	Original Descriptors (d)											De-Differentiated Descriptors (d')													
	Mean	SD	Skewness	Kurtosis	Max Value	Min Value	Max Position	Min Position	Range	Offset	Slope	MSE	Mean	SD	Skewness	Kurtosis	Max Value	Min Value	Max Position	Min Position	Range	Offset	Slope	MSE	
RMS			*	*																					
ZCR		^			^				^															^	
F0	^		●	*						*						^									
HNR	^									*															
MFCC	1								*																
	2			^								*	*								●	*	*	*	
	3											*			*						●	■			
	4								*												●				
	5			●								*			*										
	6												*					^	*		*		^		
	7						*			*			*												
	8								*																
	9		^				*		*					^			*				^	^		^	
	10	*		*		^		*	*	^		*	*		●		^	^		*	^		^	^	
	11									^				^	*			*						^	
	12			●					*				●								^		*		

^: $p < 0.1$; *: $p < 0.05$; ●: $p < 0.01$; ■: $p < 0.001$.

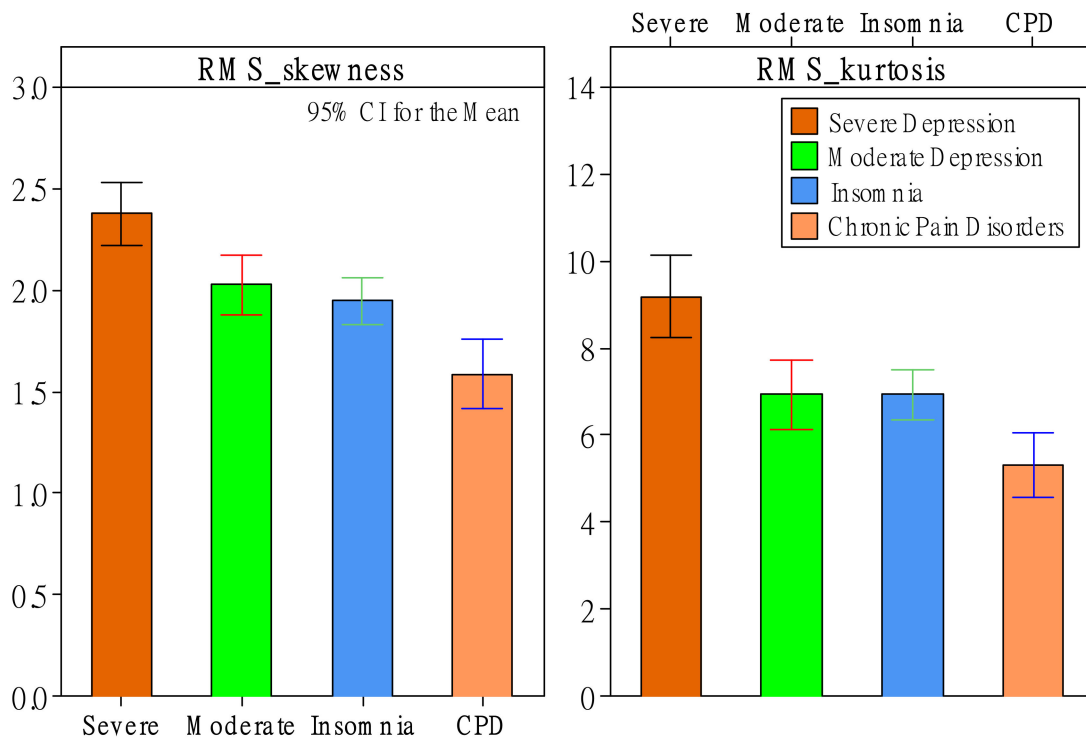


Figure 1. The effect of patient status on the root mean square (RMS).

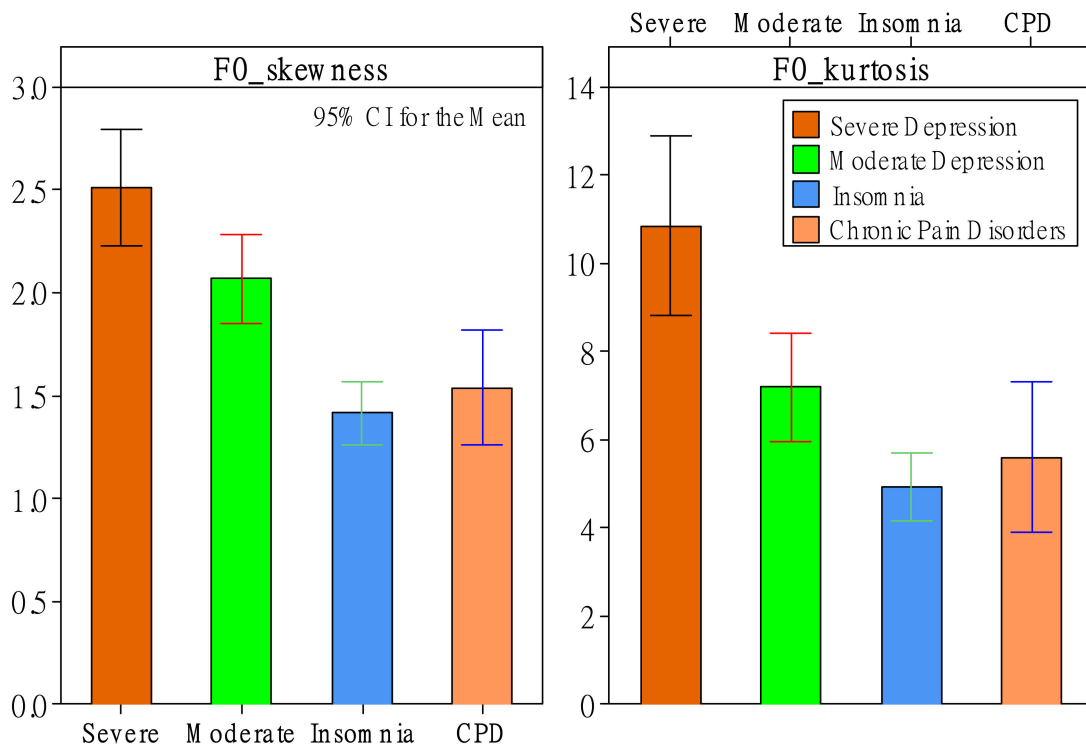


Figure 2. The effect of patient status on the pitch (F0).

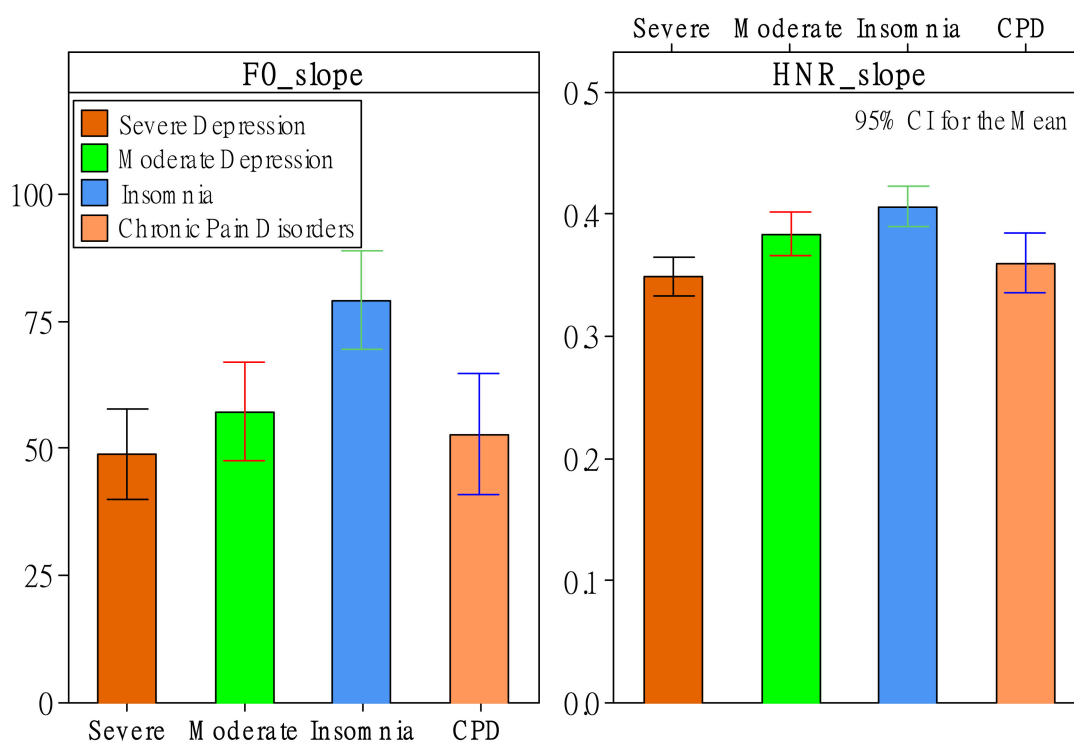


Figure 3. The effect of patient status on slopes of F0 and HNR.

3.2. Effects of Patient Status on MFCC

Regarding MFCC features, status had significant effects on 13 d features and 23 d' features. Kurtosis, again, was an essential feature in showing the differences among the patient groups. There were three significant status effects related to kurtosis found in d features (i.e., MFCCs 5, 10, and 12) and four in d' features (i.e., MFCCs 3', 5', 10', and 11'). As shown in Figure 4a, as far d features are concerned, there was a trend in that depression patients had greater kurtosis values, especially MFCC 5, compared to groups of insomnia and CPD. Where d' features are concerned, the moderate-depression group had the greatest kurtosis values among the three groups. Another critical category of features was the offset of the regression coefficient. There were four significant status effects related to offset found in d features (i.e., MFCCs 1, 4, 7, and 12) and three in d' features (i.e., MFCCs 2', 3', and 6'). As shown in Figure 4b, the CPD group had significant differences from the other three groups. The group had the greatest offset values of MFCC 1 and MFCC 2', and had the lowest offset values of MFCCs 4, 7, 12, 3', and 6'. Worth noting is that mean and slope features showed the patient differences only in d' features, not d features. There were four significant relationships of mean features (MFCCs 3', 5', 7', and 12), and three significant relationships of slope features (MFCCs 2', 3', and 12'). Regarding the mean features, as shown in Figure 4c, again, the CPD group had significant differences from the other three groups. It had the greatest mean value of MFCC 3' and had the lowest mean values of MFCCs 5', 7', and 12' in d' features. Regarding the slope features, as shown in Figure 4d, there was a trend in that depression patients had greater slope values of MFCCs 2' and 12' compared to the groups of insomnia and CPD. Again, the CPD group had significant differences from the other three groups. It had the greatest slope value of MFCC 3' and the lowest slope values of MFCCs 2' and 12'.

Other voice features that showed relatively less significant relationships (between one and two for either d or d' features) were SD, skewness, maximum value, maximum position, minimum position, range, and MSE. Regarding the SD features, as shown in Figure 4e, the severe-depression group had the lowest SD value of MFCC 10 and had the lowest SD values of MFCCs 2' and 10' in d' features. The SD value of MFCC 2' showed an ideal trend in that SD values were significantly different among the four groups. The values

increased from severe to moderate, insomnia, and then CPD. Regarding the skewness features, as shown in Figure 4f, the severe-depression group had the greatest skewness value of MFCC 2'. Regarding the maximum-value features, as shown in Figure 5a, the severe-depression group had the lowest maximum value of MFCC 9'. Regarding the maximum-position features, as shown in Figure 5b, there was a trend in that depression patients had greater maximum-position values, especially MFCC 7, compared to groups of insomnia and CPD. The moderate-depression group had the greatest maximum-position value of MFCC 9. The severe-depression group had the greatest maximum-position value of MFCC 11'. Regarding the minimum-position features, as shown in Figure 5c, again, depression patients had greater minimum-position values of MFCC 11' compared to groups of insomnia and CPD. Regarding the range features, as shown in Figure 5d, the severe-depression group had the lowest range values of MFCCs 9, 10, and 10'. Regarding the MSE features, as shown in Figure 5e, the severe-depression group had the MSE range values of MFCC 10'. However, the MSE value of MFCC 2', as shown in Figure 5f, showed an ideal trend in that MSE values were significantly different among the four groups. The values increased from severe to moderate, insomnia, and then CPD.

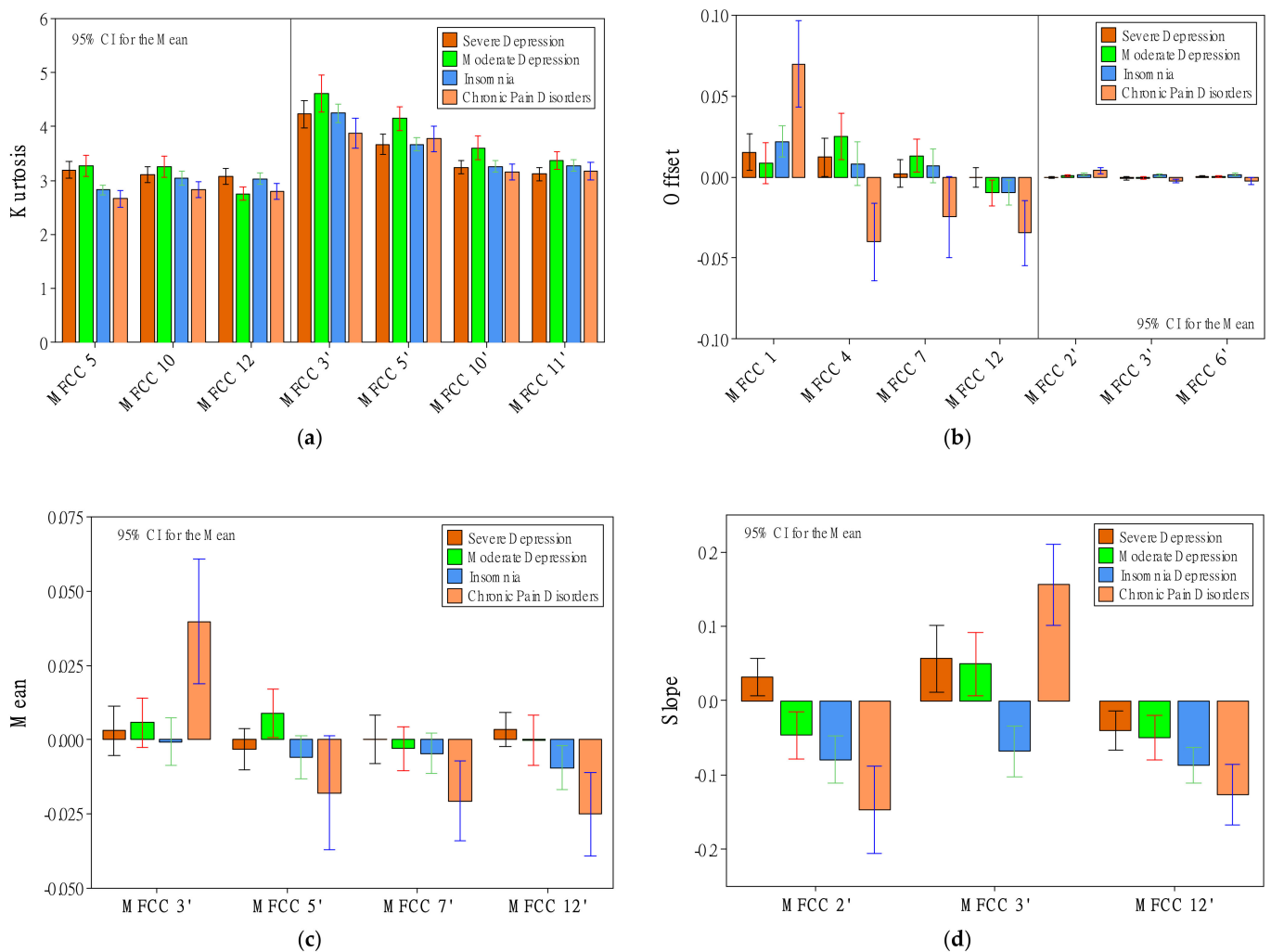


Figure 4. Cont.

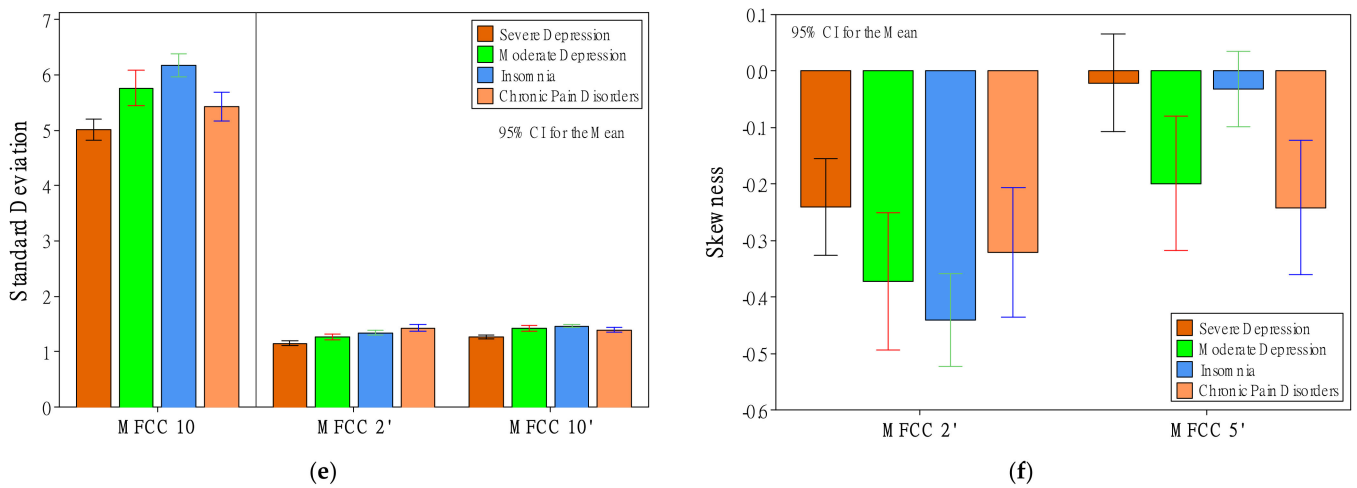


Figure 4. The effect of patient status on kurtosis (a), offset (b), mean (c), slope (d), SD (e), and skewness (f).

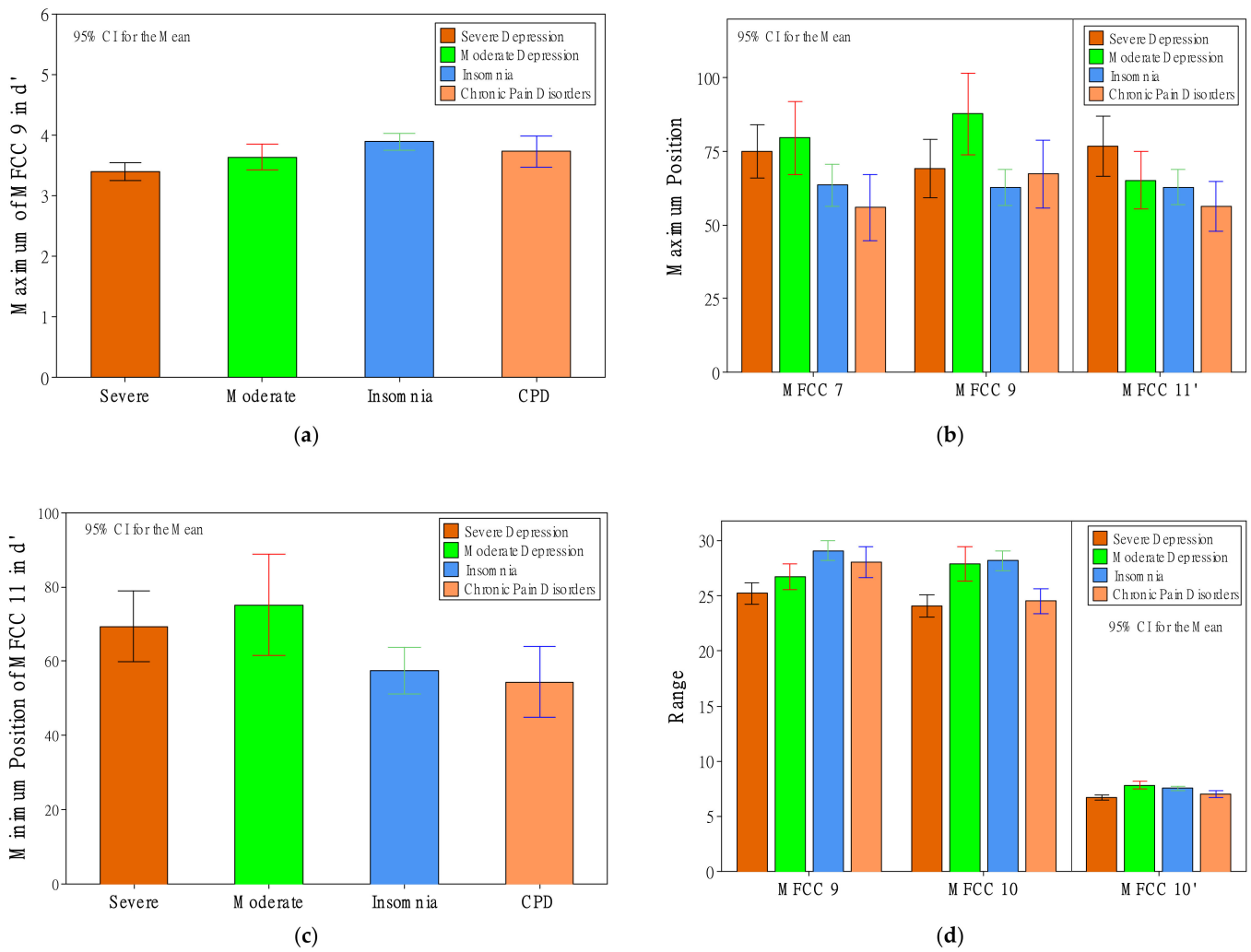


Figure 5. Cont.

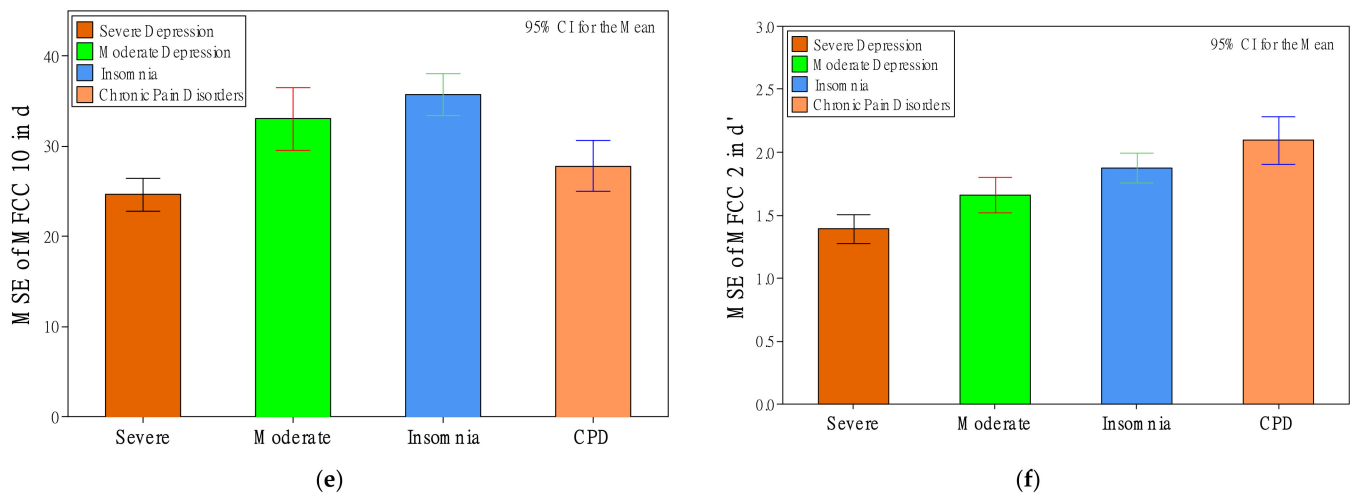


Figure 5. The effect of patient status on maximum value (a), maximum position (b), minimum position (c), range (d), and MSE (e,f).

3.3. Effects of Gender and Interaction Effects of Status and Gender

Regarding RMS, ZCR, F0, and HNR, gender had significant effects (p -value < 0.05) in four d features and five d' features, as shown in Table 2. Among d features, significant relationships mainly occurred in the ZCR category, comprising ZCR-SD, ZCR-slope, and ZCR-MSE features. The other significant relationship was the minimum value of HNR. Among d' features, ZCR-SD and ZCR-MSE features were significantly affected by gender as well. The other significant gender effects were on RMS-skewness, F0-mean, and HNR-mean features.

Regarding MFCC features, gender significantly affected 20 d features and 15 d' features. For both d features and d' features, most of the gender effects (32 out of 35) occurred in MFCCs 7–12. MFCC 1-range, MFCC 1-MSE in d' features, and MFCC 5'-mean in d' features were the only three exceptions.

As shown in Table 3, there was no significant interaction effect of status and gender on RMS-, ZCR-, F0-, or HNR-related features. There was only one interaction effect in d features and eight in d' features.

Table 2. Effects of patient gender on 384 voice features.

Effects	Original Descriptors (d)												De-Differentiated Descriptors (d')												
	Mean	SD	Skewness	Kurtosis	Max Value	Min Value	Max Position	Min Position	Range	Offset	Slope	MSE	Mean	SD	Skewness	Kurtosis	Max Value	Min Value	Max Position	Min Position	Range	Offset	Slope	MSE	
RMS															*	>									
ZCR		*												*		>									•
F0													*		>										
HNR						*		>					*		>										
MFCC	1		>					*			*														
	2				>																				
	3																								
	4							>																	
	5												*										>		
	6									>								>							

Table 2. Cont.

Effects	Original Descriptors (<i>d</i>)												De-Differentiated Descriptors (<i>d'</i>)											
	Mean	SD	Skewness	Kurtosis	Max Value	Min Value	Max Position	Min Position	Range	Offset	Slope	MSE	Mean	SD	Skewness	Kurtosis	Max Value	Min Value	Max Position	Min Position	Range	Offset	Slope	MSE
7	*				●	*				>	>		*									*	*	
8																								
9	*			>	*	*				*				*	*		>					●	*	
10	*					●	●			●			●				*					*	*	
11	*			●	●					●													>	
12		*	>						*			*			■	*	●				>			

>: $p < 0.1$; *: $p < 0.05$; ●: $p < 0.01$; ■: $p < 0.001$.

Table 3. Interaction effects of patient status and gender on 384 voice features.

Effects	Original Descriptors (<i>d</i>)												De-Differentiated Descriptors (<i>d'</i>)											
	Mean	SD	Skewness	Kurtosis	Max Value	Min Value	Max Position	Min Position	Range	Offset	Slope	MSE	Mean	SD	Skewness	Kurtosis	Max Value	Min Value	Max Position	Min Position	Range	Offset	Slope	MSE
RMS																								
ZCR																								
F0																								
HNR																								
MFCC																								
1																								
2															*							>		
3												>										*	*	
4																								
5																								
6																								
7													*											
8							>			>														
9						*																*		
10					>					>						>						*	>	
11																>								
12												*									>	●		

>: $p < 0.1$; *: $p < 0.05$; ●: $p < 0.01$; ■: $p < 0.001$.

4. Discussion

4.1. Voice Features Show the Differences among Patient Groups

This study showed the differences in voice features among patients with severe depression, moderate depression, insomnia, and CPD. To facilitate the comparison of our results with previous findings, Table 4 summarizes the relationships between patient status and critical voice features and how these voice features help in discriminating among the four groups of patients.

Surprisingly, regarding RMS, ZCR, F0, and HNR features, there were no significant status effects on any mean- or range-related features. As shown in Table 4, previous studies reported that compared to other emotions (e.g., anger, happiness, and fear), sadness results in lesser values of RMS-mean [43,44], RMS-range [43], F0-mean [43], F0 range [43], and resonant voice quality (i.e., high HNR-mean) [44]. However, this study did not find significant effects regarding these phenomena. The main reason could be due to different comparisons. Except for Kiss and Vicsi [45], who compared depression with healthy patients, previous studies compared sadness with other distinguishable emotions, such as anger, happiness, and fear. However, the four groups of patients had relatively fewer emotional differences in this study. Nevertheless, this study found statistical properties other than mean and range in d features that can help discriminate depression patients from insomnia patients and patients with CPD. Our findings show that depression groups had greater RMS-skewness, RMS-kurtosis, F0-skewness, and F0-kurtosis values. RMS-skewness and RMS-kurtosis features discriminated between groups of severe depression and CPD and discriminated these two groups from the other two groups (see Figure 1 and Table 4). Moreover, F0-skewness and F0-kurtosis features discriminated between severe- and moderate-depression patients and discriminated these two groups from the other two groups (see Figure 2 and Table 4). Furthermore, this study showed that depression groups had lower slope values of F0 and HNR, helping discriminate severe-depression, moderate-depression, and insomnia groups (see Figure 3 and Table 4). Note that there were no significant differences between the severe-depression patients and CPD patients.

The status effects on MFCC features provide more help in discriminating patient groups. Taguchi, et al. [46] compared 16 original descriptors (mean values of RMS, ZCR, F0, HNR, and 12 MFCC) of 36 major depression patients and 36 healthy controls, and reported that MFCC 2-mean was the only feature significantly affected by depression. The MFCC 2-mean value was relatively higher in depression patients than in the control group. Although we did not find a significant effect on MFCC 2-mean, we found that MFCC 2' provides the most helpful voice features among all the 384 voice features. As shown in Table 4, MFCC 2'-SD, MFCC 2'-offset, MFCC 2'-slope, and MFCC 2'-MSE were ideal voice features that can discriminate between the four patient groups. These features were significantly different among the four groups. The values increased (i.e., MFCC 2'-SD, MFCC 2'-offset, and MFCC 2'-MSE) or decreased (i.e., MFCC 2'-slope) from severe to moderate, insomnia, and then CPD.

Other than MFCC 2, we further found status effects on other MFCC features. As for the rest of the features, we will attempt to discuss them according to how they help in discriminating between our patients. First, as F0-slope and HNR-slope features, MFCC 9'-maximum, MFCC 10-SD, MFCC 10-MSE, MFCC 10'-SD, and MFCC 10'-range were helpful voice features for discriminating among severe-depression, moderate-depression, and insomnia groups. Again, these features may confuse CPD patients with one of the other three groups of patients. Second, MFCC 12-offset and MFCC 11'-maximum features discriminated between groups with severe depression and CPD and discriminated these two groups from the other two groups. Third, MFCC 9-range discriminated between groups of severe depression and moderate depression and discriminated these two groups from the other two groups. Fourth, MFCC 3'-offset, MFCC 3'-slope, MFCC 6'-offset, and MFCC 12'-offset discriminated depression patients from the other two groups and discriminated insomnia patients and patients with CPD from each other as well. Fifth, MFCC 5-kurtosis, MFCC 7-maximum value, and MFCC 11'-minimum position discriminated depression patients from the other two groups. However, this category of features could not discriminate between insomnia patients and patients with CPD. Last, there were many helpful voice features for discriminating patients with CPD from depression and insomnia. Original descriptors were all offset-related, including MFCC 1-offset, MFCC 4-offset, MFCC 7-offset, and MFCC 12-offset. De-differentiated descriptors were mean- and offset-related, including MFCC 3'-mean, MFCC 5'-mean, MFCC 7'-mean, MFCC 12'-mean, MFCC 3'-offset, and MFCC 12'-offset.

Table 4. Comparisons of this study and other studies on the depression effects on voice features.

Feature Category	Original Descriptor (<i>d</i>)			De-Differentiated Descriptor (<i>d'</i>)		
	Emotion Study	Depression Study	This Study	This Study		
RMS	Mean▼ [43,44] Range▼ [43]	-	Skewness▲		-	
			Kurtosis▲		-	
ZCR	-	-	-	-	-	
F0	Mean▼ [43] Range▼ [43]	Mean▼ [45]	Skewness▲		-	
			Kurtosis▲		-	
			Slope▼		-	
HNR	Mean▲ [44]	-	Slope: ▼		-	
MFCC 1	-	-	-	-	-	
MFCC 2	-	Mean▲ [46]	-	-	SD'▼	
			-	-	Offset'▼	
			-	-	Slope'▲	
			-	-	MSE'▼	
MFCC 3	-	-	-	-	Offset'	
			-	-	Slope'	
MFCC 4	-	Max? [19]	-	-	-	
MFCC 5	-	Mean? [19] Max? [19] Min? [19]	Kurtosis▲		-	
MFCC 6	-	-	-	-	Offset'	
MFCC 7	-	-	Max P.▲		-	
MFCC 8	-	Max? [19]	-	-	-	
MFCC 9	-	-	Range▼		Max'▼	
MFCC 10	-	Max? [19]	SD▼		SD'▼	
			MSE▼		Range'▼	
MFCC 11	-	-	-	-	Max'▲	
MFCC 12	-	-	-	-	Min P'▲	
			Offset▲		Slope'▲	

Note: ▼: depression had lesser values; ▲: depression had lesser values; ?: no comparison; bars from left to right in a bar chart show groups of severe-depression, moderate-depression, insomnia, and CPD patients, respectively, and different bar lengths show statistical differences among the four groups of patients.

Solomon, Valstar, Morriss, and Crowe [19] compared differences in original MFCC descriptors between depression and healthy participants when they were asked about a deeply emotional topic to describe their experiences with depression. As shown in Table 4, they reported that the mean of MFCC 5, maximums of MFCCs 4, 5, 8, and 10, and minimums of MFCC 5 were important features between the two groups. Although the critical features reported by Solomon, Valstar, Morriss, and Crowe [19] are not in line with our findings, both studies showed that mean- and SD-related MFCC d features may not be effective voice features for determining depression. Of course, the different results might be due to different comparisons and experimental settings.

4.2. Considerations of Gender Effects

The analyses of gender effects provide further considerations while selecting voice features for discriminating between depression and insomnia. Regarding RMS, ZCR, F0, and HNR, gender had significant effects (p -value < 0.05) on four d features and five d' features (Table 2). While previous studies reported gender effects on the pitch [26,27], voice quality [27], and pitch range [28], we found gender effects on ZCR, HNR (i.e., voice quality), and certain MFCCs (related to specific ranges of pitches). The discrepancy might be due to insufficient and unbalanced numbers of both genders in status groups. As mentioned above, although d' features were proposed by Cao, Xu, and Liu [42] to reduce individual differences, they were not effective in eliminating the effects of gender. Instead, de-differentiated descriptors help differentiate between gender differences with RMS, ZCR, F0, and HNR features. After de-differentiated processing, RMS-skewness, ZCR-MSE, F0-mean, and HNR-mean were additional descriptors that were significantly affected by gender. Regarding MFCC features, as shown in Tables 2 and 3, there were 20 and 15 significant main effects of gender on d and d' features, respectively, and 1 and 8 significant interaction effects of gender and status on d and d' features, respectively. Although the distribution of these significant effects was irregular, most of these main and interaction effects (38 out of 44) occurred in MFCCs 7–12 for both d features and d' features.

While the primary purpose of selecting critical voice features is to discriminate between patient groups, but not gender, the selection of a voice feature significantly affected by both status and gender may reduce discrimination effectiveness. Hence, two strategies are suggested for selecting critical voice features for developing discriminating models. The first strategy is to exclude the voice features significantly affected by gender to avoid interference. The second strategy is to develop individual discriminating models for female and male patients, respectively.

4.3. Contributions and Implications

This preliminary study attempted to discourse on potential voice features for discriminating among the patients with severe depression, moderate depression, insomnia, and CPD. With six severe-depression patients, four moderate-depression patients, ten insomnia patients, and four patients with CPD, the statistical analyses show many significant patient status effects on voice features that were computed by openSMILE, IS09_eotion [41]. We report speech differences between depression and insomnia that were previously scarce in the literature, and our results support the hypothesis that there are voice features that help distinguish between depression and insomnia.

While previous findings mainly focus on mean- and range-related features [43,44], or even ignore the relationships between human status and voice features while developing artificial intelligence (AI) models [47–49], this study assessed more comprehensive relationships between patient status and voice features. While the ultimate objective was to help clinical diagnoses, studies on the rationales behind significant relationships allow us to select critical features (as shown in Table 4) for further developing AI models confidently.

4.4. Limitations and Future Research

Although our analyses explained the effects of depression on certain voice features, the limited number of participants recruited in this study may not provide convincing evidence, even although we attempted to include four CPD patients in the analyses to represent a control group showing no depression and no insomnia. The number the patients and the narrow diversity of symptoms of this group make the generalization of the results difficult. However, despite these limitations, this preliminary study showed the potential of this approach and is encouraging for future studies recruiting an adequate number of patients for each group. With well-selected voice features and sufficient data, we expect the development of AI models for discriminating between patients for use in clinical recommendations. Although this study only discussed the statistically significant relationships with a p -value < 0.05 , the effects with significance < 0.1 could be tested as well when developing models. Tables 1–3 show these voice features with a symbol of ‘^’.

5. Conclusions

This preliminary study aimed at disclosing critical voice features for discriminating between depression and insomnia using statistical analyses. With six severe-depression patients, four moderate-depression patients, ten insomnia patients, and four CPD patients, the statistical results show the patient status effects on certain voice features, demonstrating the potential to develop discriminating models of depression and insomnia using voice features. The findings encourage recruiting an adequate number of patients to confirm these voice features and increasing the number of the data in future studies. The ultimate goal was to apply critical voice features with adequate data to develop a quantitative method to help discriminate between depression patients and insomnia patients.

Author Contributions: Conceptualization: R.F.L., T.-K.L. and Y.-P.L.; Data curation: R.F.L. and Y.-P.L.; Formal analysis: R.F.L. and K.-R.H.; Funding acquisition: T.-K.L. and Y.-P.L.; Investigation: R.F.L., T.-K.L. and K.-R.H.; Methodology, R.F.L., T.-K.L. and Y.-P.L.; Project administration: R.F.L., T.-K.L. and K.-R.H.; Resources: R.F.L. and T.-K.L.; Software: Y.-P.L.; Supervision: R.F.L. and T.-K.L.; Validation: R.F.L. and T.-K.L.; Visualization: R.F.L.; Writing—original draft: R.F.L., T.-K.L. and Y.-P.L.; Writing—review and editing: R.F.L., T.-K.L. and Y.-P.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Science and Technology, MOST110-2628-E-155-001, Hospital and Social Welfare Organizations Administration Commission, Ministry of Health and Welfare, grant number 10710, and Yuan Ze University.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board of Taoyuan General Hospital (protocol code TYGH106015 and date of approval 20 July 2017; protocol code TYGH107069 and date of approval 23 January 2019).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Acknowledgments: We thank Yi-Ting Chia for data collection and processing.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Lim, G.Y.; Tam, W.W.; Lu, Y.; Ho, C.S.; Zhang, M.W.; Ho, R.C. Prevalence of depression in the community from 30 countries between 1994 and 2014. *Sci. Rep.* **2018**, *8*, 2861. [[CrossRef](#)]
2. Morin, C.M.; LeBlanc, M.; Daley, M.; Gregoire, J.; Merette, C. Epidemiology of insomnia: Prevalence, self-help treatments, consultations, and determinants of help-seeking behaviors. *Sleep Med.* **2006**, *7*, 123–130. [[CrossRef](#)] [[PubMed](#)]
3. Ohayon, M.M. Observation of the natural evolution of insomnia in the American general population cohort. *Sleep Med. Clin.* **2009**, *4*, 87–92. [[CrossRef](#)]
4. Pandey, S.; Phillips, B.A. Why is the prevalence of insomnia skyrocketing? And what can be done about it? *Sleep Med.* **2015**, *16*, 555–556. [[CrossRef](#)]

5. Theorell-Haglöw, J.; Miller, C.B.; Bartlett, D.J.; Yee, B.J.; Openshaw, H.D.; Grunstein, R.R. Gender differences in obstructive sleep apnoea, insomnia and restless legs syndrome in adults—What do we know? A clinical update. *Sleep Med. Rev.* **2018**, *38*, 28–38. [[CrossRef](#)]
6. Zhang, Y.; Ren, R.; Lei, F.; Zhou, J.; Zhang, J.; Wing, Y.-K.; Sanford, L.D.; Tang, X. Worldwide and regional prevalence rates of co-occurrence of insomnia and insomnia symptoms with obstructive sleep apnea: A systematic review and meta-analysis. *Sleep Med. Rev.* **2019**, *45*, 1–17. [[CrossRef](#)]
7. Benca, R.M. Diagnosis and treatment of chronic insomnia: A review. *Psychiatr. Serv.* **2005**, *56*, 332–343. [[CrossRef](#)]
8. Taylor, D.J.; Lichstein, K.L.; Durrence, H.H. Insomnia as a health risk factor. *Behav. Sleep Med.* **2003**, *1*, 227–247. [[CrossRef](#)]
9. Taylor, D.J.; Lichstein, K.L.; Durrence, H.H.; Reidel, B.W.; Bush, A.J. Epidemiology of insomnia, depression, and anxiety. *Sleep* **2005**, *28*, 1457–1464. [[CrossRef](#)]
10. Irwin, M.R. Depression and insomnia in cancer: Prevalence, risk factors, and effects on cancer outcomes. *Curr. Psychiatry Rep.* **2013**, *15*, 404. [[CrossRef](#)]
11. Fiorentino, L.; Rissling, M.; Liu, L.; Ancoli-Israel, S. The symptom cluster of sleep, fatigue and depressive symptoms in breast cancer patients: Severity of the problem and treatment options. *Drug Today Dis. Models* **2011**, *8*, 167–173. [[CrossRef](#)] [[PubMed](#)]
12. Harvey, A.G. Insomnia: Symptom or diagnosis? *Clin. Psychol. Rev.* **2001**, *21*, 1037–1059. [[CrossRef](#)]
13. Taylor, D.; Walters, H.; Krebaum, S.; Kraft, D.; Jarrett, R. Does Residual Insomnia Predict Depressive Relapse and Recurrence in Cognitive Therapy Responders? *Sleep* **2004**, *27*, 346–347.
14. Lichstein, K.L.; Durrence, H.H.; Riedel, B.W.; Taylor, D.J.; Bush, A.J. *Epidemiology of Sleep: Age, Gender, and Ethnicity*; Psychology Press: Hove, UK, 2013.
15. DeRubeis, R.J.; Siegle, G.J.; Hollon, S.D. Cognitive therapy versus medication for depression: Treatment outcomes and neural mechanisms. *Nat. Rev. Neurosci.* **2008**, *9*, 788–796. [[CrossRef](#)]
16. Mitchell, M.D.; Gehrman, P.; Perlis, M.; Umscheid, C.A. Comparative effectiveness of cognitive behavioral therapy for insomnia: A systematic review. *BMC Fam. Pract.* **2012**, *13*, 40. [[CrossRef](#)]
17. Okajima, I.; Komada, Y.; Inoue, Y. A meta-analysis on the treatment effectiveness of cognitive behavioral therapy for primary insomnia. *Sleep Biol. Rhythm.* **2011**, *9*, 24–34. [[CrossRef](#)]
18. Girard, J.M.; Cohn, J.F. Automated audiovisual depression analysis. *Curr. Opin. Psychol.* **2015**, *4*, 75–79. [[CrossRef](#)]
19. Solomon, C.; Valstar, M.F.; Morriss, R.K.; Crowe, J. Objective methods for reliable detection of concealed depression. *Front. ICT* **2015**, *2*, 5. [[CrossRef](#)]
20. Hamilton, M. A rating scale for depression. *J. Neurol. Neurosurg. Psychiatry* **1960**, *23*, 56–62. [[CrossRef](#)]
21. Beck, A.T.; Steer, R.A.; Ball, R.; Ranieri, W.F. Comparison of Beck Depression Inventories-IA and-II in psychiatric outpatients. *J. Personal. Assess.* **1996**, *67*, 588–597. [[CrossRef](#)]
22. Asgari, M.; Shafran, I.; Sheeber, L.B. Inferring Clinical Depression from Speech and Spoken Utterances. In Proceedings of the MLSP, Reims, France, 21–24 September 2014; pp. 1–5.
23. Schramm, E.; Hohagen, F.; Grasshoff, U.; Riemann, D.; Hajak, G.; Weess, H.; Berger, M. Test-retest reliability and validity of the Structured Interview for Sleep Disorders According to DSM-III—R. *Am. J. Psychiatry* **1993**, *150*, 867–872. [[PubMed](#)]
24. Roth, T. New developments for treating sleep disorders. *J. Clin. Psychiatry* **2001**, *62*, 3–4. [[PubMed](#)]
25. Chiu, H.-Y.; Chang, L.-Y.; Hsieh, Y.-J.; Tsai, P.-S. A meta-analysis of diagnostic accuracy of three screening tools for insomnia. *J. Psychosom. Res.* **2016**, *87*, 85–92. [[CrossRef](#)] [[PubMed](#)]
26. Munson, B.; Babel, M. The phonetics of sex and gender. In *The Routledge Handbook of Phonetics*; Routledge: London, UK, 2019; pp. 499–525.
27. Simpson, A.P. Phonetic differences between male and female speech. *Lang. Linguist. Compass* **2009**, *3*, 621–640. [[CrossRef](#)]
28. Hancock, A.; Colton, L.; Douglas, F. Intonation and gender perception: Applications for transgender speakers. *J. Voice* **2014**, *28*, 203–209. [[CrossRef](#)]
29. Scherer, K.R. Vocal affect expression: A review and a model for future research. *Psychol. Bull.* **1986**, *99*, 143–165. [[CrossRef](#)]
30. Williamson, J.R.; Young, D.; Nierenberg, A.A.; Niemi, J.; Helfer, B.S.; Quatieri, T.F. Tracking depression severity from audio and video based on speech articulatory coordination. *Comput. Speech Lang.* **2018**, *55*, 40–56. [[CrossRef](#)]
31. Cummins, N.; Scherer, S.; Krajewski, J.; Schnieder, S.; Epps, J.; Quatieri, T.F. A review of depression and suicide risk assessment using speech analysis. *Speech Commun.* **2015**, *71*, 10–49. [[CrossRef](#)]
32. Gumelar, A.B.; Kurniawan, A.; Sooai, A.G.; Purnomo, M.H.; Yuniarno, E.M.; Sugiarto, I.; Widodo, A.; Kristanto, A.A.; Fahrudin, T.M. Human Voice Emotion Identification Using Prosodic and Spectral Feature Extraction Based on Deep Neural Networks. In Proceedings of the 2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH), Kyoto, Japan, 5–7 August 2019; pp. 1–8.
33. Moore II, E.; Clements, M.A.; Peifer, J.W.; Weisser, L. Critical analysis of the impact of glottal features in the classification of clinical depression in speech. *IEEE Trans. Biomed. Eng.* **2007**, *55*, 96–107. [[CrossRef](#)]
34. Alghowinem, S.; Goecke, R.; Wagner, M.; Epps, J.; Gedeon, T.; Breakspear, M.; Parker, G. A Comparative Study of Different Classifiers for Detecting Depression from Spontaneous Speech. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013; pp. 8022–8026.

35. Cohn, J.F.; Kruez, T.S.; Matthews, I.; Yang, Y.; Nguyen, M.H.; Padilla, M.T.; Zhou, F.; de la Torre, F. Detecting Depression from Facial Actions and Vocal Prosody. In Proceedings of the 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, Amsterdam, The Netherlands, 10–12 September 2009; pp. 1–7.
36. Alghowinem, S.; Goecke, R.; Cohn, J.F.; Wagner, M.; Parker, G.; Breakspear, M. Cross-Cultural Detection of Depression from Nonverbal Behaviour. In Proceedings of the 2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), Ljubljana, Slovenia, 4–8 May 2015; pp. 1–8.
37. Heydarifard, Z.; Krasikova, D. Voice and Insomnia: A Daily Study of Underlying Affective and Cognitive Mechanisms. In Proceedings of the Academy of Management Proceedings, Briarcliff Manor, NY, USA, Virtual, 26 July 2021; p. 15460.
38. Endicott, J.; Cohen, J.; Nee, J.; Fleiss, J.; Sarantakos, S. Hamilton Depression Rating Scale: Extracted from regular and change versions of the Schedule for Affective Disorders and Schizophrenia. *Arch. Gen. Psychiatry* **1981**, *38*, 98–103. [[CrossRef](#)]
39. Zimmerman, M.; Martinez, J.H.; Young, D.; Chelminski, I.; Dalrymple, K. Severity classification on the Hamilton depression rating scale. *J. Affect. Disord.* **2013**, *150*, 384–388. [[CrossRef](#)] [[PubMed](#)]
40. Che, H.H.; Lu, M.L.; Chen, H.C.; Chang, S.W.; Lee, Y.J. Validation of the Chinese version of the Beck Anxiety Inventory. *Formos. J. Med.* **2006**, *10*, 451–452.
41. Schuller, B.; Steidl, S.; Batliner, A. The Interspeech 2009 Emotion Challenge. In Proceedings of the Tenth Annual Conference of the International Speech Communication Association, Brighton, UK, 6–10 September 2009.
42. Cao, W.-H.; Xu, J.-P.; Liu, Z.-T. Speaker-Independent Speech Emotion Recognition Based on Random Forest Feature Selection Algorithm. In Proceedings of the 36th Chinese Control Conference (CCC), Dalian, China, 26–28 July 2017; pp. 10995–10998.
43. Gangamohan, P.; Kadiri, S.R.; Yegnanarayana, B. Analysis of Emotional Speech—A Review. In *Toward Robotic Socially Believable Behaving Systems—Volume I*; Intelligent Systems Reference Library; Springer: Cham, Switzerland, 2016; pp. 205–238.
44. Murray, I.R.; Arnott, J.L. Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *J. Acoust. Soc. Am.* **1993**, *93*, 1097–1108. [[CrossRef](#)] [[PubMed](#)]
45. Kiss, G.; Vicsi, K. Seasonal affective disorder speech detection on the base of acoustic phonetic speech parameters. *Acta Univ. Sapientiae Electr. Mech. Eng.* **2015**, *7*, 62–79.
46. Taguchi, T.; Tachikawa, H.; Nemoto, K.; Suzuki, M.; Nagano, T.; Tachibana, R.; Nishimura, M.; Arai, T. Major depressive disorder discrimination using vocal acoustic features. *J. Affect. Disord.* **2018**, *225*, 214–220. [[CrossRef](#)]
47. Sardari, S.; Nakisa, B.; Rastgoo, M.N.; Eklund, P. Audio based depression detection using Convolutional Autoencoder. *Expert Syst. Appl.* **2022**, *189*, 189–116076. [[CrossRef](#)]
48. Kwon, N.; Kim, S. Depression Severity Detection Using Read Speech with a Divide-and-Conquer Approach. In Proceedings of the 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Guadalajara, Mexico, (Virtual), 1–5 November 2021; pp. 633–637.
49. Zhao, Y.; Xie, Y.; Liang, R.; Zhang, L.; Zhao, L.; Liu, C. Detecting Depression from Speech through an Attentive LSTM Network. *IEICE Trans. Inf. Syst.* **2021**, *104*, 2019–2023. [[CrossRef](#)]