# Explainable AI-based clinical decision support system for hearing disorders

# Katarzyna A. Tarnowska, Ph.D. , Brett C. Dispoto, B.S., Jordan Conragan, B.S. San Jose State University, San Jose, CA

## Abstract

In clinical system design, human-computer interaction and explainability are important topics of research. Clinical systems need to provide users with not only results but also an account of their behaviors. In this research, we propose a knowledge-based clinical decision support system (CDSS) for the diagnosis and therapy of hearing disorders, such as tinnitus, hyperacusis, and misophonia. Our prototype eTRT system offers an explainable output that we expect to increase its trustworthiness and acceptance in the clinical setting. Within this paper, we: (1) present the problem area of tinnitus and its treatment; (2) describe our data-driven approach based on machine learning, such as association-and action rule discovery; (3) present the evaluation results from the inference on the extracted rule-based knowledge and chosen test cases of patients; (4) discuss advantages of explainable output incorporated into a graphical user interface; (5) conclude with the results achieved and directions for future work.

## Introduction

**Background** Tinnitus commonly referred to as "ringing in the ears", is a phantom auditory disorder estimated to affect about 15% of the global population, with patients existing on a wide spectrum of symptom severity<sup>1</sup>. Although there is currently no known cure for tinnitus, there do exist management techniques; in particular, tinnitus retraining therapy (TRT) has been shown to yield high success rates in clinical trials<sup>2</sup>. TRT is based on the neurophysiological model of tinnitus and is administered as a combination of clinical counseling and sound therapy<sup>3</sup>. The goal of TRT is the habituation of tinnitus, meaning that patients who have undergone successful treatment will still be aware of the phantom auditory signal, but will no longer be agitated by its existence. Despite its clinical success, the infrastructure surrounding tinnitus retraining therapy has a few obstacles to overcome until the treatment can be widely and efficiently adopted by medical practitioners. TRT is a highly personalized treatment that requires patients to receive care over a substantial period of time in order to be successful. Presently, physicians must be deeply experienced in TRT in order to provide effective treatment. It is relatively a niche treatment and it takes years of practice to establish the expertise necessary. Throughout treatment, an effective physician will assess patient progress and modify treatment protocol accordingly. Physicians generally make these decisions based on their knowledge of TRT heuristics along with a deeper knowledge of the neurophysiological model of tinnitus: both of which are lacking in both general-care and ear-nose-throat/audiology practitioners.

**Objectives** For the reasons outlined above, a data-driven knowledge-based clinical decision support system (CDSS) for the diagnosis and treatment of tinnitus was proposed<sup>4</sup>. This work implements and evaluates the proposed strategy by developing a graphical user interface (GUI) and a knowledge base (KB) with the inference component tested against patient cases. The requirements for a CDS system were determined with the following objectives:

- make secondary use of tinnitus electronic health records with the goal of discovering novel actionable patterns in TRT with regard to treatment outcomes,
- present interpretable and user-friendly real-time clinical advice in TRT delivery, and
- integrate seamlessly with an audiology clinical workflow increasing the likelihood of use.

In turn, it is expected that a system meeting the above criteria will help meet the broader clinical goals:

- wider adoption of an effective tinnitus retraining therapy,
- improved efficiency in the management of tinnitus symptoms, and
- improved effectiveness of treatment outcomes resulting in better care for tinnitus patients.

#### Methods

**Explainable AI for interpretability** Despite the promise of assisting human decision making through data-driven approaches, system users often find it challenging to explain and interpret the AI-algorithms behind the transformations of the system's input into a recommendable output. Physicians holding medical responsibility can hardly trust the system's results without an explanation of its underlying decision-making process. To address this issue, we propose an explainable approach for clinical system development. The CDS provides natural-language justification for its output, which is relatively simple to understand by humans. The methods follow the *precision medicine* and *personalized medicine* paradigms to help audiologists better diagnose the category of a hearing problem and treat it in an effective quantifiable manner.

**Three-tier implementation for scalability** The prototype implementation of the proposed explainable CDS strategy for tinnitus diagnosis and treatment follows a three-tier system architecture (see Figure 1). It is expected that the proposed Java-based platform for CDS will result in wider adoption of the generic CDS with seamless integration into EHR systems, as well as more maintainable, organized, and flexible changes in the CDS component. The so-called *Model-View-Controller* system design pattern was utilized to develop a CDS prototype, called eTRT (*electronic Tinnitus Retraining Therapy*). Data *Model* and back-end rule-based logic are separated from the front-end human-computer interaction module (the *View*), and the synchronization between the Model and the View data is handled through the *Controller*.



**Figure 1:** Three-tier system architecture for eTRT - a prototype CDS for tinnitus diagnosis and therapy. The user interface layer (the *View*) was developed in a graphical toolkit for Java Swing. Data *Model* is a separate system component and includes a knowledge base with a rule engine implemented in the Java expert system shell (JESS). The *Controller* synchronizes the View and the Model.

**Graphical interface for usability** The user interface (UI) of the system is of critical importance, being the only mode of interaction between the physician and the underlying CDS model. Before the system is applied in the health-care environments in a usable manner, a user-friendly graphical UI has to be designed and implemented. Within the proposed solution, the CDS component integrates into clinical TRT workflow as depicted in Figure 2. The prototype GUI was developed in a cross-platform graphical toolkit for Java, called Swing, with customary component extensions for screen development. The developed UI supports clinical processes in:

- Storing and managing the data related to:
  - Tinnitus patients demographics, medical history, audiology evaluations, and structured interviews<sup>6</sup>;
  - TRT visits diagnoses, treatment applied (sound therapy and counseling), and the outcome evaluation with standardized forms, such as *Tinnitus Handicap Inventory* (THI)<sup>7</sup> and *Tinnitus Functional Index* (TFI)<sup>8</sup>.
- Providing evidence-based diagnostic and treatment decision support with explanations and quantifiable predictive outcomes.

The sample GUI screens of the CDS supporting an audiologist in consultation with a new tinnitus patient consultation are presented in Figure 3.



Figure 2: The design of human-computer interaction that integrates eTRT in a clinical workflow.

🖻 eTRT	- 0	×	🐖 eTRT	– 🗆 X
Welcome	Select an	option	Patient Menu	Select an option
Patients	<b>Visits</b>		Add Patient Back Mon Dec 30 11:03:40 PS	View/Edit Patients
	<b>W</b>		Add New Patient Patient Information Additional Information Connecting	Pre-filed Information THC# Pre-filed Date Mon Dec 30 20 Tinnihus/hyperacuis
Analytics	Other		Work Status	Tinnitus Etiology noise exposure
			Educational Degree	Hyperacusis Onset
Diffice of Dr. Jastreboff Aon Dec 30 11:03:37 PST 2019		×	Additional Comments Additional Comments	Holomaticine Fromby
Visit Patient Information			Cancel	Now Visit Sa
Visit ID: 1 Date: Wed Jan 15			🕅 eTRT	-  X
Patient: Test Patient 1 Visit 0 THC: 2 Diagnosis	:5		Treatment Recom	mendations
Interview > Audiology > M	edical Other Diagn	ose	Patient: Test Patient 1	THC: 25
Problem: HT Category: 4	Protocol: 4		Primary Recommendation Actions: change instrument t	io GHS
Follow-up: A 🗸 Instrument: GHH 🗸	Next Visit: 2/18/20 R	EM:	Gain: 41.0%	
Comments: devices sent to pt			Explanation: Recommend ch was induced by	anging GHH to GHS for a male whose tinnitu noise
Instrument Details REM Co Details I	Details Recommendation	nd it	Other Recommendations Recommended Actions	Improvement Explanation
Cancel		Save	Back	Done

**Figure 3:** Sample GUI-based clinical TRT workflow with eTRT system: (1) new patient data entry; (2) new visit data entry; (3) treatment recommendations inferred from the current patient/visit data and eTRT knowledge base.

**Rule-based knowledge representation for flexibility** Within the logic layer, a flexible and interpretable rule-based data model was incorporated which is expected to augment the clinician's knowledge to provide better healthcare for tinnitus. The clinical decision rules are automatically exported from the output files of the data mining software (LISp-Miner) and encoded into CDSS using the Encoder component (see Figure 4). This architecture allows for an efficient update of the knowledge base once new raw data on treatment outcomes becomes available (self-adaptation). The preliminary knowledge discovery used a clinical dataset describing 555 unique patients and 3000 TRT visits, recorded by Dr. Jastreboff over several years of clinical practice at Tinnitus and Hyperacusis Center of Emory University School of Medicine<sup>5</sup>. The dataset was anonymized in a clinic before sharing, therefore no IRB protocol was involved. The dataset was pre-processed, including data cleansing, data transformation, and feature extraction, described in more detail in<sup>5</sup>. The proposed machine learning models for decision support in the diagnosis and treatment include *association rules*. The extracted rules were validated with Dr. Jastreboff as a domain expert in TRT.



Figure 4: The high-level architecture of a data-driven CDSS for tinnitus with the *Encoder* component automatizing knowledge engineering, *Knowledge Base* with *Inference Engine* providing decision support in diagnosis and treatment.

**Diagnostic rules** TRT protocol differentiates five categories of the hearing problem: category 0 indicates tinnitus as a minimal problem; 1–tinnitus as a significant problem; 2–tinnitus as significant and hearing loss as existing; 3–tinnitus irrelevant, but hyperacusis (decreased sound tolerance) is significant; 4 is characterized by prolonged tinnitus/prolonged exacerbation of hyperacusis. The accurate categorization is critical, as it determines the further treatment protocol, and therefore the effectiveness of TRT. The diagnosis process defined by the TRT protocol was modeled with a decision rule concept, which is defined as the following:

A **decision rule** is a rule *r* in the form ( $\phi \Rightarrow \delta$ ), where  $\phi$  is called *antecedent* (or assumption) and  $\delta$  is called *descendant* (or thesis). Each rule is characterized by *support* and *confidence*. *Support(r)* is defined as the number of objects matching the rule's antecedent. *Confidence(r)* is the relative number of objects matching both the rule's antecedent and descendant of the rule. Listing 1 presents a sample diagnostic rule encoded in eTRT using JESS syntax. The diagnostic pattern consists of the premises and the conclusion, a confidence metric, and a natural-language explanation.

Listing 1: A sample diagnostic rule encoded in eTRT using JESS syntax.

```
(defrule C1-HLpr-Hpr-Tpr
```

```
(Interview {hpr >= 0 && hpr < 0.5 & hlpr >= 0 && hlpr < 0.5 && tpr >= 6 && tpr < 8}) => (add (new Diagnosis 1 85% "hyperacusis and hearing loss indicated as low, but tinnitus indicated as high")))
```

**Treatment rules** The treatment decisions in TRT are supported with actionable knowledge encoded into eTRT. The action rule concept is a novel way in machine learning proposed by Ras and Wieczorkowska<sup>10</sup>. Since its introduction in 2000, the application of action rules was proposed, among others, for business, medicine, and music indexing<sup>10,11,12,13</sup>. Action rule is defined as the following logical term:

Action rule r is a term  $[(\omega) \land (\alpha \rightarrow \beta) \Rightarrow (\theta \rightarrow \psi)]$ , where  $(\omega \land \alpha) \Rightarrow \theta$  and  $(\omega \land \beta) \Rightarrow \psi$  are classification rules,  $\omega$  is a conjunction of stable attribute values,  $(\alpha \rightarrow \beta)$  shows changes in flexible attribute values, and  $(\theta \rightarrow \psi)$  shows the desired effect of the action. Now we give an example assuming that a is a stable attribute, b is a flexible attribute, and d is a decision attribute. Terms  $(a, a_2), (b, b_1 \rightarrow b_2), (d, d_1 \rightarrow d_2)$  are examples of *atomic actions*. Expression  $r = [(a, a_2) \land (b, b_1 \rightarrow b_2)] \Rightarrow (d, d_1 \rightarrow d_2)$  is an example of an action rule saying that if value  $a_2$  of a in a given object remains unchanged and its value of b will change from  $b_1$  to  $b_2$ , then its value of d is expected to transition from  $d_1$  to  $d_2$ . Listing 2 presents a sample action rule extracted from tinnitus datasets and encoded into eTRT using JESS syntax. In this rule, the stable part is defined by gender (male) and etiology (noise), which constitutes the fixed patient's profile. The flexible (changeable) attribute is the instrument used in sound therapy and the recommendation indicated by the rule's conclusion is changing the instrument's model from GH hard to GH soft to improve THI score with 80% confidence. The explanation of the rule is encoded in natural language by the *Encoder* component.

Listing 2: A sample action rule for treatment encoded in eTRT using JESS syntax.

(defrule Gm-NTI-GHH

(Patient {gender == "M" && tEtiology == "NTI"})
(Instrument {it == "GHH"})
=> (add (new Treatment "GHH -> GHS" 80 "change GHH to GHS
for a male whose tinnitus was induced by noise exposure")))

Action rule modeling is especially promising in the field of medical data, as a doctor can examine the effect of treatment decisions on a patient's improved state. This technique is also particularly useful for building knowledge-based decision support systems since it provides actionable advice needed by practitioners.

**Inference engine** There are hundreds of such rules extracted from the dataset in the knowledge discovery process. Each rule represents a small piece of the expert's knowledge available from the clinical dataset through machine learning. Rules, such as these presented in Listings 1 and 2, are numerous which presents a challenge for encoding them in the eTRT knowledge base. The manual encoding would be very inefficient and error-prone. Therefore, the automatic conversion was implemented as the Encoder component, which extracts rules from the LISp-Miner output files, parses them, and translates them into JESS syntax, including explanations encoded in natural language. Once rules are encoded into KB by Encoder, the *inference* or *deduction* controls the application of the rules to the patient cases. The method for the inference applied within this research is based on an efficient pattern-matching algorithm called Rete<sup>14</sup>, which is provided by the JESS framework. The deduction algorithm matches the current patients' data entered into the system with the machine-learned diagnostic/treatment rules in the knowledge base. If the left-hand side of the rule (antecedent) is matched with the current patient, the right-hand side (consequent) of a rule is executed. Consequent clauses decide about the diagnosis/treatment decision suggested and displayed to the physician via GUI.

#### Results

**Association patterns discovery** In the experiments on decision rule discovery LISp-Miner 4ft-Miner module was used<sup>9</sup>. In the pattern formulation, the descendant of the rule (the decision attribute) was defined as the TRT category (0-4). For the antecedent part, attributes describing a patient, such as demographics etiology, pharmacology, audiometry, severity, and effect on life were chosen. The examples of the found associations between the variables describing the tinnitus patient and the TRT category are provided in Table 1. These rules are interpreted as follows:

- If hyperacusis  $H_{pr}$  and hearing loss  $HL_{pr}$  not indicated as problems, but tinnitus  $T_{pr}$  indicated a problem *then* a patient falls under Category 1 with 85% confidence.
- If audiometric values of  $L_2$  (audiogram at 2kHz for the left ear) is greater or equal to 50 and  $R_6$  (audiogram at 6kHz for the right ear) is less or equal to 75, *then* a patient falls under Category 2 with 87% confidence.

 Table 1: Sample association rules between patient characteristics and the TRT category.

Diagnostic association rule	Confidence
$H_{pr}(<0;0.5)) \land HL_{pr}(<0;0.5)) \land T_{pr}(<6;8)) \Rightarrow Category(1)$	85%
$L_2 \ge 50 \land R_6 \le 75 \Rightarrow Category(2)$	87%

Actionable patterns discovery In the experiments on actionable pattern discovery LISp-Miner Act4ft module was used. The actionable pattern was formulated by choosing the patient's profile (e.i. age, gender, etiology) in the stable (fixed) part, and treatment methods as flexible (or changeable in the course of treatment). Treatment methods in TRT include sound therapy with different categories, types, and models of instrument, real-ear measurements (REM) supporting instrument fitting set to optimal numerical parameters, and counseling delivered within individual therapy. The goal of the actionable pattern (or the desired effect) is to decrease the severity of tinnitus as indicated by the *total score* of tinnitus handicap inventory. To indicate the change, the additional, temporal attribute was added and its values were imputed as calculated distance-based percentage change metric between visits. Table 2 presents examples of extracted action rules, which indicate changes in the settings of the instruments, or changing the length of a particular treatment that lead to patient's improvement. These rules are interpreted as follows:

Table 2: Examples of discovered action rules for recommending treatment in TRT.

Treatment action rule	Conf.
$T_{side}(yes) \land OMTI(yes) : (Ins_{vis(01)}GHH \to V) \land FU(0 \to T) \Rightarrow Ch(better)$	82%
$Ins(SG): (Mix_{RSL}(<11;12) \rightarrow <9;10)) \Rightarrow Ch(better)$	100%
$FU(A) \land Ins_{vis(01)}(GHI) \land Freq_{LE}(<3000; 3150)) : (treat(<5; 6) \to <6; 8))$	
$\Rightarrow Ch(better)$	88%

- If tinnitus was induced by other medical condition (OMTI), and as a side effect of taking medications ( $T_{side}$ ), then changing the sound generator model GH hard (GHH) to the Viennatone model (V) at the first visit and changing the follow-up contact to the telephone-based (T) improves patient with 82% confidence.
- If the current treatment involves sound generator SG, then changing mixing point for the right ear  $Mix_{RSL}$  from < 11; 12) to < 9; 10) improves a patient's state with 100% confidence.
- If the current treatment involves audiology (FU(A)) with the GHI instrument, and frequency in the left ear measured by REM - $Freq_{LE}$  in the range of < 3000; 3150) then prolonging that treatment from 5-6 weeks to 6-8 weeks brings improvement with 88% confidence.

**Inference evaluation** The evaluation objective is to determine whether the prototype system does what it was intended to and at an adequate level of accuracy. The system is expected to generate accurate, patient-specific, and interpretable clinical suggestions. This will encourage efficient and effective use of tinnitus retraining therapy for the management of hearing disorders. The evaluation method includes:

- 1. Identifying a set of representative test cases of patients from the dataset not used for building the model.
- 2. Running inference on the chosen test cases entered into the system (see Figures 5–7).
- 3. Performing quantitative and qualitative evaluation of the system based on the results from the above.

The metrics used for this evaluation of the system include:

- Accuracy the number of correct predictions versus the total number of predictions. The predictions were compared with the actual diagnosis/treatment decisions from Dr. Jastreboff, who is considered the "gold standard" in TRT, as the founder and years-long practitioner of the method<sup>2</sup>.
- Coverage the number of test cases matched against the knowledge base.
- Interpretability if the recommendations were provided with human-understandable explanations.

**Test cases** The goal was to identify the smallest possible *representative* set of test cases. A test patient for each etiology and each category of the hearing problem was selected from the test dataset (see Table 3). The chosen test cases reflect the heterogeneity of the hearing problem and patient profile.

**Table 3:** Patient test cases - patient profile, etiology of their hearing problem, the actually diagnosed category and the actual treatment protocol as determined by the TRT founder.

Test case	Patient profile	Etiology	Diagnosis	Treatment protocol
1	male, age 38	noise exposure	Category 4	Category 4
2	male, age 49	ear surgery	Category 3	Category 3
3	female, age 77	hearing loss	Category 2	Category 2
4	male, age 53	stress-related	Category 1	Category 1
5	male, age 36	car accident	Category 0	Category 1



**Figure 5:** The diagnostic/treatment inference results for test case 1 (noise-based, middle-aged male) based on audiometry: (1) primary diagnosis of category 4 with 66.7%, and (2) treatment recommendation for changing the instrument type with the expected decrease in tinnitus severity by 41 percentage points.

Sound Tolerance		eTRT -	- X	
Oversensitivity	⊖Yes ⊖No			
o reisensiarity	0.111.0.111	Diagnosis Inference		
Physical Discomfort	⊖Yes ⊖No	Patient Information		
Description of Troublesome Sound	3			
	Concerts Restaurants	Patient: Test Patient 2	HC· 1	
	Housekeeping			
	Driving Childcare	Primary Diagnosis		
Activities Prevented/Affected	Movies Sports	Category: 3		
		Oulogory. 5		
	Social Work	Confidence: 100.0%		
	Church Other			
		LL3 in range 85:91), hyperacusis anno	vance >=8. H effect on	
Severity	0 1 2 3 4 5 6 7 8	Explanation: life $\geq = 8$ and H severity $\geq = 7.5$		
	0 1 2 5 4 5 0 7 8			
Annoyance		Other Diagnoses		
	0 1 2 3 4 5 6 7 8	<sup>9 10</sup> Category Confidence Explanation	1	
Effect on Life		□ IO0.0 LL3 in range <85;91), hyperacusis effect on life >= 8,and		
0 1 2 3 4 5 6 7 8 9 10		<b>•</b>		
Additional Comments		Back	Done	
		Duok		

**Figure 6:** The diagnostic inference for test case 2 based on audiometry/initial interview. The explanation for the category 3 with 100% confidence included a high score for hyperacusis as a problem.

🕅 eTRT	- 🗆 ×	
Visit in Progress	THC: 20 Search	Diagnosis Inference Patient Information
Patient Information		Patient: Test Patient 4 THC: 20
Visit ID: 203 Date: Thu Jan 09 2020		Primary Diagnosis
Patient: Test Patient 4	Visit 3	Category: 1 Confidence: 04.4%
Diagnosis		Explanation: R3 in range 15;20) and T annoyance >= 8
Interview > Audiology > Medical	Other Diagnose	Explanation           1         0.36.         R3 in range <15.20 and T sevently >= 8           1         73.8.         Patient feels frustrated because of tinnitus(TH-E-10) but tinnit.
Problem: T Category: 1	Protocol: 1	P etrt – X
Follow-up: C V Instrument: GHS V N	Next Visit:	Treatment Recommendations
Comments: reinstruction re:sound level and use of SG		Patient: Test Patient 4 THC: 20
Treatment		Drimany Recommendation
Instrument REM Counselin	g 💊 Recommend	Actions: change instrument from GHS to GHI, use it for 9-14 weeks
Details Details Details	Treatment	Gain: 34.8%
Cancel	Save	Explanation: Cat1, current length greater than 22 weeks

**Figure 7:** The diagnostic/treatment inference results for test case 4: (1) category 1 was inferred based on the audiometry results and initial interview (annoyance over tinnitus high); (2) recommendation included the change of the sound instrument from GH soft and shorten its application time to 9-14 weeks with an expected gain of 34.4 percentage points.

**Result summary** Tables 4 and 5 provide the diagnostic and treatment inference results for all test cases. The diagnosis prediction was 80% accurate and covered 100% of cases. The average confidence in the primary diagnosis inference was 83.51 %. The only incorrect prediction was for test case 5. After closer investigation, this case was annotated by the physician as a "discrepancy in information" in interview data, and "inconsistent results" is audiological evaluation, which are the reasons that misled the predictive model (as an "outlier" data point). Moreover, the actual protocol followed was the same as for the category predicted by the system. The treatment recommendations were generated for 3 out of 5 patient test cases. The other two cases were not covered, that is, no action rule was matched with the patient profile, due to a limited number of rules encoded manually in KB at the time of testing. For all the tested cases, both the diagnostic and the treatment recommendations were explained with a human-comprehensible message/reason. The explanations were provided by means of the premises of the rules in KB that were matched against the current patient's profile/visit data. The predictions' probabilities were quantified by means of the matched rules' confidence metric.

Test case	Actual	Predicted	Pred. Conf.	Explanation
1	Cat 4	Cat 4	66.7%	LSD <= 100, L4 < 10, and LL3 < 75
2	Cat 3	Cat 3	100%	LL3 in < 85; 91), Hyper. Annoy >=8,
				$H_{Effect} >= 8$ , and H Sev $>= 7.5$
3	Cat 2	Cat 2	96.2%	LR8 >= 999, R6 >= 75, and $T_{sv}$ >= 8
4	Cat 1	Cat 1	94.4%	LL3 in $< 15; 20$ ) and Tin. annoy. $>= 8$
5	Cat 0	Cat 1	60.3%	patient often irritable by tinnitus (E14)
				and tinnitus makes him anxious (E22)

**Table 4:** Actual versus predicted TRT category for the test patients, the confidence, and the explanation behind the prediction.

**Table 5:** Recommendation generated for the test cases 1,4, and 5. Each recommendation is supported by a predicted gain in the patient's improvement and explanation based on the patient's personalized profile. Due to a limited knowledge base at the time of testing, no action rules matching profiles in test cases 2 and 3 were found, but are expected once a full KB is built.

Test case	Recommended Action(s)	Gain	Explanation
1	Change instrument from GHH to GHS	41 pp	a male whose tinnitus
			was induced by noise
4	Change instrument from GHS to GHI,	34.8 pp	Cat1, instr. duration
	use it for 9-14 weeks		greater than 22 weeks
5	Change Freq LE from $< 2800; 3000$ )	8.4 pp	Instrument used GHS
	to < 2670; 2800) in REM		

#### Discussion

The application of novel approaches in actionable data mining helps uncover links between clinical variables, with the goal to develop optimal strategies for tinnitus management. This is expected to open new horizons for TRT, which does not have a stagnant protocol but continues to evolve based on information gathered from treatments of patients<sup>2</sup>. The proposed CDS strategy was implemented and preliminarily evaluated in the eTRT prototype. The next step is the usability study with the target users and later within real patient consultation. Its future deployment in the clinical setting is expected to improve TRT delivery, tinnitus management, and clinical outcomes, such as THI total score and habituation time. The system's output is transparent, and as such, will increase its acceptance in the clinical setting. The clinical decision-makers offered the predictive models, which cannot answer questions, such as "Why did you predict this patient to fall into this category?" or "Why did you recommend these treatment actions for this patient?" will be hesitant to use the system. The lack of transparency is a problem for the clinician who wants to understand the way the model works to help them improve their service. While AI technologies are powerful, the adoption of these algorithms in health care has been slow because doctors and regulators cannot verify their results. For fields such as health care, where mistakes can have catastrophic effects, the "black-box" aspect of AI makes it difficult for doctors and regulators to trust it. Explainable clinical systems have the opportunity to disrupt the health care sector because of their ability to diagnose and produce results efficiently. The trend of explainable AI has grown in recent years and looks set to continue.

From the design phase of the eTRT prototype, through implementation up to the evaluation, the researchers' efforts have focused on the human-computer interaction and the AI explainability. The result is the prototype system that utilizes the explainable AI (XAI) approach by providing justification for its results and is human-understandable. Through rule-based inference, intelligent decisions are made and the automatic decision-making process is traced. The advantage of the approach is not only being similar to human reasoning but also involving a minimal set of predictor variables to infer diagnosis or treatment and therefore reducing cognitive load on the clinical user. The end result is developing a prototype that not only generates new insights but is also more credible. The final outcome of this research is expected to be the realization of "precision medicine" or "personalized medicine" in tinnitus practice, which takes into account individual variability in demographics and etiology of hearing disorder for each patient.

#### Conclusion

Our contribution is implementing a novel evidence-based CDS strategy for a niche medical condition. This promising tool will help physicians optimize diagnosis and treatment for tinnitus, and will be particularly useful for practitioners, not that experienced or familiar with TRT. The developed prototype CDSS supports clinical decision-making with a high degree of accuracy, covers a wide range of heterogeneous cases, and provides natural language-based interpretation within GUI. The system also provides a degree of certainty for the suggestions and alternatives for a primary diagnostic/treatment hypothesis. The limitations of this study were identified for the future work directions, which includes the following:

- Usability study involving actual users of the system (audiology clinicians) to determine usefulness and acceptance of the system in the clinical setting and testing the system in a real-time patient consultation;
- Machine learning tuning and further model calibration to optimize predictive accuracy and coverage;
- Expanding the knowledge base with more rules and including data from more TRT clinical experts and clinics;
- Developing a strategy to integrate the CDS component within electronic health records (EHR) systems of audiology clinics.

Although this work is specific to audiology, the proposed data-driven AI-based approach to developing a knowledgebased clinical decision support system is applicable to a wide range of disease, where the lack of experts blocks the delivery of effective treatment for the majority of patients.

#### References

- 1. American Tinnitus Association. Understanding the facts [Internet], Available from: www.ata.org/understanding-facts. [Cited 29 Jul 2020]
- 2. Jastreboff PJ. 25 years of tinnitus retraining therapy. HNO. 2015 Apr;63(4):307-11.
- 3. Jastreboff PJ, Hazell JWP. Tinnitus retraining therapy: implementing the neurophysiological model. Cambridge: Cambridge University Press; 2004.
- 4. Tarnowska KA, Ras ZW, Jastreboff PJ. Decision support system for diagnosis and treatment of hearing disorders. The case of tinnitus. Series in Computational Intelligence 685, Springer; 2017.160 p.
- 5. Tarnowska KA, Ras ZW, Jastreboff PJ. Mining for actionable knowledge in tinnitus datasets. In: Thriving Rough Sets. International: Springer; 2017. p.367-395.
- 6. Jastreboff MM, Jastreboff PJ. Questionnaires for assessment of the patients and treatment outcome, Sixth International Tinnitus Seminar, 1999.
- 7. Newman CW, Wharton JA, Jacobson GP. Retest stability of the tinnitus handicap questionnaire. Ann Otol Rhinol Laryngol 1995;104:718-23.
- 8. American Academy of Audiology. Tinnitus functional index. [Internet] Available from: https://www.audiology.org/news/tinnitus-functional-index [Cited 29 Jul 2020]
- 9. Simunek M. LISp-Miner control language description of scripting language implementation. Journal of Systems Integration. 2014;5(2):28–44.
- 10. Ras ZW, Wieczorkowska A. Action-rules: how to increase profit of a company. In: Principles of data mining and knowledge discovery. Springer; 2000. p.587–592.
- 11. Tarnowska K, Ras Z, Daniel L. Recommender system for improving customer loyalty. Studies in Big Data 55, Springer; 2019. 124 p.
- 12. Wasyluk H, Ras ZW, Wyrzykowska E. Application of action rules to HEPAR clinical decision support system. Experimental and Clinical Hepatology Bd. 2008;4(2):46–48.
- 13. Ras ZW, Wieczorkowska A. Advances in music information retrieval. Studies in Computational Intelligence, Springer; 2010.
- 14. Forgy CL. Rete: A fast algorithm for the many pattern/many object pattern match problem. Artificial intelligence. 1982;19(1):17–37.