HHS Public Access

Author manuscript

Ear Hear. Author manuscript; available in PMC 2019 September 01.

Published in final edited form as:

Ear Hear. 2018 ; 39(5): 1035–1038. doi:10.1097/AUD.0000000000000561.

A Retrospective Estimate of Ear Disease Detection Using the "Red Flags" in a Clinical Sample

Niall A.M. Klyn[1], **Samantha Kleindienst Robler**[2,3], **Razan Alfakir**[4], **Donald W. Nielsen**[5], **James W. Griffith**[6], **Deborah L. Carlson**[7], **Larry Lundy**[4], **Sumitrajit Dhar**[1,8], and **David A. Zapala**[4]

[1]Roxelyn and Richard Pepper Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL, USA

[2]Division of Audiology, Mayo Clinic, Scottsdale, AZ, USA

[3]Department of Audiology, Norton Sound Health Corporation, Nome, AK, USA

[4]Department of Otorhinolaryngology, Mayo Clinic, Jacksonville, FL, USA

[5]Don Nielsen Consulting, LLC, Dublin, OH, USA

[6]Department of Medical Social Science, Northwestern University, Chicago, IL, USA

[7]Department of Otolaryngology, University of Texas Medical Branch, Galveston, TX, USA

[8]The Hugh Knowles Center, Northwestern University, Evanston, IL, USA

Abstract

Objectives—The purpose of this study was to evaluate the specificity and sensitivity of two red flag protocols in detecting ear diseases associated with changes in hearing.

Design—The presence of red flag symptoms was determined in a chart review of 307 adult patients from the Mayo Clinic Florida departments of Otorhinolaryngology and Audiology. Participants formed a convenience sample recruited for a separate study. Neurotologist diagnosis was the criterion for comparisons.

Results—Of the 251 patient files retained for analysis, 191 had one or more targeted diseases and 60 had age- or noise-related hearing loss. Food and Drug Administration (FDA) red flags sensitivity was 91% (CI 86–95%) and specificity was 72% (CI 59–83%). American Academy of Otolaryngology (AAO-HNS) red flags sensitivity was 98% (CI 95–99%) and specificity was 20% (CI 11–32%).

Conclusions—Stakeholders must determine which diseases are meaningful contraindications for hearing aid use, and whether these red-flag protocols have acceptable levels of sensitivity and

Corresponding author: Niall Klyn, Roxelyn and Richard Pepper Department of Communication Sciences and Disorders, Northwestern University, 2240 Campus Dr, Evanston, IL 60208, USA. niall.klyn@northwestern.edu.

Financial Disclosures/Conflicts of Interest:

Research was supported by NIH/NIDCD (R21/33 DC013115) [Dhar, Zapala] and the Knowles Hearing Center at Northwestern University [Dhar, Zapala, Nielsen]

specificity. As direct-to-consumer models of hearing devices increase, a disease detection method that does not require provider intercession would be useful.

INTRODUCTION

The FDA recently announced they would no longer enforce the mandatory medical evaluation or waiver protocols for individuals seeking hearing aids (FDA 2016). This policy change leaves a gap in regulatory guidance for audiologists and hearing aid dispensers. Medical evaluation is potentially important because changes in an individual's hearing could be a symptom of a disease associated with risks to an individual's life or health (Kleindienst et al. 2016). Failing to detect such diseases can have significant ramifications for the individual and the healthcare system. For many diseases with symptomatic changes in hearing, timely treatment can lead to improved patient outcomes and a reduced burden on the healthcare system. The potential seriousness of undetected diseases was part of the stated rationale of the earlier FDA surveillance scheme. As part of that scheme, the FDA set forth seven "red flags" that each warrant a full medical evaluation prior to procurement and use of hearing aids (21C.F.R.801.421 2016). In addition, the American Academy of Otolaryngology-Head and Neck Surgery (AAO-HNS) proposed its own set of red flags to provide guidance on when a patient should be evaluated by a physician (AAO-HNS 2014).

With the recent change in FDA policy, the onus is entirely on audiologists and hearing aid dispensers to dutifully screen and appropriately triage individuals suspected of having insidious ear disease. The FDA and AAO-HNS red flag criteria, therefore, may continue to serve as important tools ensuring consumer safety in hearing healthcare. Though the current iteration of the FDA red flags has been in use for over 20 years, we could not find published, peer-reviewed estimates of their effectiveness in screening for diseases related to changes in hearing. This brief report evaluates the FDA and AAO-HNS red flags in a sample of patients with disease-related etiologies and those with age- or noise-related changes in hearing as determined by a board-certified neurotologist (American Board of Otolaryngology 2017). This sample allowed for estimation of the sensitivity and specificity of the screening tools. The stated purpose of both sets of red flags is to suggest medical evaluation when warranted (21C.F.R.801.421 2016, AAO-HNS 2014). This brief report will focus on the red flags' sensitivity and specificity in detecting disease, and does not comment on whether a hearing aid would be an appropriate option for the study sample. As far as is possible with the large set of diseases and conditions considered by the red flags, this report follows the STARD protocol in reporting its results (Bossuyt et al. 2015)

MATERIALS AND METHODS

Subjects

Patients (n = 307, 48% female) were recruited from the Otorhinolaryngology and Audiology Departments of the Mayo Clinic Florida between June, 2014 and August, 2015 as part of development of the Consumer Ear Disease Risk Assessment (CEDRA, Kleindienst et al. 2017) questionnaire. Participants were between 40–80 years of age (mean = 62.9, SD = 9.8), and were seeking care for ear- or hearing-related complaints. This convenience sample was chosen to have a high percentage of targeted disease-related diagnoses, and was limited to

25% of the patients having age- or noise-related changes in hearing. An a priori power analysis indicated an initial sample size of 300 was sufficient for the Kleindienst et al. study, with 80% power to detect an area under the curve of 0.63 or greater with alpha = .05. No separate power analysis was performed for the current analysis. Participants received treatment according to the standard-of-care at Mayo Clinic, and the study was approved by the Mayo Clinic IRB.

Procedures

As part of the standard-of-care at Mayo Clinic, audiology examination, case history, and background data were entered into the electronic health record (EHR), which were later extracted for this study's analysis. All study participants received an examination by a board-certified neurotologist (American Board of Otolaryngology 2017) whose findings were used to determine the presence/absence of any of 104 diseases that may include otologic symptoms (see Text, Supplemental Digital Content, for a complete list; Kleindienst et al. 2016). All examinations were performed according to the practice standards used by Mayo Clinic during the study. The FDA and AAO-HNS red flags were extracted from the audiometric data, background notes, and case history, all of which were accessible to the neurotologist. Only examination information collected equal-to or less than 90 days before or after the neurotologic exam was included in the analysis because these diseases have varying time courses. This resulted in 251 patients for the present analysis.

Two independent coders (NK and DZ) analyzed the de-identified dataset to determine the presence or absence of the red flag symptoms. Both coders had access to the neurotologist diagnosis, but completed their analysis without reference to it. Initial agreement between the two coders was 91%. Examination of the discordant classifications showed that disagreements resulted from different interpretations of the red flags. The study team discussed possible interpretations and decided on a single uniform interpretation for this study. The red flags and the study team's operationalization of each may be found in Table 1. The results here reflect complete final agreement between the two coders.

Data Analysis

For each of the FDA and AAO-HNS red flag sets, a patient was coded as a "refer" in the presence of any red flag sign or symptom, and as a "pass" only when no red flags were present. Sensitivity was calculated as the number of correctly identified diseased cases divided by the total number of diseased cases. Specificity was calculated as the number of correctly identified non-diseased cases divided by the total number of non-diseased cases. The sensitivity and specificity of the screening tools are reported using the neurotologist's diagnosis as the criterion, with the exact binomial 95% confidence intervals in parentheses.

RESULTS

The sensitivity and specificity for the FDA red flags was 91% (CI 86–95%) and 72% (CI 59–83%), respectively. The sensitivity and specificity for the AAO-HNS red flags was 98% (CI 95–99%) and 20% (CI 11–32%), respectively. See Table 2 for cross tables. We could not find peer-reviewed estimates of the cumulative prevalence of these diseases, so the positive

predictive value (PPV) and negative predictive value (NPV) were estimated using three prevalence levels: 0.5%, 2%, and 5%. The highest prevalence rate chosen is close to the highest described by the recent National Academies of Science, Engineering, and Medicine report (Committee on Accessible and Affordable Hearing Health Care for Adults, 2016), which reflects the most common diseases that may affect hearing. The PPVs and NPVs are summarized in Table 3.

DISCUSSION

The data presented here provide the first published estimates of the sensitivity and specificity of the FDA and AAO-HNS red flags in screening for any of the 104 ear diseases described in Kleindienst et al. (2016). Sensitivity and specificity are characteristics of the tests themselves, and do not vary with the prevalence of the disease. The PPV and NPV estimates provide some context for these characteristics by factoring in how rare or common the diseases are in the general population, and are provided for illustrative purposes only, as the true cumulative prevalence of ear diseases is unknown. These results highlight the relatively low PPV of these tests: for the AAO-HNS red flags in the highest prevalence case, only 6% of individuals referred for further diagnosis have a disease. For the FDA red flags in the highest prevalence case, that number rose to 14.6%. Conversely, under reasonable prevalence estimates, passing either set of red flags indicated that the individual likely has no disease, with >99% passing the test being disease-free.

The present study highlights two potential concerns in implementing the red flags. First, an individual seeking a hearing aid must see a provider (e.g. audiologist or hearing instrument specialist) to obtain the required audiometric data included in both the FDA and AAO-HNS red flags. Though this step potentially avoids individuals unnecessarily seeing a physician for non-medically treatable conditions, it still requires the outlay of time and money for the provider assessment. With increasing direct-to-consumer sales of hearing devices, a disease detection method with acceptable levels of sensitivity and specificity that does not require provider participation could be more efficient. Kleindienst et al. (2017) developed a consumer questionnaire that requires no examination and possesses test characteristics comparable to the FDA red flags, and with better specificity than the AAO-HNS red flags.

The second potential concern in using the red flags is that some items in both the FDA and AAO-HNS red flag descriptions are ambiguously defined. For example, does the FDA red flag “audiometric air-bone gap equal to or greater than 15 decibels at 500, 1,000, and 2,000 Hz” apply to gaps at all three frequencies, any frequency, or an average of the three? In the initial coding, one author coded a patient reporting “pressure” or “fullness” as a flag for both the FDA and AAO-HNS, and the other author did not. For this project the final coding was determined by consensus among the research team, but the ambiguous definitions leave open the possibility of variance in how these red flags are used. Indeed, in preliminary analyses (not shown), these different interpretations were found to change the sensitivity and specificity of the red flags. This could cause individuals to not be referred when they have a disease or non-diseased individuals to pursue unnecessary care, simply depending on the dispenser’s or audiologist’s interpretation of the red flags.

This study has limitations. First, we studied a clinical sample (n = 251), that does not represent the population prevalence of ear disease. The high prevalence of disease, however, provided a more rigorous test of the red flag's sensitivity than would have been possible in a random sample of the hearing aid seeking population. The concern is therefore that the diseased and control individuals in this study may be systematically different than the general population. An ongoing research project is evaluating the red flag criteria and CEDRA in a larger, more representative population of individuals seeking hearing healthcare. A second limitation is that the red flag symptoms were derived from a retrospective analysis of case history and examination data. The symptoms noted by the clinicians therefore had to be interpreted to determine if they matched any red flag. The high agreement between coders indicates that the possibility of the interpretative process skewing the results is small, but that different interpretations of the red flags themselves affect the outcomes. The ongoing research project discussed above is explicitly examining the red flag criteria as interpreted by practicing audiologists. The 104 diseases targeted here have not been independently evaluated as meaningful contraindications of hearing device acquisition and use. The ratings provided by neurotologists in Kleindienst et al. (2016) covered a range of severity should the disease go undetected. Here, however, the red flags were tested for their effectiveness at determining the presence or absence of a disease, not the potential consequence of missed detection. Whether an otolaryngologist would clear a patient for hearing aid use, regardless of the presence of a disease or condition other than presbycusis, is part of the ongoing research described above. Finally, the authors note that neurotologist diagnosis is not a strictly defined reference test as set forth in the STARD protocol (Bossuyt et al., 2015). Board certification in neurotology represents the highest medical subspecialty credential for assessing ear disease, and provides the best available standardization for any effort to replicate these results.

For any screening effort, the relative costs and benefits of the procedure must be weighed. Whether screening for these diseases is appropriate and necessary, and the acceptable sensitivity and specificity for a screener, are questions that must be considered by the public and policymakers. The results presented here can help inform such decision-making.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

All authors contributed equally to this work. N.K. and D.Z. provided analyses. S.K. reviewed analyses and provided critical interpretation. D.Z., L.L., and S.K. contributed to data collection. S.K., R.A. and D.Z. pulled and de-identified data for analysis. N.K. wrote the initial draft of this work. All authors reviewed, edited, and approved the final paper.

References

21 C.F.R. 801.421 (2016).

American Academy of Otolaryngology-Head and Neck Surgery. [Accessed January 7, 2017] Position Statement: Red Flags Warning of Ear Disease. 2014. http://www.entnet.org/?q=node/912

American Board of Otolarynology. Neurotology Subcertification Exam. 2017. Retrieved from https://www.aboto.org/neuro_certification.html

Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, … Kressel HY. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. Radiology. 2015; 277(3):826–832. [PubMed: 26509226]

Committee on Accessible and Affordable Hearing Health Care for Adults; Board on Health Sciences Policy; Health and Medicine Division; National Academies of Sciences, Engineering, and Medicine. Hearing Health Care for Adults: Priorities for Improving Access and Affordability. Washington, DC: National Academies Press; 2016.

Food and Drug Administration, Center for Devices and Radiological Health. [Accessed April 6, 2017] Immediately in Effect Guidance Document: Conditions for Sale for Air- Conduction Hearing Aids. Guidance for Industry and Food and Drug Administration Staff. 2016. https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm531995.pdf

Kleindienst SJ, Dhar S, Nielsen DW, et al. Identifying and Prioritizing Diseases Important for Detection in Adult Hearing Health Care. American Journal of Audiology. 2016; 25(3):224.doi: 10.1044/2016_AJA-15-0079 [PubMed: 27679840]

Kleindienst SJ, Zapala DA, Nielsen DW, et al. Development and initial validation of a consumer questionnaire to predict the presence of ear disease. JAMA Otolaryngol Head Neck Surg. 2017

Table 1

Operationalization of the Red Flags for the Current Study

FDA Red Flags	**Key criteria used in chart review**
Visible congenital or traumatic deformity of the ear.	Search for “deformity,” or “malformation,” of the “pinna”, or “external ear”.
History of active drainage from the ear within the previous 90 days.	Search for “drain,” “pus,” or “otorrhea.”
History of sudden or rapidly progressive hearing loss within the previous 90 days.	Search for “sudden,” “rapid,” “acute,” and “hearing loss” within 90 days.
Acute or chronic dizziness.	Search for “dizziness,” or “vertigo.”
Unilateral hearing loss of sudden or recent onset within the previous 90 days.	Search for “unilateral” or “asymmetric” within 90 days.
Audiometric air-bone gap equal to or greater than 15 dB at 500 Hz, 1000 Hz, and 2000 Hz. *	Calculate air-bone gap at 500 Hz, 1000 Hz, and 2000 Hz, then calculate the average.
Visible evidence of significant cerumen accumulation or a foreign body in the ear canal.	Search for “blockage,” “foreign,” “debris,” “object,” “cerumen obstructing clear view of the tympanic membrane,” or “wax obstructing clear view of the tympanic membrane”, “impaction”
Pain or discomfort in the ear. *	Search for “pain,” “discomfort,” “fullness,” “pressure,” or “otalgia.”
AAO-HNS Red Flags	
Hearing loss with a positive history of ear infections, noise exposure, familial hearing loss, TB, syphilis, HIV, Meniere’s disease, autoimmune disorder, ototoxic medication use, otosclerosis, von Recklinghausen’s neurofibromatosis, Paget’s disease of bone, ear or head trauma related to onset.	Hearing loss as $PTA_{.5,\ 1,\ 2\ kHz}$ > 20 dB HL, and search for “noise,” “familial/family,” “TB/tuberculosis,” “syphilis,” “HIV/AIDS,” “Meniere’s,” “autoimmune,” “ototoxic,” “otosclerosis,” “Recklinghausen,” “neurofibromatosis,” “Paget,” or “trauma”
History of pain, active drainage, or bleeding from an ear. *	Search for “pain,” “discomfort,” “pressure,” “fullness,” “otalgia,” “drain/drainage,” “blood,” “bleeding,” “pus,” or “otorrhea.”
Sudden onset or rapidly progressive hearing loss.	Search for “sudden,” “rapid,” or “acute.”
Acute, chronic, or recurrent episodes of dizziness.	Search for “dizziness” or “vertigo.”
Evidence of congenital or traumatic deformity of the ear.	Search for “deformity,” or “malformation,” of the “pinna”, or “external ear”.
Visualization of blood, pus, cerumen plug, foreign body, or other material in the ear canal.	Search for “blood,” “bleeding,” “pus,” “blockage,” “foreign,” “debris,” “object,” “cerumen,” or “wax.”
An unexplained conductive hearing loss or abnormal tympanogram.	Tympanograms with pressure < −100 daPa, or classified as type “B” or “C.”
Unilateral or asymmetric hearing loss (a difference of greater than 15 dB Pure Tone Average between ears); or bilateral hearing loss > 30 dB. *	Calculated $PTA_{.5,1,2\ kHz}$ for each ear and subtracted. For bilateral, flagged if both ears’ PTA > 30 dB HL.
Unilateral or pulsatile tinnitus.	Search for “tinnitus” with “pulsatile” or “unilateral.”
Unilateral or asymmetrically poor speech discrimination scores (a difference of greater than 15% between ears); or bilateral speech discrimination scores <80%.	Examine scores from NU-6 for each ear and subtract. For bilateral, flagged if both ears’ speech scores were < 80%.

* Alternate operationalizations were attempted for these red flags, which resulted in moderate changes to the sensitivity and specificity. The authors finalized the analysis using the descriptions listed in the table.

Table 2

Cross Table of Red Flags and Reference for Disease Presence

FDA	Diseased	Age- or noise-related loss	Totals	AAO-HNS	Diseased	Age- or noise-related loss	Totals
Flagged	174	17	191	Flagged	187	48	235
Passed	17	43	60	Passed	4	12	16
Totals	191	60	251	Totals	191	60	251

Of the 191 individuals with targeted diseases, 42 had multiple diagnoses. 118 diagnoses were of sensorineural or labyrinthine disorders, 14 were related to the posterior fossa or central auditory nervous system, 79 were conductive, middle ear, or mastoid-related, and 22 were eighth nerve-related. See Kleindienst et al. (2016) for details and categorizations of the 104 targeted diseases.

Table 3

Estimated Positive Predictive Value (PPV) and Negative Predictive Value (NPV) at Three Prevalence Levels

	0.5% prevalence		2% prevalence		5% prevalence	
	PPV	**NPV**	**PPV**	**NPV**	**PPV**	**NPV**
FDA	1.6%	99.9%	6.2%	99.7%	14.6%	99.3%
AAO-HNS	0.6%	99.9%	2.4%	99.8%	6.1%	99.5%