

VALIDITÉ DES EXAMENS AUDIOMÉTRIQUES DE DÉPISTAGE: UTOPIE OU RÉALITÉ?

Josée Gauthier

Direction de la santé publique, Hôpital Maisonneuve-Rosemont
Santé au travail et environnementale, Équipe Bruit et Audition¹
75 rue de Port-Royal est, Montréal, Québec, H3L 3T1

SOMMAIRE

Une analyse de la validité concomitante des examens audiométriques de dépistage a été réalisée en procédant à une *analyse de variance multivariée sur mesures répétées* (MANOVA) des seuils audiométriques mesurés en clinique et en dépistage sur un même groupe de travailleurs industriels. D'autres types de comparaisons plus qualitatives ont aussi été effectués. Globalement, les seuils mesurés en clinique et en dépistage correspondent étroitement. Cependant, l'analyse de variance a mis en évidence des interactions *Test * Fréquence* et *Test * Oreille* significatives. Ces interactions sont essentiellement liées aux dispositifs d'examens (étalonnage et procédure automatisée) et ne sont pas de nature à remettre en cause la validité de l'examen. Toutefois, ces résultats témoignent de la fragilité des conditions de validité et mettent en exergue la nécessité du contrôle rigoureux et régulier de l'ensemble des sources potentielles d'erreurs. À cet égard, la procédure d'analyse utilisée dans cette étude peut constituer un outil de supervision indirecte pour les audiologistes.

ABSTRACT

An analysis of the validity of audiometric screening tests has been performed by a *repeated measures analysis of variance* (MANOVA) on hearing thresholds measured in a screening context, and the corresponding thresholds obtained by clinical tests among the same industrial workers. The results support the feasibility of valid audiometric tests in the context of screening for noise-induced hearing loss. However, they also demonstrate the importance of strict control of all potential sources of variation to obtain valid thresholds. The method used in this study is easy to set up and particularly sensitive to systematic sources of variation; therefore it could be integrated into the procedures applied by audiologists supervising auditory screening.

1. INTRODUCTION

Au Québec, on effectue chaque année des milliers d'examens audiométriques de dépistage auprès de travailleurs exposés à des niveaux de bruit nocifs pour leur audition. La validité et la

fiabilité de tels examens sont primordiales car l'information en découlant sert diverses fins dont, entre autres, informer l'individu de son état de santé auditive en relation avec son milieu de travail, documenter la pertinence de procéder à une demande d'indemnisation pour surdit  professionnelle, motiver

¹ Nouvelle dénomination pour le Programme Régional-audiologie consécutive à la récente restructuration des programmes de santé publique dans la région de Montréal

la réduction de l'exposition au bruit, etc. On peut donc facilement concevoir qu'un certain laxisme au niveau de la qualité des examens peut devenir préjudiciable à la fois pour les travailleurs et pour la crédibilité des interventions.

L'examen audiométrique tonal aérien est une procédure particulièrement vulnérable à l'influence de certaines sources d'erreurs systématiques ou aléatoires qui commandent un contrôle rigoureux des conditions dans lesquelles il se déroule.

Ces sources d'erreurs s'établissent globalement comme suit:

- Étalonnage de l'audiomètre;
- Niveau de bruit ambiant;
- Positionnement des écouteurs et tension de l'arceau;
- Procédure de recherche et de détermination des seuils;
- Intervention de l'examineur;
- Familiarisation du sujet avec la tâche;
- État du sujet au moment de l'examen.

Sans entrer dans le détail, mentionnons que l'ensemble des conditions minimales à respecter à l'égard de ces différentes sources d'erreurs ont fait l'objet de normes internationales sur lesquelles s'est aligné le Bureau de normalisation du Québec afin d'établir ses propres normes pour les examens de dépistage en milieu de travail [1]. C'est ainsi que l'on assume généralement la validité des examens audiométriques lorsque l'on peut établir le respect intégral des normes relatives au dispositif d'examens et à la procédure de recherche de seuils utilisés de même qu'à l'état du sujet au moment de l'examen.

Dans un contexte de dépistage, il est facile d'établir la conformité aux normes pour ce qui est du dispositif d'examens comme tel. Il l'est cependant beaucoup moins pour tout ce qui touche les aspects sur lesquels l'examineur exerce son propre contrôle (procédure d'examens, respect des prérequis, etc.). La réalisation des examens étant généralement confiée à des tiers alors que la responsabilité de la qualité et de la validité des examens réalisés par un service donné est assumée par un audiologiste, il importe de développer des outils qui lui permettent de vérifier la validité des examens. Comme la supervision directe constitue une procédure peu efficiente dans un contexte de dépistage, des modalités de contrôle indirect doivent être mises en place.

C'est dans cette perspective que l'équipe d'audiologistes du Programme Régional-Audiologie (PRA), centre spécialisé en

bruit et audition et dispensant des services audiologiques cliniques et de dépistage à des travailleurs industriels, a procédé à l'étude rapportée ici.

Plus spécifiquement, les objectifs de l'étude étaient 1 - de vérifier la validité des examens audiométriques de dépistage effectués par le PRA et 2 - de développer un outil de supervision indirecte de la réalisation d'examens audiométriques par des tiers.

2. MÉTHODOLOGIE

On peut considérer que, de façon générale, les seuils audiométriques mesurés en clinique audiolinguistique sont valides et fiables et ce, pour diverses raisons. En effet, le contrôle optimal des diverses sources d'erreurs y est rigoureusement observé et les examens audiométriques sont effectués dans le cadre d'une évaluation complète dont les différentes épreuves doivent concorder entre elles. De plus, l'audiologiste qui interprète et signe les résultats est soumis à un code de déontologie qui le responsabilise de sa pratique. L'ensemble de ces facteurs permet l'utilisation des seuils mesurés en clinique spécialisée comme étalon pour évaluer la validité des mesures effectuées en dépistage.

On établit ainsi ce qu'il est convenu d'appeler la validité concomitante [2] des seuils audiométriques de dépistage. Cette façon de procéder a comme principale vertu de permettre une évaluation *in situ* de la validité car les examens sont effectués de la manière habituelle sans souci de performance à l'égard d'une procédure expérimentale.

2.1 Base de données

Le PRA assumait le suivi clinique d'une certaine proportion des travailleurs évalués par son service de dépistage. On dispose ainsi d'une banque de données facilement accessible.

Le PRA assurait un contrôle minutieux et un respect rigoureux des normes de pratique [1, 3-5] en ce qui a trait aux examens auditifs effectués par ses services et ce, tant en dépistage qu'en clinique. Ceci sous-tend entre autres, un calendrier d'étalonnage complet des appareils, un contrôle subjectif régulier de même que le respect strict des prérequis aux examens. De plus, les infirmières effectuant depuis plus de dix (10) ans les examens de dépistage au PRA ont reçu pour ce faire une formation complète, régulièrement mise à jour, et

disposaient à tout moment du support d'une équipe d'audiologistes.

Afin de caractériser au mieux la fonction auditive des individus, les examens auditifs de dépistage effectués au PRA comportaient, en plus de l'examen audiométrique tonal aérien, un examen visuel du conduit auditif externe et un examen tympanométrique. Le premier a pour objet d'identifier la présence éventuelle d'obstacle à la transmission de l'énergie sonore à la chaîne tympano-ossiculaire alors que le second vient compléter l'information fournie par l'examen audiométrique en établissant l'état actuel de ce mécanisme conductif. La mesure des seuils audiométriques était effectuée pour la gamme de fréquences de 500 Hz à 6 000 Hz avec un audiomètre à enregistrement automatisé (de type "Békésy") afin de limiter au mieux le biais de l'examineur [6,7].

2.2 Sélection des dossiers

Pour colliger l'ensemble des données pertinentes, une recherche a été entreprise dans les registres administratifs afin d'identifier les dossiers de travailleurs ayant été vus au PRA tant en clinique qu'en dépistage ($n = 69$). Nous n'avons conservé que les dossiers dont l'intervalle entre les examens cliniques et de dépistage était inférieur à 2 ans (7 exclusions). Ce critère de sélection nous semblait légitime car le temps écoulé entre les examens était suffisamment long pour éviter un éventuel effet d'apprentissage et suffisamment court pour limiter la dégradation de l'audition liée à l'exposition au bruit et au processus de vieillissement. Nous avons par ailleurs écarté tous dossiers pour lesquels nous ne disposions pas de la totalité des informations (ex.: seuils non-mesurés, résultats tympanométriques absents, etc.) (6 exclusions). Nous avons de plus retranché les dossiers dont les résultats tympanométriques suggéraient un potentiel de variation de la mesure des seuils audiométriques (5 exclusions) (amplitude du tracé non-mesurable, pression de l'oreille moyenne non-mesurable ou inférieure à - 100 daPa) [8]. C'est ainsi qu'un échantillon de cinquante et un (51) dossiers a été constitué pour fins d'analyse.

L'âge moyen des travailleurs retenus est de 49,4 ans (E.T. = 10,8 ans). L'âge moyen relativement élevé de cet échantillon est explicable par le fait que la population de travailleurs vus à la clinique du PRA, l'était dans un contexte d'indemnisation pour surdité professionnelle. En ce qui a trait à l'intervalle entre les examens, il est en moyenne de 0,44 an (E.T. = 0,25 an), l'examen de dépistage étant toujours antérieur

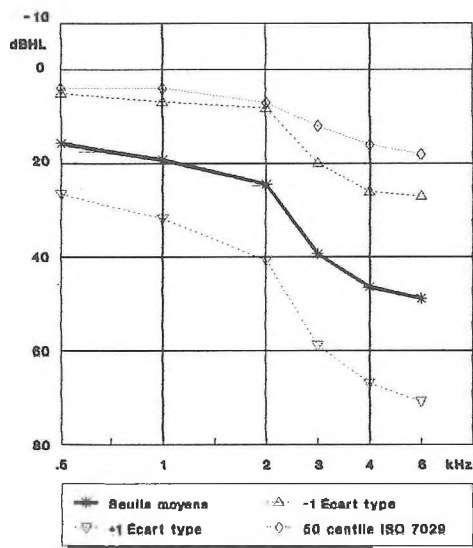
à l'examen clinique. Les valeurs minimale et maximale observées à ce niveau sont respectivement 0,08 an et 1,08 an.

2.3 Procédure

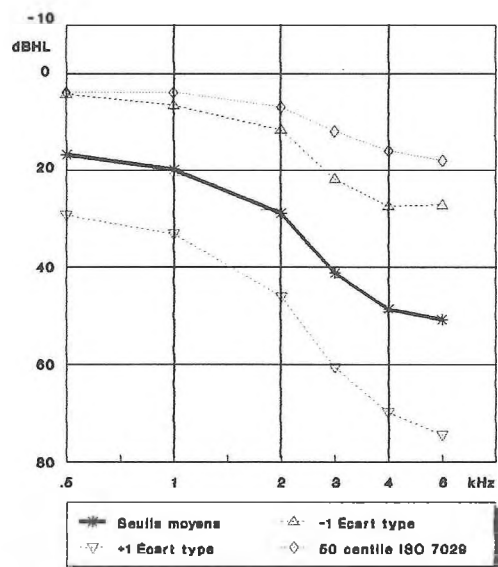
La majorité des études consultées au moment d'établir notre stratégie d'analyse proposait une procédure consistant en une comparaison des statistiques descriptives des différences observées (moyennes et écarts-types) [9-17], sans égard à la latéralité (oreilles droite et gauche confondues) [9-11,13-17], basée sur les corrélations de Pearson [12,14,15,17,18] et sur les résultats de tests bilatéraux de t de Student [12,15,17].

Un des paramètres importants de la variabilité audiométrique est la marge de manoeuvre permise par les normes en ce qui a trait aux niveaux de sortie et à la précision en fréquences [19] des audiomètres utilisés pour chacun des examens audiométriques comparés. Les niveaux de sortie des audiomètres peuvent effectivement varier, entre écouteurs et entre audiomètres, à l'intérieur de ± 3 dB de 500 à 4 000 Hz et de ± 5 dB à 6 000 Hz et ce, tant en clinique qu'en dépistage [1,3]. Puisque la majorité des études portaient sur la fiabilité des examens audiométriques effectués sur un même appareil et de façon très rapprochée dans le temps, ce paramètre s'avérait relativement bien contrôlé. Ceci permet d'expliquer pourquoi les auteurs ne jugeaient pas utile d'effectuer des analyses en séparant les deux oreilles : il n'y avait alors aucune raison de croire que la variabilité audiométrique soit différentielle suivant ce facteur. Cependant, pour les études où les examens comparés étaient effectués sur des audiomètres différents ou que l'intervalle de temps les séparant était relativement long, trop peu d'études [9,10,16] ont considéré les différences potentielles entre les niveaux de sortie des audiomètres par fréquence et par écouteurs. Il est donc légitime de croire que les résultats de certaines études aient pu être entachés par des biais systématiques non-identifiés.

Aussi, pour établir la validité concomitante de l'examen audiométrique de dépistage, nous avons d'abord procédé à une analyse de variance multivariée (MANOVA) sur mesures répétées [20] avec trois critères de classification intra-sujets soient la fréquence (500 Hz, 1 000 Hz, 2 000 Hz, 3 000 Hz, 4 000 Hz et 6 000 Hz) l'oreille testée (droite vs gauche) et le type d'examen (clinique vs dépistage). Nous avons de plus apprécié la distribution relative, par fréquence, des écarts individuels observés en fonction de la distribution attendue en vertu de l'erreur audiométrique. Enfin, nous avons répliqué la



Oreille droite (n = 51)



Oreille gauche (n=51)

FIGURE 1: Moyennes et écarts-types des seuils audiométriques mesurés en clinique

procédure suggérée par les différentes études consultées afin d'en apprécier les éventuelles conséquences au niveau de l'interprétation.

L'ensemble des résultats ainsi obtenus est présenté et analysé ci-après.

3. RÉSULTATS ET ANALYSE

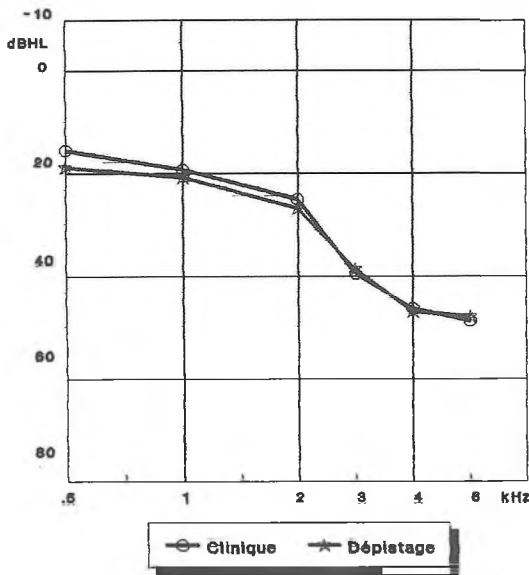
Comme en témoigne la figure 1, les seuils audiométriques moyens mesurés en clinique montrent des pertes d'audition bilatérales et symétriques relativement importantes pour l'ensemble des travailleurs de l'échantillon. Les valeurs moyennes de seuils observés sont beaucoup plus importantes que les valeurs du 50^e centile de la distribution de l'effet de l'âge sur l'audition établie selon la norme ISO 7029 pour un groupe d'hommes de 50 ans [21]. L'ampleur des pertes mesurées et la configuration audiométrique observée sont caractéristiques de la population effectivement rejointe par le service clinique du PRA. Les écarts types relativement élevés qui augmentent de façon concomitante avec les fréquences témoignent vraisemblablement des différences interindividuelles aux niveaux de l'exposition au bruit et des effets du bruit sur l'audition combinées à un éventuel effet

d'âge, lui même variable d'un individu à l'autre.

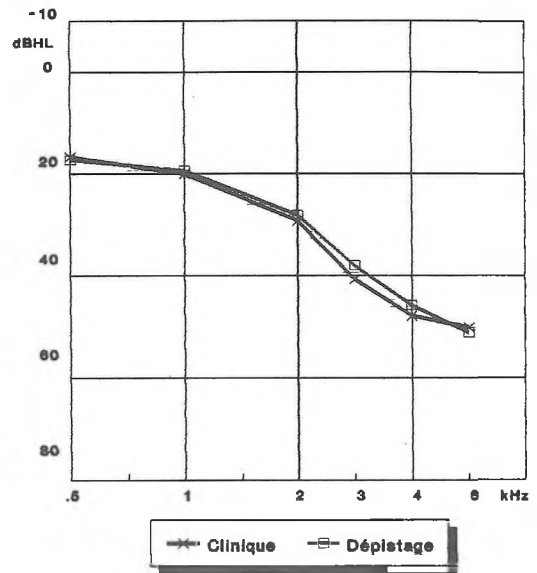
La figure 2 établit visuellement la comparaison des seuils mesurés en clinique avec ceux obtenus en dépistage et ce, pour chacune des oreilles. Malgré l'étroite correspondance des valeurs moyennes observées, on note un écart entre les examens clinique et dépistage qui semble différentiel en fréquences et entre oreilles. En effet, des écarts entre les examens clinique et dépistage sont visuellement appréciables à l'oreille droite surtout à 500 Hz mais aussi, bien que dans une moindre mesure, à 1000 Hz et 2 000 Hz. À l'oreille gauche, ce sont à 3 000 Hz et à 4 000 Hz que de tels écarts sont notés. L'analyse de variance effectuée reflète objectivement ces observations.

3.1 Analyse de variance (MANOVA)

Les différences observées entre les seuils cliniques et les seuils de dépistage dépendent effectivement tant de l'oreille considérée (interaction *Test * Oreille* : $F(50,1) = 3.63$; $p = 0.063$) que de la fréquence observée (interaction *Test * Fréquence* : $F(250,1) = 4.81$; $p = .000$). Aucune interaction significative n'est cependant notée entre *Oreille et Fréquence* ($F(250,5) = 1.38$; $p = 2.32$) ce qui soutient que le facteur *Test* est le premier déterminant des interactions



Oreille droite n=51 oreilles



Oreille gauche n=51 oreilles

FIGURE 2: Comparaison oreille par oreille des seuils moyens mesurés en clinique avec les seuils moyens mesurés en dépistage.

TABLEAU 1 Comparaison Clinique - Dépistage, analyse oreille par oreille

Fréquences	Oreille droite (n = 51)			Oreille gauche (n = 51)		
	Différence ¹ (dBHL)		Corrélation de Pearson (p<.001)	Différence ¹ (dBHL)		Corrélation de Pearson (p<.001)
	Moyenne	Écart type		Moyenne	Écart type	
500	-3,35	5,99	.846	-0,59	7,07	.838
1000	-1,49	7,51	.827	0,63	6,30	.881
2000	-1,78	6,56	.934	1,12	6,31	.939
3000	0,86	6,72	.942	2,69	5,71	.961
4000	-0,61	7,45	.935	2,02	6,38	.957
6000	0,82	9,84	.899	-1,02	9,23	.920

1 (Seuil clinique = Seuil dépistage)

observées.

La présence d'interactions significatives nous empêche de juger de la validité de l'examen de dépistage à partir des résultats de l'analyse de variance quoique les résultats n'établissent pas de différence statistiquement significative entre l'examen de

dépistage et l'examen clinique au niveau des seuils mesurés (Test : F(50,1) = 0.39 ; p = .536). Cependant, une analyse subjective plus détaillée des différences observées permet d'établir le poids relatif des différentes fréquences, en fonction de la latéralité, dans les interactions mises en évidence. On peut ainsi déduire une certaine appréciation de la validité de

l'examen et abstraire les hypothèses explicatives les plus plausibles des interactions identifiées.

À partir des informations résumées au tableau 1, on constate que les différences moyennes entre les résultats obtenus en clinique et ceux de dépistage varient entre moins de 1 dB et un peu plus de 3 dB et que, globalement, les mesures de seuils à l'oreille droite vont dans le sens de meilleurs seuils en clinique qu'en dépistage tandis que l'effet inverse est noté à l'oreille gauche. Les écarts visualisés à la figure 2 sont évidemment supportés par les valeurs moyennes correspondantes, les plus importantes survenant à 500 Hz à l'oreille droite (- 3,35 dB) ainsi qu'à 3 000 Hz (+2,69 dB) et 4 000 Hz (+2,02 dB) à l'oreille gauche. Par ailleurs, les moyennes et écarts types des différences se comparent avantageusement à ce que l'on peut retrouver dans la littérature [10, 11, 17, 22]. Les coefficients de corrélation de Pearson, tous supérieurs à 0,83 et statistiquement significatifs, montrent qu'en dépit des différences observées plus haut, les deux types d'examens comparés mesurent essentiellement la même chose ce qui soutient la validité de l'examen de dépistage.

Compte tenu d'effets différents selon les oreilles et les fréquences, il apparaît clairement que ce ne sont pas des phénomènes physiopathologiques qui sont à la base des différences observées. En effet, d'une part, en ce qui concerne les résultats obtenus à 500 Hz, une différence significative de sensibilité à cette fréquence particulière est difficilement explicable puisqu'un contrôle serré d'éventuels problèmes de conduction a été effectué au moment de la sélection des dossiers. D'autre part, bien que les fréquences de 3 000 Hz et 4 000 Hz soient parmi les plus sensibles aux effets cumulatifs de l'âge et du bruit et que le sens des différences observées soit compatible avec une telle détérioration (seuils pires en clinique qu'en dépistage), il demeure impensable que le bruit et l'âge aient eu un effet systématiquement supérieur à une oreille en particulier. Force est donc de constater que les différences observées témoignent d'erreurs systématiques liées à la mesure des seuils tant en dépistage qu'en clinique. Si l'on entend favoriser la validité des examens audiométriques réalisés par nos services, il est impératif d'en identifier les sources de façon à apporter les correctifs qui s'imposent.

Nous avons écarté la possibilité de l'effet de familiarisation en cours de test pour expliquer les différences observées à 500 Hz parce que la première oreille testée par nos appareils automatisés est toujours l'oreille gauche. Nous avons trouvé le

phénomène explicatif recherché en nous attardant au fonctionnement automatisé de nos audiomètres. En effet, une fois la dernière fréquence testée à l'oreille gauche, l'appareil procède automatiquement au changement d'oreille et débute la mesure du seuil à 500 Hz à l'oreille droite à un niveau supra-liminaire. Ce changement subit dans le niveau de présentation du signal peut nécessiter un certain temps d'ajustement de la réponse du sujet examiné pour en ramener la précision à un niveau liminaire. Le temps consenti par l'appareil pour la mesure des seuils à chacune des fréquences, à savoir 30 secondes, ne suffit vraisemblablement pas pour couvrir le temps d'ajustement nécessaire à une certaine proportion d'individus testés. Ce phénomène entraîne une erreur systématique de surestimation du seuil à cette fréquence et à cette oreille et pourrait même expliquer les écarts obtenus à 1 000 Hz et 2 000 Hz à la même oreille. Les différences individuelles observées ne nous amènent pas à invalider les résultats à cette fréquence en dépistage car elles n'excèdent pas plus l'erreur audiométrique au niveau individuel que les différences observées aux autres fréquences (voir figure 3). Quoiqu'il en soit, nous avons considéré important de pallier ce problème pour optimiser la qualité des examens réalisés par notre service de dépistage.

En ce qui a trait au phénomène observé à 3 000 Hz et à 4 000 Hz à l'oreille gauche, c'est en étudiant les rapports périodiques de vérification d'étalonnage des appareils utilisés que nous avons trouvé l'explication la plus plausible. En effet, malgré le strict respect des normes y étant relatives [1,3], nous avons constaté une légère non-équivalence des niveaux de sortie entre les dispositifs utilisés en dépistage et celui utilisé en clinique. Par surcroît, cette non-équivalence était différente selon qu'il s'agissait de l'écouteur droit ou de l'écouteur gauche. Sans entrer dans le détail, mentionnons qu'en général les niveaux de sortie tendaient à être plutôt inférieurs aux niveaux attendus en vertu des normes et ce, tant pour les audiomètres de dépistage que pour l'appareil clinique. En ce qui a trait à la symétrie des niveaux de sortie d'un écouteur à l'autre, elle était relativement bonne en dépistage avec une différence moyenne inférieure à 1 dB, favorisant des niveaux plus élevés à l'écouteur gauche sur toute la gamme de fréquences audiométriques utilisées. En clinique, les niveaux de sortie étaient comparables entre écouteurs de 500 Hz à 2 000 Hz mais présentaient des différences légèrement supérieures à 1 dB jusqu'à 6 000 Hz, l'écouteur droit émettant les niveaux plus élevés. Concrètement, on favorisait systématiquement en dépistage la mesure de seuils relativement meilleurs à gauche alors qu'en clinique, la situation inverse survenait et ce, particulièrement à

3 000 Hz, 4 000 Hz et 6 000 Hz. Cet état de fait influençait inévitablement la comparaison des seuils mesurés en clinique à ceux obtenus en dépistage et pouvait même accentuer les différences observées.

Comme ces différences étaient en sens inverse d'une oreille à l'autre, la considération des résultats sans égard à la latéralité (section 3.3) devrait, sinon masquer, du moins atténuer les écarts observés à 3 000 Hz à 4 000 Hz. Il est à noter qu'à la fréquence de 6 000 Hz, pourtant soumise aux mêmes tendances des niveaux de sortie, aucune différence sensible n'a été mise en évidence. Nous croyons que l'erreur de mesure plus importante à cette fréquence [23] explique, du moins en partie, l'absence de relation entre les différences observées entre les seuils mesurés et entre les niveaux de sortie des appareils. En effet, les différentes sources de variation auxquelles la mesure de seuil en hautes fréquences est plus sensible (ex.: placement des écouteurs, affaissement même partiel du conduit auditif externe, présence d'acouphènes, etc.) confèreraient, selon toute vraisemblance, un poids relatif moindre aux niveaux de sortie effectifs des audiomètres.

3.2 Écarts Clinique - Dépistage observés versus attendus

L'erreur inhérente à l'examen audiométrique a été documentée et on peut retrouver dans la littérature une estimation de cette dernière pour des conditions optimales de mesure et pour un même équipement. En effet, dans un contexte d'expérimentation en laboratoire, l'erreur-type de mesure (1.96 Se) du dispositif utilisé a été établie à environ 5 dB pour la gamme de fréquences de 500 Hz à 4 000 Hz et à plus de 8 dB à 6 000 Hz [23].

Les histogrammes présentés à la figure 3 nous renseignent sur la distribution relative de l'ensemble des écarts clinique - dépistage individuels observés en fonction des fréquences audiométriques.

De façon générale, on constate que les différences observées, sans égard à la latéralité, se situent généralement entre -5 dB et 5 dB et se distribuent relativement bien de part et d'autre de ces valeurs. On note cependant que l'étendue de toutes les distributions est légèrement asymétrique vers la gauche témoignant de la tendance globale des seuils cliniques à être meilleurs que ceux mesurés en dépistage. Cet état de fait est parfaitement compatible avec le plus grand contrôle, en

clinique, des conditions de réalisation des examens.

En assumant une erreur de mesure de 5 dB de 500 Hz à 4 000 Hz et de 8 dB à 6 000 Hz [23], on s'attendrait dans des conditions optimales de validité et de fiabilité à ce que 95% des différences observées entre des seuils mesurés successivement sur un même dispositif se situent à l'intérieur d'un intervalle de 10 dB de 500 Hz à 4 000 Hz et de 16 dB à 6 000 Hz. La proportion des différences observées ici qui se situent à l'intérieur de la marge d'erreur attendue est de l'ordre de 90 % et plus. Compte tenu du contexte, les différences qui excèdent cet intervalle demeurent donc dans des proportions tout à fait acceptables, à savoir moins de 10%.

En effet, les conditions de mesures des seuils que nous comparons ne sont pas idéales au plan de la fiabilité d'autant plus que le caractère empirique de la présente étude n'a pu permettre le contrôle de certains facteurs. Ainsi, la présence éventuelle d'acouphènes, l'utilisation de bruit de masque à certaines fréquences lors de l'examen clinique, l'affaissement du conduit auditif externe non identifié au moment du dépistage sont autant de facteurs qui, ajoutés au laps de temps écoulé entre les examens, ont pu augmenter la variation audiométrique.

Si l'on s'attarde spécifiquement aux fréquences pour lesquelles les différences visuellement les plus significatives ont été notées lors d'une appréciation oreille par oreille, soit 500 Hz, 3 000 Hz et 4 000 Hz, on peut relever à la figure 3 certains éléments d'information complémentaires. À 500 Hz la tendance à mesurer des seuils meilleurs en clinique demeure évidente même si on ne considère pas les oreilles de façon individuelle. En effet, outre les différences observées dans l'intervalle -5 dB à 5 dB, 29,5% des différences sont négatives comparativement à 5 % de positives. À 3 000 Hz, la tendance demeure inversée mais devient un peu moins claire dans un contexte d'analyse sans égard à la latéralité (7,9 % de différences négatives contre 18,7 % positives). À 4 000 Hz, la distribution des écarts observés sans égard à la latéralité s'avère relativement symétrique (17,7 % vs 21,7 % respectivement pour les écarts négatifs et positifs).

Ces différents constats nous amènent à remettre en cause la légitimité de procéder à des analyses statistiques sans égard à la latéralité fondées sur le principe qu'il n'y a aucune raison de croire que l'erreur de mesure soit différente d'une oreille à l'autre. Tel que mentionné précédemment, bien que cette assertion soit fondée du point de vue de la réponse psycho-

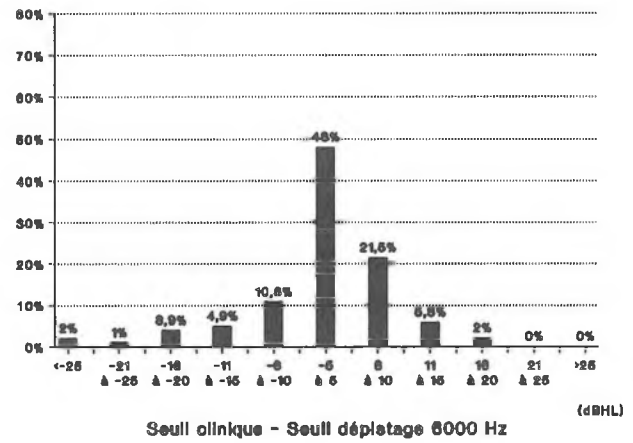
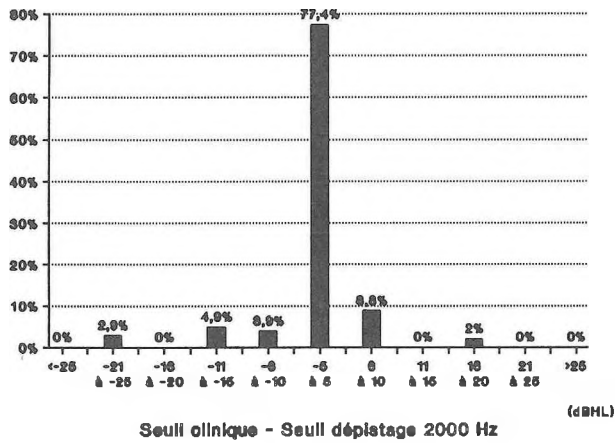
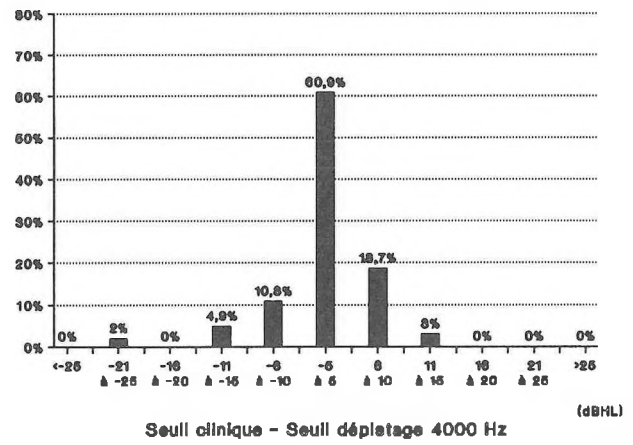
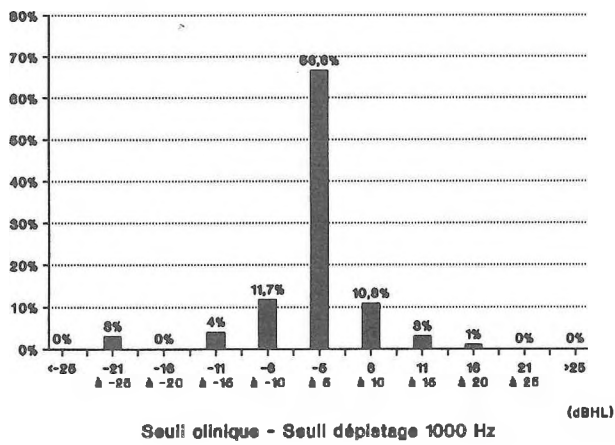
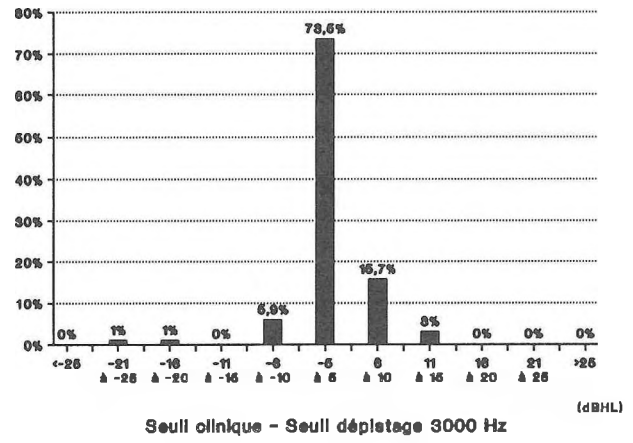
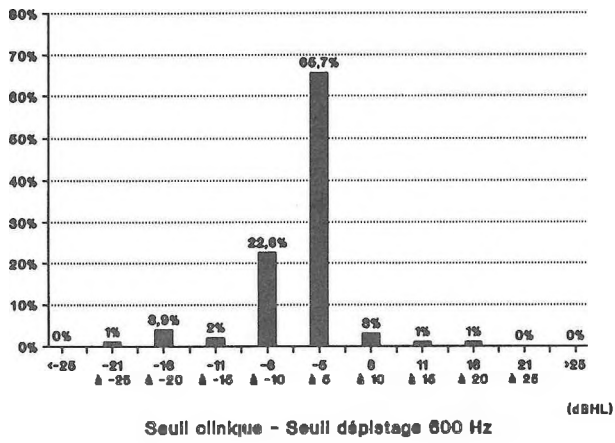
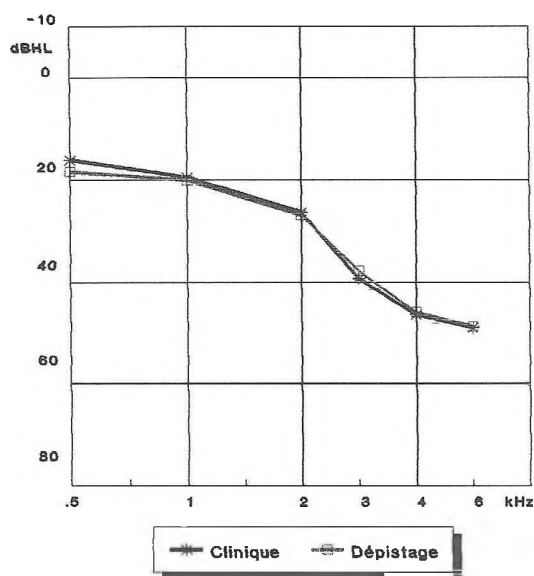


FIGURE 3: Répartition selon les fréquences des différences observées entre les seuils mesurés en clinique et les seuils mesurés en dépistage (n = 102 oreilles).

acoustique du sujet, elle ne tient cependant nullement compte des différences éventuelles susceptibles d'intervenir au niveau des appareils de mesure comme tels.

3.3 Analyse classique sans égard à la latéralité



Echantillon épuré n=102 oreilles

FIGURE 4: Comparaison des seuils moyens mesurés en clinique avec les seuils moyens mesurés en dépistage (n=102 oreilles).

Si l'on considère la figure 4 où les écarts clinique - dépistage sans égard à la latéralité sont visuellement appréciables, on constate un rétrécissement flagrant des écarts observés précédemment. Seules persistent des différences à 500 Hz et à 3 000 Hz. À partir des informations résumées au tableau 2, on constate que les différences moyennes entre les résultats obtenus en clinique avec ceux de dépistage sont effectivement minimales puisque inférieures à 1 dB sauf à 500 Hz et 3 000 Hz tel que déjà noté. À 500 Hz, les seuils observés sont en moyenne de 2,27 dB meilleurs en clinique qu'en dépistage tandis qu'à 3 000 Hz ils sont meilleurs en dépistage pour une différence moyenne de 1,60 dB. Les résultats des tests de *t* de Student démontrent que ces différences sont les seules qui sont significatives du point de vue statistique.

Compte tenu de ces résultats et des coefficients de corrélation de Pearson, somme toute excellents, d'aucuns concluraient en la validité des examens audiométriques de dépistage réalisés au PRA. Compte tenu de l'absence d'information en fonction de

TABLEAU 2 Comparaison Clinique -Dépistage, (n = 102 oreilles).

Fréquences	Différence ¹ (dBHL)		Corrélation de Pearson (p<.001)
	Moyenne	Écart type	
500	-2,27 ²	6,86	.828
1000	-0,57	7,51	.829
2000	-0,59	6,96	.915
3000	1,60 ²	5,95	.954
4000	0,54	7,12	.942
6000	0,18	9,63	.909

1 (Seuil clinique - Seuil dépistage)
2 Statistiquement significatif en vertu d'un test bilatéral de *t* de Student, p<.05

l'oreille testée, il serait très difficile, voire impossible, d'établir les sources d'erreurs expliquant les écarts observés à 500 Hz et 3 000 Hz. Il serait par contre aisé de suggérer la détérioration de l'audition à 3 000 Hz comme hypothèse explicative la plus plausible et ainsi, d'excuser le service de dépistage. À 500 Hz, compte tenu des contraintes de bruit de fond et de la plus grande permissivité à cet égard des normes édictées en dépistage, il serait facile d'en arriver à pointer ce problème particulier plutôt que tout autre.

C'est ainsi, que même une étude de validité comme la nôtre qui, par essence, se doit d'identifier tout problème entachant la mesure de seuils, aurait facilement pu faillir à cet égard si nous avions simplement procédé comme nous le suggérait la littérature audiolinguistique.

4. DISCUSSION

La démarche d'analyse rapportée dans ce qui précède avait notamment pour but de vérifier la validité des examens audiométriques réalisés par le service de dépistage du PRA et partant, de démontrer la faisabilité d'examen audiométriques valides dans un contexte de dépistage. Les résultats obtenus soutiennent cette hypothèse puisque les seuils mesurés par le service de dépistage du PRA correspondent étroitement à ceux obtenus en clinique pour le même groupe de travailleurs. Les seules différences observées ne sont pas de nature à remettre en cause la validité des examens mais témoignent davantage de

l'importance que revêt la vérification régulière de l'étalonnage des appareils et le contrôle périodique de toutes les autres sources potentielles de variation. À cet égard, la procédure utilisée, à savoir l'analyse de variance multivariée sur mesures répétées (MANOVA) s'est avérée une forme nouvelle d'appréciation de notre service de dépistage qui nous a permis d'améliorer la qualité des examens que nous réalisons en palliant les problèmes qu'elle a mis en évidence. C'est ainsi que dans le cas de la mesure du seuil à 500 Hz à l'oreille droite en dépistage, nous avons ajouté des éléments de contrôle à la procédure de recherche de seuils. Nous avons, par ailleurs, amélioré la symétrie des écouteurs et favorisé l'équivalence entre les différents niveaux de sortie des audiomètres de nos services.

En ce qui nous concerne, cette façon d'apprécier la qualité des examens constitue un outil d'intérêt pour la supervision des préposés aux examens et le contrôle de qualité des résultats audiométriques émanant de notre service de dépistage. Il est évident que les variations systématiques que nous avons mises à jour n'ont eu, si tant est, qu'une légère influence sur les résultats individuels. Certains types d'erreur peuvent cependant entraîner des variations beaucoup plus importantes qui vont avoir une influence sensible sur les résultats individuels pouvant entraîner des conséquences plus ou moins importantes tant pour l'individu que pour la collectivité. Nous avons d'ailleurs récemment démontré qu'un certain laxisme au niveau des conditions de validité des examens de dépistage pouvait occasionner 10 fois plus de références inutiles en clinique spécialisée [24] entraînant ainsi des frais pour le travailleur et générant des coûts sociaux indus.

L'application de la même procédure à des données émanant de d'autres services de dépistage nous a montré qu'elle était sensible à différentes sources de variation audiométrique telles notamment *l'intervention de l'examineur*, *la procédure de recherche et de détermination des seuils* ou le *respect des prérequis*. Il va sans dire que l'identification de problème à l'un ou l'autre de ces niveaux permet l'ajustement correspondant en termes de formation des préposés et d'encadrement immédiat. En somme, cette procédure représente un outil très intéressant pour tous les audiologistes ayant à superviser la réalisation d'examens auditifs effectués par des tiers. Elle permettrait notamment de détecter des problèmes entachant la qualité des examens et pouvant même en compromettre la validité qui, compte tenu de l'erreur audiométrique, demeureraient insoupçonnés à l'échelle des résultats individuels.

Nous avons par ailleurs démontré que l'utilisation stricte de la procédure d'analyse statistique suggérée par nos lectures nous aurait conduits à des conclusions erronées. Non seulement l'utilisation de cette procédure d'analyses statistiques passe outre le fait que les comparaisons de moyennes effectuées ne sont pas issues de deux échantillons indépendants puisqu'il s'agit de données colligées en deux temps auprès d'un même groupe d'individus mais elle impose une interprétation fondée sur plusieurs tests statistiques univariés portant sur des variables indéniablement dépendantes l'une de l'autre. Mais, il y a pire au plan des conséquences éventuelles.

Point n'est besoin de démontrer que l'erreur de mesure étant aléatoire, elle tend vers 0 avec l'augmentation du nombre de cas étudiés. En corollaire, tout biais systématique est ainsi mis en exergue ce qui constitue le principe de base de la procédure d'analyse des études de validité et de fiabilité. Or, bon nombre de ces études ne considèrent pas les niveaux de sortie effectifs des audiomètres lorsqu'elles comparent (ou "parent") des seuils audiométriques alors qu'il s'agit là d'une source potentielle de biais systématique évidente et facile à contrôler. Paradoxalement, en imposant le pairage des données audiométriques, les études de validité et de fiabilité laissent ainsi davantage de prise aux biais systématiques liés aux dispositifs utilisés. De fait, le pairage des données résulte en l'addition des biais liés à chacun des éléments de la paire. C'est ainsi que, dépendamment de leur sens respectif, le pairage contribuera soit à amplifier les biais soit à les camoufler. Il va de soi que l'identification de tels biais systématiques est encore plus compromise par le fait que les données audiométriques sont souvent traitées sans égard à la latéralité.

Finalement, la majorité des études en audiologie partent de la prémisse que la conformité aux normes établit la qualité de leur base de données, ce qui est en soi légitime. Nous avons été à même de constater que le strict respect des normes n'exclut nullement l'éventualité de biais systématiques qui peuvent facilement passés inaperçus. Or, toute stratification d'échantillon de seuils mesurés suivant un ou plusieurs paramètres impose le même type de considération quant à la qualité des données puisque cette opération sous-tend nécessairement une comparaison des résultats d'une strate à l'autre. Un biais systématique de mesure pourrait éventuellement compromettre le bien-fondé des conclusions qu'on tirera de données reflétant une double influence à savoir celle du paramètre étudié et celle du biais. Pour cette raison, il nous apparaît plus qu'opportun que la méthodologie de ce type

d'étude accorde aussi une attention spéciale à la correspondance des niveaux de sortie des appareils et à la symétrie entre les écouteurs de chacun des dispositifs utilisés pour la collecte des données.

En bref, on peut dire que la correspondance aux normes favorise effectivement la validité des mesures audiométriques effectuées mais ne constitue pas à elle seule un gage de validité des résultats du traitement statistique qu'on en fera et donc, par extension, des conclusions qui en découleront. Ce faisant, les spécifications particulières rapportées ci-haut, à savoir l'équivalence des niveaux de sortie entre audiomètres et la symétrie entre écouteurs, devraient s'appliquer d'emblée à partir du moment où des données audiométriques sont colligées par l'intermédiaire de plusieurs dispositifs ou que la période de collecte s'allonge dans le temps et ce, que ce soit dans la perspective immédiate ou éventuelle de traitement statistique de ces données.

5. CONCLUSION

D'après l'ensemble des résultats obtenus par notre analyse, il apparaît clairement que l'obtention de seuils valides dans un contexte de dépistage de la surdité professionnelle peut constituer une réalité. Il n'en demeure pas moins cependant que l'actualisation de cette réalité ne peut s'établir que sur la base d'un contrôle rigoureux de toutes les sources d'erreurs citées précédemment. Dans cette perspective, la procédure que nous avons utilisée s'est avérée un outil intéressant pour tout audiologiste ayant à superviser la réalisation d'examen de dépistage. En effet, facile à opérationnaliser, elle permet de mettre au jour d'éventuels biais systématiques de mesure et ainsi d'en identifier les causes probables. La supervision indirecte des préposés aux examens peut ainsi être améliorée et conséquemment, la validité et, plus globalement, la qualité des examens audiométriques.

Finalement, nous avons pu constater que même une légère non-équivalence asymétrique des niveaux de sortie entre les différents audiomètres utilisés peut engendrer un biais pouvant éventuellement mener à des conclusions basées sur des artefacts de mesure. Or, aucune norme relative à l'étalonnage des audiomètres cliniques ou de dépistage et à sa vérification périodique ne formule de recommandation à cet égard. À notre avis, ceci constitue une lacune qui ajoute à la variabilité audiométrique et qui mériterait qu'on s'y attarde dans la perspective d'une meilleure correspondance des seuils

audiométriques individuels d'un service à un autre. Cette prise en compte s'avère cependant indispensable dans les études portant sur un échantillonnage de données audiométriques.

REMERCIEMENTS

De judicieux commentaires ont été apportés sur les versions préliminaires de cet article par Pauline Fortier, Marjolaine Hamel, Raymond Hétu, Monique Lalonde, Odette Lemoine, Louise Paré et Madeleine Reny. L'expertise de Jean Lambert et sa gracieuse collaboration ont été particulièrement appréciées et ont permis de privilégier les analyses statistiques les plus appropriées. Ginette Gendron-Desroches a pour sa part fourni une précieuse collaboration en assurant le majeure partie du travail de bureautique que cet article a impliqué. Je suis par ailleurs redevable à l'ensemble de l'équipe du PRA pour ce qui a trait à la collecte et à la saisie des données utilisées et aussi pour m'avoir confié la responsabilité de la réalisation de cette étude. Je remercie finalement les réviseurs qui ont accepté de lire et de critiquer le contenu scientifique de cet article. J'ai pris connaissance de l'ensemble de leurs commentaires et les ai intégrés au mieux pour la version finale.

RÉFÉRENCES

- [1] Bureau de normalisation du Québec, "Normes recommandées pour les laboratoires d'examen auditif en milieu de travail", BNQ 5780-900, janvier 1984
- [2] Betz, N. E. & Weiss, D. J. Handbook of measurement and evaluation in rehabilitation. Baltimore, University Park Press, 1976.
- [3] American National Standard Institute, "Specification for audiometers". ANSI S3.6, 1989.
- [4] American National Standard Institute, "Criteria for permissible ambient noise during audiometric testing", ANSI S3.1, 1991.
- [5] American National Standard Institute, "Equivalent threshold force levels for audiometric bone vibrators", ANSI S3.26, 1981
- [6] Howell, R.W. et B.P.R. Hartley, "Variability in audiometric recording", Brit. J. Indust., Med. 29, 432-435, 1972.
- [7] Hartley, B.P.R., Howell R.W., Sinclair, A. et D.A.D. Slaterry, "Subject variability in short-term audiometric recording", Brit. J. Industr. Med., 30, 271-275, 1973.
- [8] Gauthier, J., Hamel, M., Lalonde, M., et O. Lemoine, "Évaluation de l'audition dans le cadre d'un programme de santé et de prévention en santé au travail", Modules 2 et 3, DSC Hôpital Saint-Luc, 1986.
- [9] Robinson, D.W., "Audiometric configurations and repeatability in

- noise-induced hearing loss", ISVR Technical report No. 123, University of Southampton, 1984.
- [10] Lutman, M.E., Cane, M.A. et P.A. Smith, "Comparison of manual and computer-controlled self-recorded audiometric methods for serial monitoring of hearing", *Brit. Journal of Audiology*, 23, 305-315, 1989.
- [11] Mets, J.T., "Reliability of industrial audiometry as a screening method for incipient noise-induced hearing loss", *S.Af. Med. J.*, 71 (1) : 35-39, 1987.
- [12] Valente, Michael, Valente, Maureen, Goebel, Joel, "Reliability and intersubject variability of the real ear unaided response", *Ear and Hearing*, 12(3) : 216-220, 1991.
- [13] Laukli, E., Fjermedal, O., "Reproductibility of hearing threshold measurements", *Scand. Audiol.*, 19(3) : 187-190, 1990.
- [14] Gauz, M.T., Ahroon, W.A. et S.D. Roberts, "High frequency Bekesy audiometry: II. Threshold test procedure, reliability and validity", *J. of Auditory Research*, 21, 21-28, 1981.
- [15] Jerlvall, L., Arlinger, S., "A comparison of 2 dB and 5 dB step size in pure-tone audiometry", *Scand. Audiol.*, 15, 51-56, 1986.
- [16] Arlinger, Stig D., "Normal hearing threshold levels in the low-frequency range determined by an insert earphone", *J. Acoust. Soc. Am.* 90(5) : 2411-2414, 1991.
- [17] Dobie, R.A., "Reliability and validity of industrial audiometry: Implications for hearing conservation program design", *Laryngoscope*, No. 93, 906-927, 1983.
- [18] High, W., Glorig, A. et J. Nixon, "Estimating the Reliability of auditory thresholds measurements", *J. of Auditory Research*. Vol. 1, 247-262, 1961
- [19] Woodford, Charles M., "The effect of calibration tolerance values on obtained audiometric threshold : A hearing conservation perspective", *Am. Ind. Hyg. Assoc. J.* 45(2) : 110-116, 1984.
- [20] Norusis, Marija J., SPSS for Windows: Advanced Statistics, release 6.0, SPSS inc, Chicago, 1993, 578 pages.
- [21] Organisation internationale de normalisation, "Acoustique - Seuil normal d'audition par conduction aérienne en fonction de l'âge et du sexe pour les personnes otologiquement normales", ISO 7029, 1984.
- [22] Gosztonyi, R.E., Vassallo, L.A. et Sataloff J., "Audiometric reliability in industry", *Arch. Environ Health*, Vol. 22, 1971.
- [23] Hétu, R., "Critical analysis of the effectiveness of secondary prevention of occupational hearing loss". *J. of Occup. Med.*, 21 (4), 251-254, 1979.
- [24] Lemoine, O., Gauthier, J., "Impact de la validité des examens auditifs de dépistage sur les interventions de suivi", *Programme régional -Audiologie*, Montréal, Juillet 1993.