Journal of the Association for Research in Otolaryngology (2023) 24:499–511
https://doi.org/10.1007/s10162-023-00909-y

JARO

ORIGINAL ARTICLE

Neural Degeneration in Normal-Aging Human Cochleas: Machine-Learning Counts and 3D Mapping in Archival Sections

Pei-zhe Wu[1,2] · **Jennifer T. O'Malley**[1,2] · **M. Charles Liberman**[1,2]

Received: 12 January 2023 / Accepted: 3 September 2023 / Published online: 13 November 2023

Abstract

Quantifying the survival patterns of spiral ganglion cells (SGCs), the cell bodies of auditory-nerve fibers, is critical to studies of sensorineural hearing loss, especially in human temporal bones. The classic method of manual counting is tedious, and, although stereology approaches can be faster, they can only be used to estimate total cell numbers per cochlea. Here, a machine-learning algorithm that automatically identifies, counts, and maps the SGCs in digitized images of semi-serial human temporal-bone sections not only speeds the analysis, with no loss of accuracy, but also allows 3D visualization of the SGCs and fine-grained mapping to cochlear frequency. Applying the algorithm to 62 normal-aging human ears shows significantly faster degeneration of SGCs in the basal than the apical half of the cochlea. Comparison to fiber counts in the same ears shows that the fraction of surviving SGCs lacking a peripheral axon steadily increases with age, reaching more than 50% in the apical cochlea and almost 66% in basal regions.

Keywords Machine learning algorithm · Spiral ganglion cells · Human temporal bone · Age-related hearing loss · Normal aging

Introduction

The classic approach to the analysis of cochlear neural health in otopathological studies of human temporal-bone sections has been to manually count the spiral ganglion cells (SGCs) using a light microscope equipped with a gridded eyepiece [1, 2]. This process is extremely labor intensive, since there can be up to 4000 cells in one set of slides from a healthy normal ear (consisting of every 10th 20-µm section through the cochlea). Stereological approaches offer a more time-efficient approach to the estimation of SGC counts [3, 4]. However, due to the spiral geometry of the cochlear duct, neither of these approaches accurately associates SGC counts with the region of the cochlear epithelium they innervate, i.e., stereology can compute only the total SGC count per cochlea, while the classic method bins the ganglion spiral into four half-turns.

Given the stereotyped size and shape of SGC nuclei in this type of histological material, we reasoned that SGC counting could be automated by applying machine learning (ML) to digitized scans of the slide sets. In addition to the benefit of being fully unbiased and automated, the algorithm can be enhanced by aligning digitized images of the serial sections, which allows algorithmic registration of the SGC locations in 3 dimensions (3D) and pairing of the SGC positions to the organ of Corti position based on the radial projection of the peripheral axons of the SGCs. Here, we developed custom software that automatically identifies and crops the region of interest, i.e., the cochlea, from a series of digitized slides, aligns the images in 3D, and passes them to a ML pipeline that identifies the SGCs, registers them in 3D, and computes the cochlear frequency region to which each SGC projects.

We validated the ML algorithm by comparing it to manual SGC counts from the same slides. We then applied this algorithm to 62 normal-aging ears from another human study [5], for which hair cell counts and other histopathological analyses were carried out, but for which cochlear neural survival was assessed by counting auditory-nerve peripheral axons, rather than SGCs. This was done because (1) counting SGCs is time-consuming, and (2) in an acoustic hearing context, peripheral axons are more relevant than

✉ Pei-zhe Wu
Peizhe-Wu@uiowa.edu

[1] Eaton-Peabody Laboratories, Massachusetts Eye and Ear Infirmary, 243 Charles St., Boston, MA 02114-3096, USA

[2] Department of Otolaryngology, Harvard Medical School, Boston, MA 02115, USA

SGCs, because peripheral axons degenerate faster than the cell bodies [6–8], and a SGC that is disconnected from the organ of Corti is non-functional without a cochlear implant.

Here, by combining the SGC counts with our prior counts of peripheral axons in a set of human ears with varying degrees of age-related hearing loss, we can estimate the proportion of surviving SGCs lacking peripheral axons. This issue has taken on added importance since the discovery that the de-afferentation of surviving inner hair cells is a major aspect of the histopathology of presbycusis [9] and may underlie the difficulties in understanding speech in a noisy environment experienced by those with sensorineural hearing loss [10]. Reconnecting the SGCs to the organ of Corti in regions where the inner hair cells survive may ultimately be possible via neurotrophin-based therapies [11], but little is known about the prevalence of disconnected SGCs in different hearing loss etiologies. In a more general context, the automation of SGC counting and the detailed 3D mapping of SGC locations should be useful in many otopathological studies, for example, in efforts to understand the differences in word-recognition performance in different cochlear implant users [12].

Methods

Subjects and Groups Archival slide sets of celloidin-embedded, serially sectioned human temporal bones from the Massachusetts Eye and Ear collection [13] were analyzed in the present study. Each slide set was processed as described elsewhere [14] and contained every 10th serial section, cut at 20 µm, and stained with hematoxylin & eosin (H&E). Included in the final data for this study were a subset ($n=62$) of cases selected for a prior study of normal aging [5], i.e., those that, in addition to satisfying the medical-history criteria we used to define "normal aging," also had adequate staining of SGCs for reliable ML analysis. The selected individuals ranged in age from birth to 95 years and included 41 males and 21 females. All procedures and protocols for the study of human tissues included informed consent and were approved by the Institutional Review Board of the Massachusetts Eye and Ear.

Image Acquisition and Training-Set Annotation The archival slide sets included here were digitized into "eslides" with a scanner (Leica Aperio AT2) that automatically tiles best-focus images with a high-resolution 20×objective, numerical aperture (N.A.) = 0.75, and stitches the tiles into a composite that can be seamlessly zoomed from low to high power. The pixel size of the highest-resolution image is 0.503 µm. The scanner automatically detects the best focal plane and scans the slide at that focal plane only. Two models were trained on two separate sets of images. In total, 7 temporal bone cases were included in the training material for the "blue" model and 6 cases in the "red" model. These case were chosen to include a wide range of staining intensities and color balance, i.e., "blue" when the hematoxylin dominates and "red" when the eosin dominates, since both variables significantly affect the ML results. In each temporal bone, 2–4 eslides were selected to include both cross sections and tangential sections through Rosenthal's canal. Each eslide was then cropped into "samples" (3600×2400 pixels per sample) for the final ML pipeline. All samples from temporal bones in the same model were randomly allocated to either the training set (70% of all samples), the validation set (20%), or the test set (10%). In the final training material, the blue model included 77, 22, and 11 samples in the training, validation, and test sets, respectively. The red model included 59, 16, and 9 samples in the training, validation, and test sets, respectively. The training materials were annotated by hand in Photoshop, painting the nucleus, not the cytoplasm, because the boundary of the SGC cytoplasm is often faint and difficult to delineate. These annotations were validated by assessing the physical slide under a Nikon E800 microscope with a high N.A. 40×objective.

Manual SGC Counting For validation purposes, all the SGCs in two sets of human temporal-bone slides were manually counted under a Nikon E600 microscope equipped with a high N.A. 40×objective and a computer-aided anatomy system (Neurolucida). An experienced observer focused through the entire thickness of the slide and counted all the SGCs with a nucleus. The slide was projected onto a computer monitor to facilitate hand annotation. The locations of the SGCs were marked, and the x, y, z coordinates generated by the software were exported to custom software for further analysis.

Machine Learning Pipeline Our ML pipeline ultimately converts an eslide, which is a 3-channel (RGB) 8-bit image (Fig. 1A), into a binary image, the "mask," in which all the pixels in SGC nuclei are labeled as "1" and all other pixels are labeled as "0" (Fig. 1C). As schematized, we split the eslide into pieces, i.e., the "sample," and reassembled them at the end to avoid overloading the computer memory. The ML pipeline we used (Fig. 1B) was adopted from an open-source semantic-segmentation package (Pavel Yakubovskiy, 2020, GitHub, URL: https://github.com/qubvel/segmentation_models.pytorch) using the *"resnet34"* backbone, *"imagenet"* encoder, *"DiceLoss"* loss function, and "*Adam"* optimizer. The basic architecture is known as a *U-Net*, because of the two paths it contains: (1) a decoder path which decreases the dimensionality of the image in stepwise fashion and is used to capture the nature of the elements and (2) an encoder path, which increases the dimensionality while

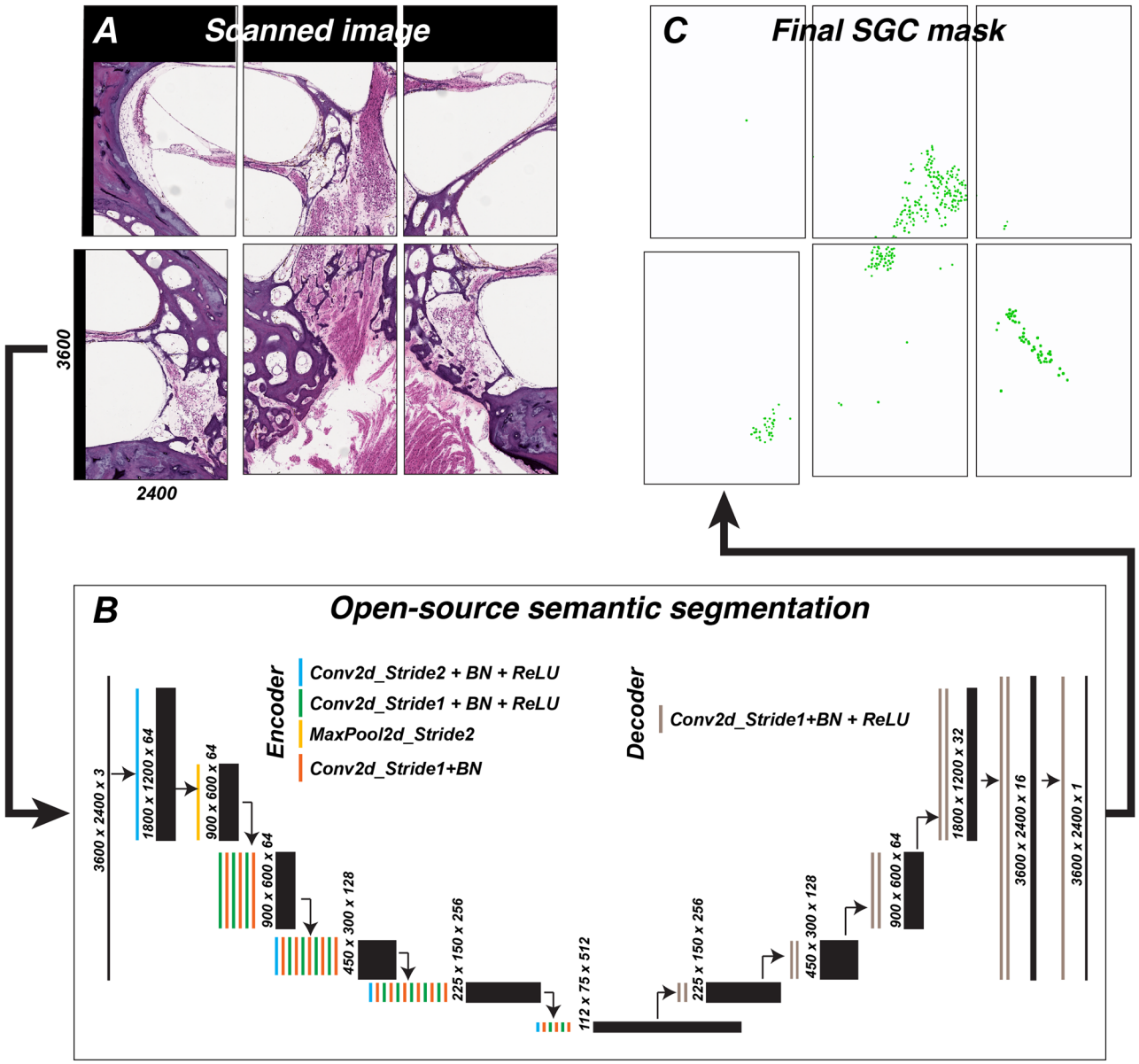

Fig. 1 The ML pipeline condenses each RGB image (**A**) into a matrix of progressively smaller three-dimensional arrays (black boxes in **B**) with progressively more parameters in the third dimension (encoding arm on the left in **B**) and then progressively reconstitutes (decoding arm to the right in **B**) a 2D image that ultimately contains only the "mask" (**C**). **B** The size of the image in each step of the algorithm is shown in length × width × depth. The mathematical operations are illustrated with colored stripes (legend in **B**). **C** The final mask is a single-channel image with the resolution of the original image (3600 × 2400), in which the locations of SGC nuclei shown in green. See "Methods" for further details

enabling a precise localization of the elements of interest. Completing these tasks requires a variety of mathematical operations, as indicated in the schematic including *2D convolution*, *MaxPooling*, and *Rectified Linear Unit (ReLU)* that are carried out in a specific sequence. Each of these operations carries a set of "weights." Training the model is the process of calculating the correct "weights" by supplying the pipeline with hand-traced "masks" and the "samples" on which they are based. The finalized set of "weights" under a specific "pipeline" is called a "model." The training materials were augmented with image augmentation, which includes flipping, shifting, scaling, padding, random cropping, and adding Gaussian noise. The model was trained with 80 epochs and was updated when the "loss," i.e., the difference between the algorithm-generated mask and the hand-annotated mask, became smaller than previous epochs. As noted above, we trained separate "blue" and "red" models, but the final mask was derived by merging the masks

from both. The performance of the pipeline was evaluated via an $F1$ score, calculated from the samples (3600 × 2400 pixels) and expressed as a mean ± standard deviation of the mean (SEM).

$$F1\text{score} = \frac{2 * \textit{True Positive}}{2 * \text{True Positive} + \text{False Negative} + \text{False Positive}}$$

At the end of pipeline, the SGCs identified by the algorithm were overlayed on the eslide for the observer to assess the success of the operation. The centroid of each SGC nucleus was calculated with *cv2.connectedcomponents* (a function in OpenCV library) for 3D reconstruction of the SGCs.

Reconstruction of the Organ of Corti Spiral Custom code was written in Python to automatically identify the cochlea in each eslide then crop, and align, the serial images. The algorithm used for image alignment was *MOTION_EUCLIDEAN* (a function in OpenCV library). Landmarks for reconstruction of the cochlear spiral were manually positioned at the intersections of the heads of inner and outer pillar cells in each section through the organ of Corti, which resulted in ~ 100 landmarks along the spiral in each case (gray circles in Suppl. Fig. 1). The landmarks for the organ of Corti were interpolated in 3D in MATLAB, using slide number and the section thickness (20 µm) to estimate the z depth and resampled into 1000 equally spaced points along the spiral. Irregularities introduced by stretching and shrinking of the sections were minimized by smoothing using *Smoothing Splines* (a function in MATLAB). The percent distance from the apex (d in the equation below), along the organ of Corti spiral, was converted into cochlear frequency using a map for humans [15], modified to produce the apical-most and basal-most best frequencies of 0.1 and 20 kHz, respectively:

$$Hz = 165.4 * (10^{2.1 * \frac{d}{100}} - 0.4)$$

Reconstruction of Rosenthal's Canal and Mapping of SGCs The spiraling Rosenthal's canal, which houses the SGCs, was reconstructed in a similar fashion. The landmarks for Rosenthal's canal in each section were placed on the thin bony wall separating the ganglion from the adjacent scala tympani, positioned at the point of maximum curvature of this bony partition (black circles in Suppl. Fig. 1). The cochlear frequencies for the Rosenthal's canal spiral were then derived by matching each position along organ of Corti spiral to the nearest location on the canal spiral, using the 3D coordinate system. This process introduces some inaccuracies in the cochlear apex (see "Discussion"). The cochlear frequency for each SGC was then defined by the cochlear frequency of the closest spot on Rosenthal's canal spiral.

Statistical Analysis Regression lines and confidence intervals in Figs. 3 and 8 were computed using the *fitlm* function @MATLAB. The correlation coefficient used in Steiger's Z test were computed using the *corrcoef* function @MATLAB.

Results

Developing the Machine-Learning Model

The human temporal bone library at the Massachusetts Eye and Ear (MEE) [13] includes specimens (> 2500 temporal bone cases) processed over a > 60-yr span, and the slide sets vary in section condition and staining, which poses a significant challenge for automatic recognition. The two most salient challenges are section flatness and staining intensity/color balance. Some of the slide sets processed before 1970 could not be analyzed, because the sections were not sufficiently flat, and the digitized images of Rosenthal's canal area were not uniformly in focus (Suppl. Fig. 2A). Others collected before 1970 had poor color differentiation between the nucleus and cytoplasm, such that identifying SGCs from the eslides was unreliable (Suppl. Fig. 2B).

Among the cases with sufficient color differentiation and good section flatness, the relative color balance between the hematoxylin (blue) and eosin (red) varied from blue dominance (Fig. 2A) to red dominance (Fig. 2B), due to fading of the hematoxylin. Models trained on the reddish sets failed to identify SGCs on the bluish sets and vice versa. Normalizing the RGB colors or converting RGB into grayscale images did not boost the performance across different slide sets. Therefore, we trained two models using images from exemplars from each of the color sets. The final mask was generated by superimposing the masks from the two models. The combined model worked equally well on both bluish and reddish slide sets (Fig. 2A_2, B_2). The $F1$ score of the final model is 0.87 ± 0.02 (mean ± SEM) when compared to the ground truth in the test set.

To estimate the number of cases in the MEE library with sufficient image quality for the SGC analysis, we applied the model to one mid-modiolar section from each of 181 cases [2, 5]. These cases were originally processed from 1953 to 2013 and thus represent a wide range of staining intensity. As each case is analyzed, the software superimposes the silhouettes of all the "found" SGC nuclei onto the eslide image, allowing the user to assess the overall accuracy of the ML result by toggling the superimposed silhouettes on and off (e.g., Fig. 2). In this way, the results were easily sorted by visual inspection into "successful" cases where virtually all the SGCs were correctly identified by the algorithm

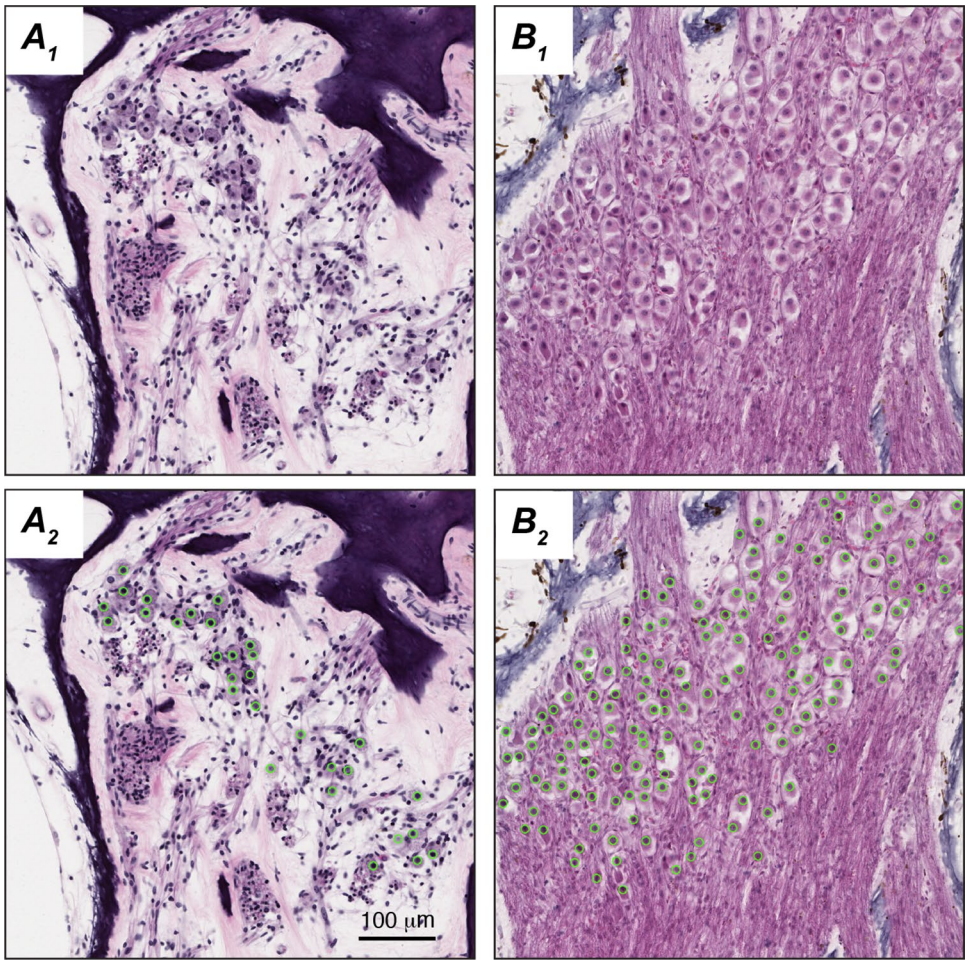

Fig. 2 The ML algorithm performs equally well in a range of staining conditions and SGC densities. Each pair of images shows one cross section through Rosenthal's canal from an eslide ($\mathbf{A_1}$, $\mathbf{B_1}$) and the same cropped image with the superimposed masks of the SGC nuclei identified by the ML algorithm shown in green circles ($\mathbf{A_2}$, $\mathbf{B_2}$). The scale bar in $\mathbf{A_2}$ applies to all panels

and "unsuccessful" cases, where a great number of SGCs were "missed" by the ML pipeline. The lack of success arose because either (1) a large area was out of the focus (Suppl. Fig. 2A) or (2) the staining contrast was insufficient (Suppl. Fig. 2B). Of the 181 cases, 30% (18/60) of the pre-1970 cases and 80% (97/121) of the post-1970 cases had sufficient color differentiation and section flatness for the ML algorithm. Of note, many of the poorly stained cases cannot be unambiguously analyzed by an experienced human observer either.

Evaluating the Machine-Learning Model

To evaluate the ML model performance more quantitatively, we compared automatic SGC counts to manual SGC counts in every section of two temporal-bone slide sets using a computer-aided anatomy program (Neurolucida) to track the x, y, z positions of the manually placed cell markers. The ML analysis required < 30 min per case (including the time to add markers for 3D reconstruction of ganglion and organ of Corti spirals) compared to ~ 30 h per case by the manual approach.

The overall correlation between the machine count and the manual count was very high ($r^2 = 0.945$, Fig. 3). A human-level $F1$ score can be calculated by superimposing the ML results with the x, y, z coordinates exported from Neurolucida (Fig. 4A). The $F1$ score of the final model is 0.87 ± 0.02 (mean ± SEM) when compared to manual count using Neurolucida. Superimposing the two sets of validation results and analyzing the mismatches by examination of the original histological section identified several sources of disagreement (Fig. 4). Most of the false negatives, i.e., SGCs identified by manual count, but missed by the algorithm, were SGCs located near the tops or the bottoms of the 20-µm sections. Since the eslide images are captured from one focal plane toward the middle of the z-depth, the nuclei at the top and bottom of the section appear blurry (Fig. $4B_2$ blue arrows) and are not captured by the algorithm. A smaller subset of false negatives occurs in areas with densely packed SGCs: closely spaced pairs of nuclei (Fig. $4B_1$ green arrows), which can arise when two SGCs overlie each other in the z dimension, are sometimes identified by the algorithm as one nucleus. The analysis identified very few false positives. On rare occasions (< 0.1% of all cells identified

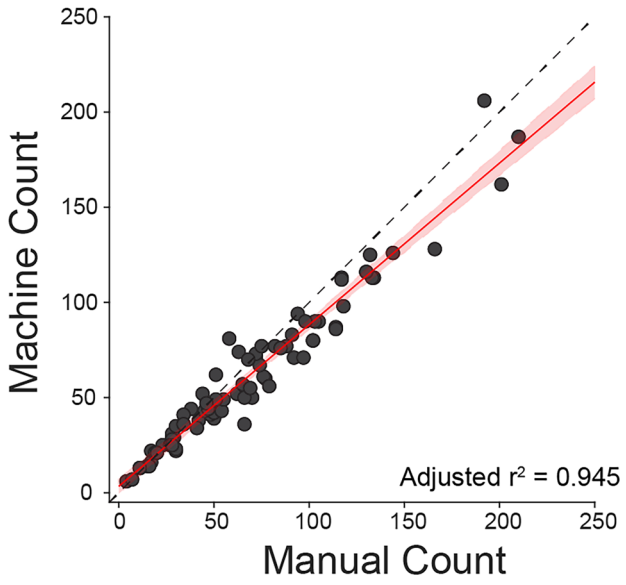

Fig. 3 ML counts are highly correlated with manual counts from the same sections. Data from all 81 sections through both ears of one temporal-bone donor, chosen for its dark staining that made identification of SGCs more challenging. The best-fit straight line and 95% confidence intervals are shown. The dashed black line is the equivalency line. Both ML and manual counts represent the total number of SGCs per slide, i.e., including all sections through Rosenthal's canal

by the algorithm), red blood cells or Schwann cells were misclassified as SGCs. These observations help explain why the comparison in Fig. 3 suggests a small, but systematic, undercounting by the algorithm, which increases as the cell density increases.

Mapping SGCs to the Organ of Corti

In classic otopathological studies of human temporal bones [16], SGC counts are summed within each of the four half turns of Rosenthal's canal, resulting in four "segment counts" from base to apex. This classic method provides a relatively crude view of the correlation between patterns of SGC loss and damage to other cell types in the cochlear duct. A more nuanced view can be obtained with the present dataset, because the computer-driven cell counts also provide location data, allowing 3D reconstruction of the SGCs and mapping to the region of the organ of Corti to which they project (Fig. 5B_1,B_2,B_3). We accomplished this mapping by assuming a radial trajectory from the ganglion spiral to the cochlear spiral. According to our analysis, the length of the cochlear spiral innervated by a unit length of Rosenthal's canal remains relatively constant at the base then grows toward the apex. On average, the basal 50% of

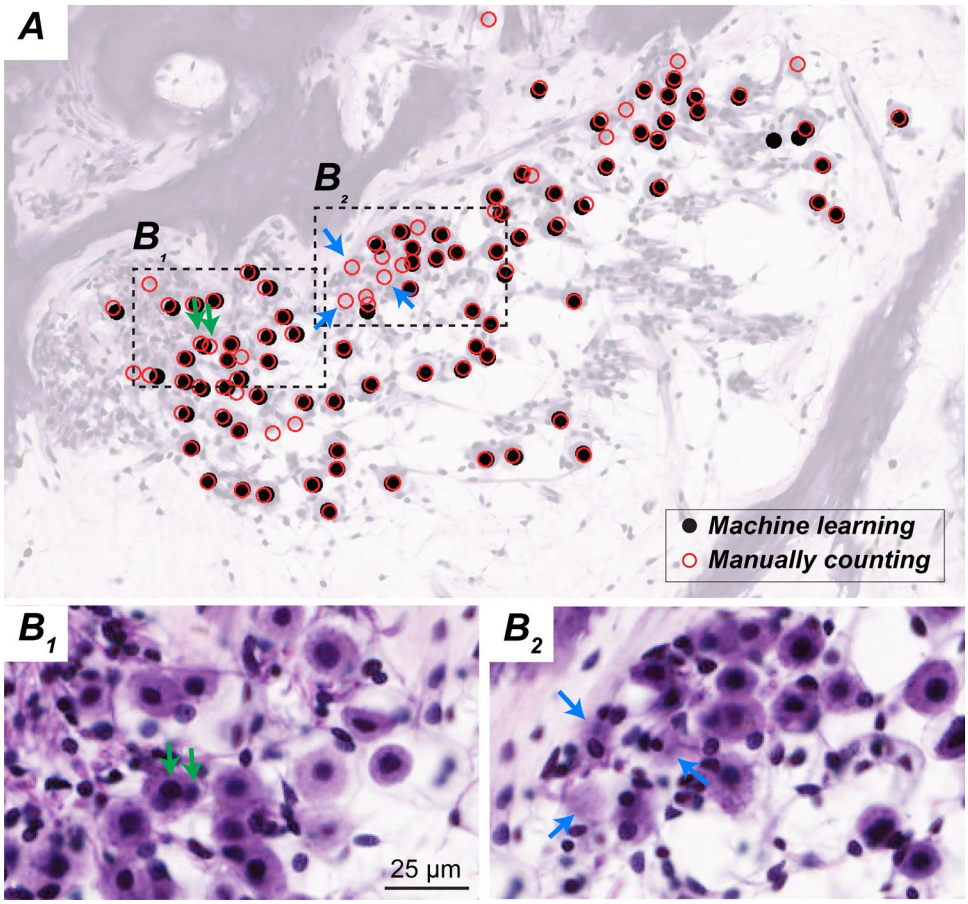

Fig. 4 Comparison of ML-generated SGC locations (black) with the locations identified by an observer viewing the physical slides (**A**) with a 40×DIC objective in a computer-aided anatomy system (red). The areas within the dashed rectangles in **A** are shown in $\mathbf{B_1}$ and $\mathbf{B_2}$. Green arrows point to paired nuclei that were identified as one SGC by the algorithm. Blue arrows point to nuclei that are out of focus and therefore ignored by the algorithm. Scale bar in $\mathbf{B_1}$ also applies to $\mathbf{B_2}$. Images in $\mathbf{B_1}$ and $\mathbf{B_2}$ were gamma adjusted to better differentiate the cytoplasm from the nucleus

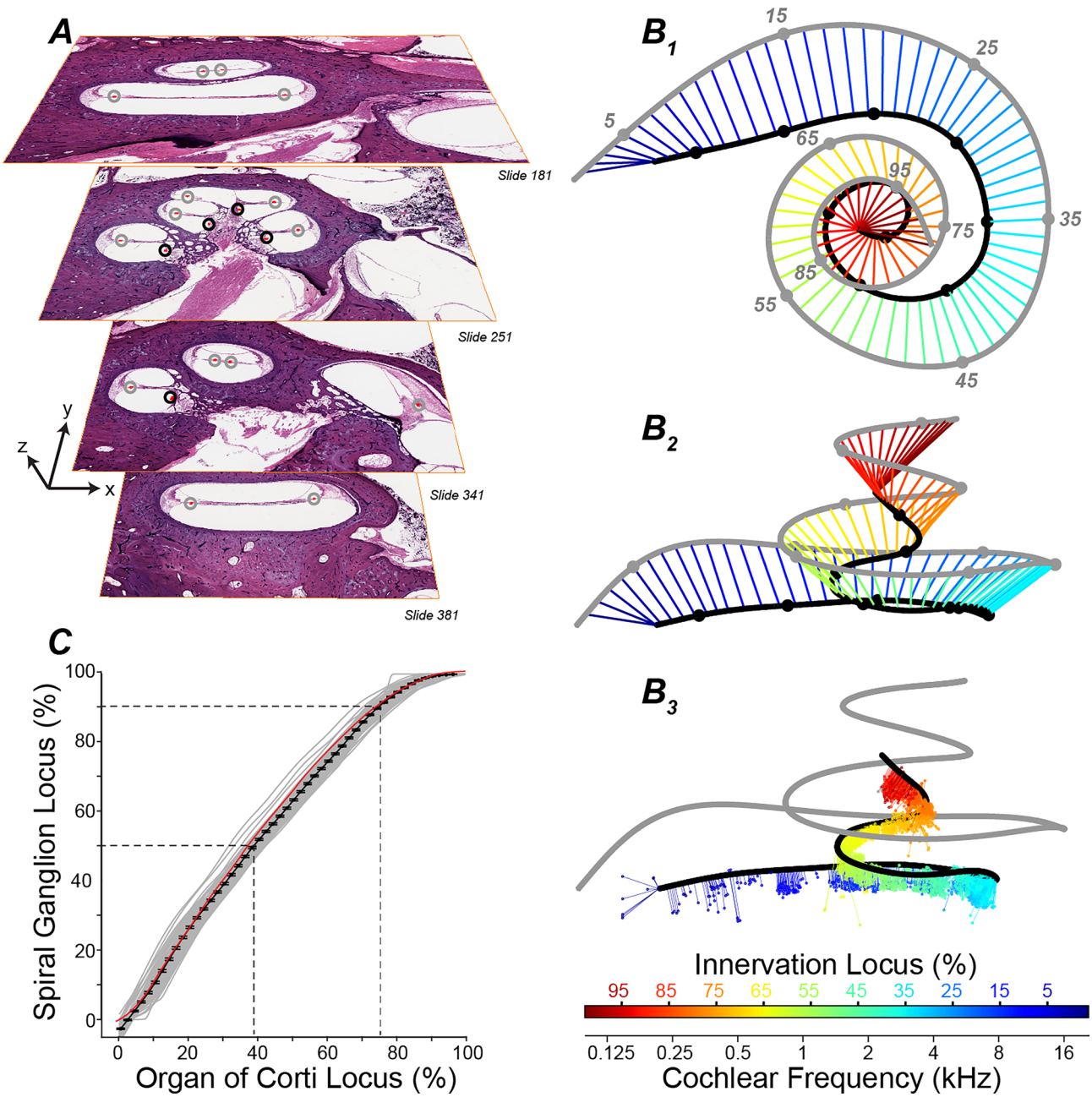

Fig. 5 3D reconstruction and frequency mapping of SGCs. **A** Selected slides showing the placement of landmarks for reconstruction of the organ of Corti (gray circles) and Rosenthal's canal (black circles). A more detailed view of the landmark placement can be found in Suppl. Fig. 1. **B_1, B_2, B_3** Mapping SGCs to hair cells is done in three steps. First, the percentage from base of the cochlear spiral (gray numbers in **B_1**) is converted to cochlear frequency using Greenwood Function [15]. Second, each 1% increment along the organ of Corti (gray line) is projected to the nearest point along Rosenthal's canal (black line), as shown by lines color-coded to cochlear frequency. **B_1** and **B_2** show orthogonal projections of these relations. Third (**B_3**), each SGC is projected to the nearest point along Rosenthal's canal. **C** The correspondence between SGC locus and hair cell locus from all 62 cases in this study: individual cases shown in gray, mean (±SEM) shown in black. The red line is the mapping function derived from a prior whole mount study of human cochleas [17]. Key in **B_3** applies to **B_1, B_2,** and **B_3**

a human cochlea is innervated by SGCs in the basal 60% of Rosenthal's canal, while the apical 25% of the cochlea is innervated by the apical 10% of Rosenthal's canal (Fig. 5C). At the apical end, i.e., above the 80% point (~0.3 kHz), the ganglion spiral is extremely tight, such that a compact cluster of SGCs at the end of the canal innervates a large region of the organ of Corti, and the peripheral axons take a curvilinear course from ganglion to hair cell [17]. Despite the problems

inherent in mapping the apical regions, this result matches exceptionally well with a prior study in which the mapping was measured directly in 8 human cochleas processed as microdissected whole mounts (red line in Fig. 5C) [17].

To further validate the accuracy of our mapping approach, we divided the SGC counts per mm of cochlear length by the number of inner hair cells per mm of the sensory epithelium at corresponding cochlear frequencies [9] to estimate the number of SGCs innervating each IHC as a function of cochlear location (Fig. 6). The density of SGCs in the youngest cases in our sample (< 10 yrs) peaks in the middle of the cochlea at ~ 14 neurons per IHC and falls off dramatically at cochlear regions apical to the 0.5 kHz region and basal to the 10 kHz region (Fig. 6). Since the peripheral processes of SGCs are always unbranched [18], these SGC values should match those for peripheral axon counts in the osseous spiral lamina: indeed, peripheral axon counts peak at ~ 15 fibers per IHC, according to a prior human study of cochlear wholemounts [19]. As expected, the estimates of SGCs per inner hair cell decrease with increasing age, across the entire cochlear frequency spectrum (Fig. 6).

Comparing Age-Related Loss of SGCs and Peripheral Axons

Since peripheral axons of auditory-nerve fibers often degenerate before their cell bodies in the spiral ganglion [8, 20], SGC counts will generally be higher than peripheral axons counts. Comparison of these two measures of cochlear neural survival from two cases from the temporal-bone archives of the Massachusetts Eye and Ear with histories of occupational noise exposure [21] illustrates this trend (Fig. 7). These two individuals had similar downsloping audiograms and showed similar degrees of peripheral axon loss (Fig. 7C, F). However, the 95-year-old, who lived much longer after retirement from his noisy occupation, had many fewer SGCs surviving, especially in the basal half of the cochlea.

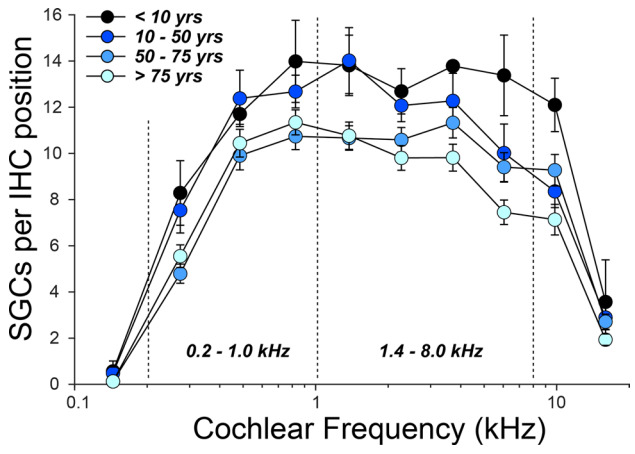

Fig. 6 Innervation density at different age groups, expressed as SGCs per IHC position. Data from all 62 normal-aging ears, divided by age group (0–10 years old: $n=3$; 10–50 years old: $n=6$; 50–75 years old: $n=23$; 75–100 years old: $n=30$). Each point (mean ± SEM) is derived by binning the raw SGCs counts according to their inferred innervation locus (10% length bins), multiplying by ten to account for the intervening sections, and then dividing by the number IHCs normally present in that span of the organ of Corti, as measured in a prior study of human cochleas processed as whole mounts [9]

In our prior studies of archival temporal bones [5, 22], we counted peripheral axons rather than SGCs, because SGCs that are disconnected from the organ of Corti are nonfunctional in the absence of a cochlear implant. However, our algorithm provides a unique opportunity to describe the frequency-specific SGC survival from the prior histopathological analysis of "normal aging" in which we had previously quantified hair cells [23], auditory-nerve peripheral axons, and the stria vascularis, but not the SGCs [5]. Of the 76 temporal bone cases from the prior studies we applied this algorithm to, 62 (82%) of these cases were classified to be "successful," corresponding to those with the flattest sections and the most robust staining, as described above. The result reveals a number of important trends (Fig. 8). First, there is relatively little age-related loss of SGCs in the apical half of the cochlea (1.4% per decade) compared to the basal half (3.3% per decade), and the difference was statistically significant ($n=62$, $Z=-0.199$, $p<0.05$ *by Steiger's Z test*). Second, the age-related loss of peripheral axons outstrips the loss of SGCs, in both apical and basal halves of the cochlea. Thus, in both cochlear regions, the fraction of surviving SGCs that lack a peripheral axon steadily increases with age reaching more than 50% in the apical half of the cochlea and almost 66% in the basal half of the cochlea. Lastly, the data show that the SGCs are significantly more resilient to degeneration than the IHCs (or OHCs), especially in the basal half of the cochlea, where the IHC loss increases steeply with increasing age.

Discussion

Technical Considerations in the Analysis of SGC Survival

The classic method of analyzing SGCs in humans is to count all the SGCs in every 10th section, then multiply by 10, and adjust by a "correction factor" [1–3, 24], to eliminate double counting for cells (or nuclei) split between sections [25, 26]. To speed the analysis, some animal studies have counted SGCs within a small, randomly placed region of interest inside each ganglion cross section [27]. A more formalized approach to random sampling and estimation of total SGC count is implemented in design-based stereology, which has been used in cat [28], rat [29], mouse [4], and human [3, 30,

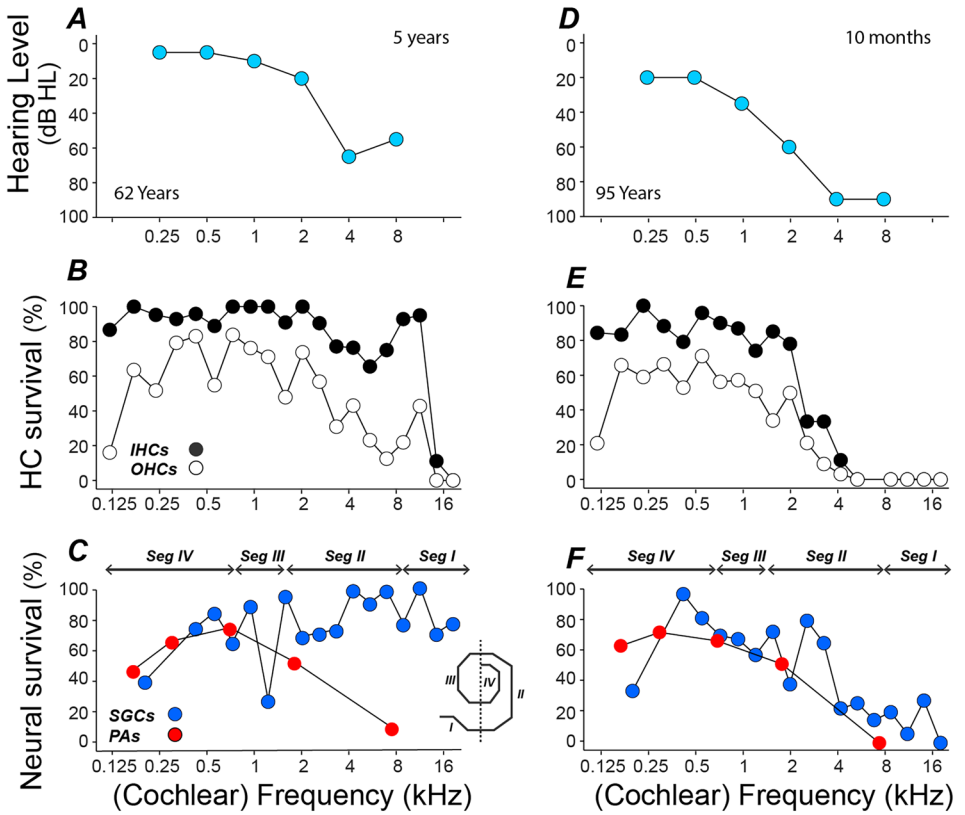

Fig. 7 Audiogram (**A**, **D**), hair cell survival (**B**, **E**), and neuronal survival (**C**, **F**) from a 62-year-old male (**A**, **B**, **C**) who was exposed to noise from working in a foundry for three years and a 95-year-old male (**D**, **E**, **F**)who was exposed to high-level noise in series of scientific experiments [21]. The interval between audiogram and death is noted in the top right corner of each audiogram (**A**, **D**). Hair cell counts (**B, E**) and peripheral axon (PA) counts (**C, F**) are from a prior study, where preparation and analysis techniques are described in detail [5]. Hair cell survival is assessed as described in a prior study [23]. OHC survival is the mean of all three rows. SGC and PA survivals (**C, F**) are normalized to the 95th percentile of the data from all ages in the present study. The segment map for SGC analysis is shown in the inset in **C**. Key in **B** applies to **B** and **E**. Key in **C** applies to **C** and **F**

31]. However, stereology is limited by its inability to assess SGC survival as function of cochlear location.

The machine-learning method described here has significant advantages over previous approaches. First, it speeds the process of counting every SGC in a standard temporal-bone slide set without any loss of accuracy. Since our computer-driven counts were based on digital images, scanned and stitched at one focal level with a 20 × objective, the algorithm slightly undercounted SGCs, when compared to meticulous manual counts using high-power DIC objectives to through-focus each section and computer-aided anatomy to tag each cell and avoid undercounting in high-density areas. However, the classic approach was not that comprehensive, using visual counts through gridded eyepieces [1, 2], which must also fail to capture all the SGCs, especially in areas of high SGC density. Thus, our ML method yielded virtually identical results when applied to the same slide sets used in some of these prior studies using classic methods.

In the manual analysis, the classification of SGC nuclei in Rosenthal's canal is not always unambiguous. Thus, it is not clear that discrepancies between the manual and algorithmic identification always represent errors on the part of the machine-driven analysis (e.g., arrow in Fig. 3A). The ML result has the advantage that it should be less subject to user bias and criterion drift. A further ambiguity in the analysis of SGCs is the distinction between type I and type II neurons, innervating inner and outer hair cells, respectively [32]. In humans, the nuclei of both SGC types are spherical [33], and, although their mean diameters differ (type I are larger), the size distributions overlap. Thus, manual-counting studies do not typically attempt to differentiate the two types [1, 2], and the present ML algorithm does not either. Since the type II neurons constitute only 5–10% of the SGC population [34], the counts will not differ greatly, whether or not the type II cells are included.

An additional advantage of the computer-driven approach is that it improves the power of the analysis by registering SGCs in a 3D coordinate space and relating them algorithmically to organ of Corti location (Figs. 5 and 7) in a much more precise mapping than prior methodologies. The accuracy of our mapping technique is seen in the close agreement between our estimates of SGCs per inner hair cell in the normal ear (Fig. 6) and a prior study that cut cross sections through the osseous spiral lamina and counted auditory-nerve peripheral axons in young normals [35]. The latter showed a peak of 14.8 fibers/per inner hair cell, at the 1.0 kHz region, compared to our estimate of 14.0 fibers per inner hair cell at the same locus.

This 3D mapping functionality could also be useful in studying the variation in performance of cochlear implant users by relating SGC survival to speech scores. Past attempts using the classic method of binning SGC

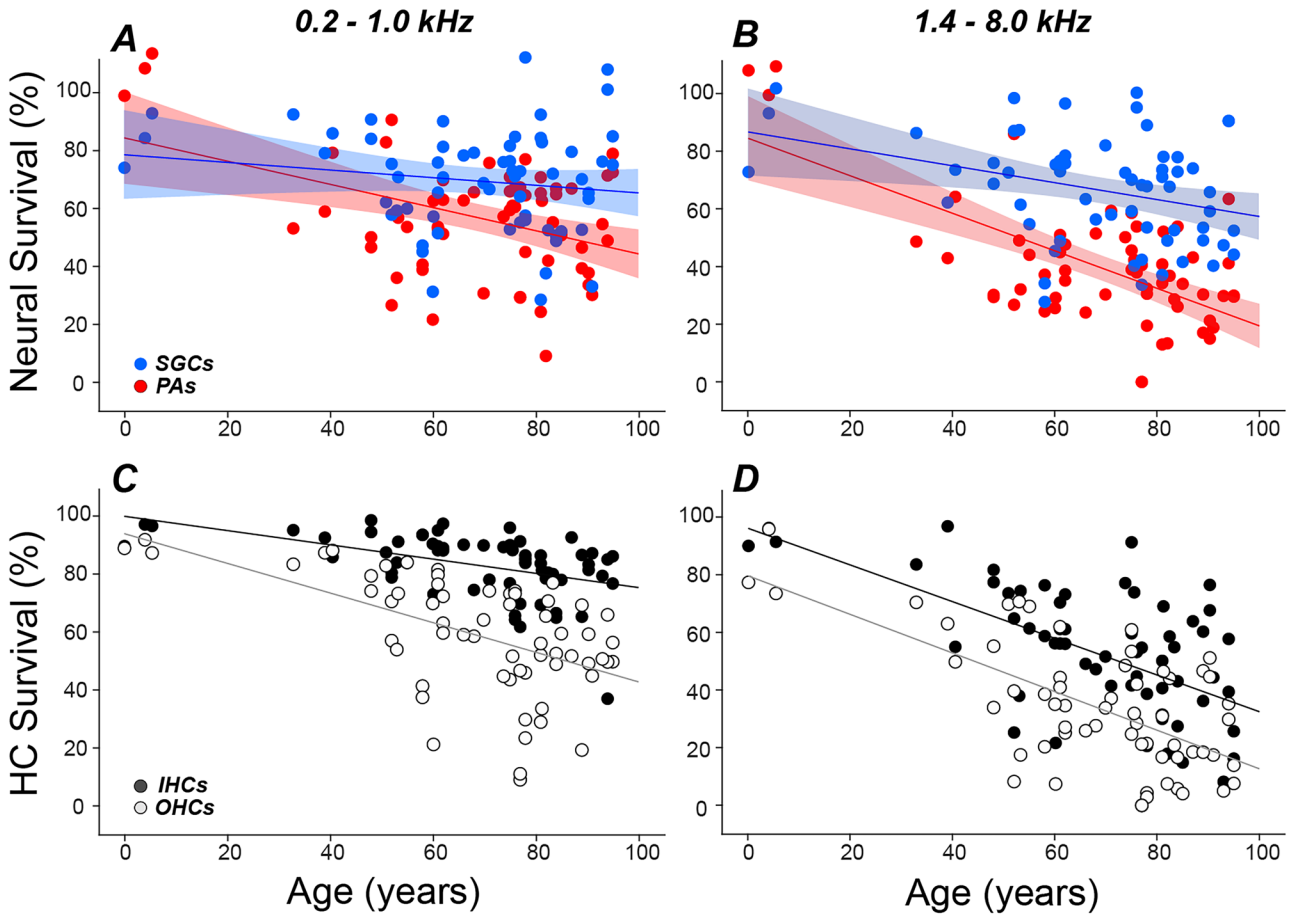

Fig. 8 Age-related loss of peripheral axons is faster than the loss of SGCs in both the apex (**A**) and the base (**B**) of the cochlea. Data summarized here are extracted from all the normal-aging cases analyzed in the present study ($n=62$). Each point shows the cell survival (as coded in the key in **A** and **C**) in a single case averaged separately for the apical (**A** and **C**, 0.25–1.0 kHz inclusive) and the basal (**B** and **D**, 1.4–8.0 kHz inclusive) regions of the cochlea. The SGC data and peripheral axon data are normalized to the 95th percentile of all cases in this study. The best-fit straight lines are shown for all four measures; however, the 95% confidence limits are shown only for peripheral axons and SGCs (**A, B**). The inner hair cell (IHC), outer hair cell (OHC), and auditory-nerve fiber peripheral axon data are from a prior study, where details of the measurement techniques can be found [5, 23]. Key in **A** applies to **A** and **B**. Key in **C** applies to **C** and **D**

survival into half-turn segments have not been very successful [36–38]. The present analysis yields 17 bins along the cochlear frequency axis (Fig. 6) and should be able to more accurately characterize "dead zones" when present. Furthermore, with small adjustments, our algorithm can compute distances from each SGC to each cochlear implant electrode, the locations of which can be derived from the capsular fibrous sheath surrounding the cochlear electrode arrays [39]. Such data could also be used to power modeling studies of current spread in implanted ears [40].

The approach used here could easily be applied to any serial or semi-serial set of cochlear specimens in any mammalian species, assuming the sections have been stained in a way that reveals the SGCs and their nuclei with sufficient contrast, and given an investment of a few hours to hand-annotate a training set of images. Of course, with new material, there are alternate approaches to produce 3D image stacks of unsectioned cochleas, for example, using light-sheet fluorescent microscopy [41], in which any cellular elements that can be immunostained could also be counted. Such approaches, if applied to image stacks of the requisite resolution, could be more accurate than the reconstructions produced here, but these more modern methods cannot be applied to the unique archival libraries of sectioned human temporal bones [13].

Age-Related Degeneration of SGCs vs. Peripheral Axons in Humans

SGC degeneration depends on a number of factors, including age [1], damage to hair cells and supporting cells [42], and hearing loss etiology [14]. Our study found a slower age-related SGC loss (2.5% per decade) than a prior study (5.0% per decade) [1]. However, the prior study included

cases with hearing loss from many etiologies (e.g., congenital anomalies, Meniere's disease, vestibular Schwannoma, ototoxic drugs, and sudden deafness) whereas our study was based on a "normal-hearing" group excluding subjects with an otologic diagnosis other than presbycusis [5]. Our data agree closely with a more recent study [2] (3.2% per decade), who included cases with a variety of otologic diagnoses, but excluded cases with significant hair cell loss who skew SGC losses toward greater values.

Our study found slower SGC degeneration in the apical (1.4% per decade) than the basal (3.3% per decade) half of the cochlea, which mirrors the apical-basal gradient of hair cell and supporting cell degeneration [5, 42, 43]. These apical-basal gradients are consistent with the view that neurotrophins released from the sensory epithelium are critical for SGC survival [44, 45]. A prior study also found a steeper age-related SGC loss in the basal-most of their four SGC half-turn segments, which failed to reach statistical significance [2]. The difference may arise because the prior study excluded cases with complete wipeout of the organ of the Corti, which is most often seen in the basal half of the cochlea and which is typically associated with more severe loss of SGCs [42].

The degeneration of SGCs often proceeds from the periphery such that cell bodies, with intact central axons, can survive long after loss of the peripheral axons connecting them to hair cells [14]. This phenomenon is key to the success of cochlear implants. However, the degree and speed of the cell body vs. axonal degeneration depend on hearing loss etiology in ways that are incompletely understood. In animal studies of acoustic injury, the loss of peripheral axons appears several weeks after the insult, while loss of SGCs (and central axons) progresses for months to years [7, 46–48]. Correspondingly, a human temporal bone study, using serial section ultrastructure, showed that 68.5% of the SGCs in the 4-kHz region of a subject with longstanding noise-induced hearing loss lacked peripheral axons [20].

With respect to age-related hearing loss, one study suggested that a subset of aging individuals, with "neuritic presbycusis," show significant loss of peripheral axons despite "normal" SGC populations [8]. Our prior study of "normal-aging" humans suggested that loss of peripheral axons from surviving SGCs was the rule, rather than the exception, in presbycusis [9], based on an observed peripheral axon loss of 7.7% per decade compared to the 3.2% per decade loss of SGCs reported previously [2]. Here, we provide the first quantitative comparison of SGC vs. peripheral axon survival in data extracted from the same ears. As predicted, age-related loss of peripheral axons outstrips age-related SGC loss by a factor of 2–3: 6.2 vs. 3.3% per decade in the basal cochlea and 4.0 vs. 1.4% per decade in the apical half of the cochlea. Our prior study showed that the age-related loss of inner hair cells was similar to that of SGCs, not peripheral axons [9]. These results suggest that, in aging ears, > 50% of surviving SGCs throughout the cochlea are lacking peripheral axons and thus are disconnected from the surviving inner hair cells they once contacted.

Understanding the proportion and cochlear locations of SGCs lacking peripheral axons has translational relevance, given that this type of deafferentation could potentially be reversed. After ototoxin-induced destruction of the organ of Corti, cochlear perfusion of neurotrophins can promote neurite extension from surviving SGCs along the denuded basilar membrane, despite the lack of hair cell targets [49]. Similarly, neurotrophin application to the round window after noise-induced cochlear deafferentation can restore SGC synapses to surviving inner hair cells and rescue the associated functional deficits [11] and genetically forced overexpression of neurotrophins in cochlear supporting cells at middle age can halt the further age-related loss of synaptic connections between auditory nerve fibers and inner hair cells in aged mice [50]. Given the deleterious effects of this type of primary neural degeneration on hearing in noise [10], reconnecting this large population of SGCs to their surviving hair cell targets should significantly reduce the dysfunction of age-related hearing loss.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10162-023-00909-y.

Author Contribution PZW, JTO, and MCL conceived the study. PZW and JTO carried out the data collection. PZW created the computer algorithm and carried out the data analysis and the statistical analyses. MCL and PZW wrote the manuscript, and JTO edited it.

Funding Research supported by grants from the National Institute on Deafness and other Communicative Disorders: P50 DC 015857 and U24 DC 020849.

Data Availability The data that support the findings of this study are available from the corresponding author upon reasonable request.

Code Availability The Code that support the findings of this study are available from GitHub.com, (https://github.com/PeizheWu/SGC-counting-and-3D-mapping). Further support is available from the corresponding author upon reasonable request.

Declarations

Competing Interests The authors declare no competing interests.

References

1. Otte J, Schunknecht HF, Kerr AG (1978) Ganglion cell populations in normal and pathological human cochleae. Implications for cochlear implantation. Laryngoscope 88(8 Pt 1):1231–1246
2. Makary CA et al (2011) Age-related primary cochlear neuronal degeneration in human temporal bones. J Assoc Res Otolaryngol 12(6):711–717

3. Ishiyama G et al (2011) Spiral and vestibular ganglion estimates in archival temporal bones obtained by design based stereology and Abercrombie methods. J Neurosci Methods 196(1):76–80
4. Schettino AE, Lauer AM (2013) The efficiency of design-based stereology in estimating spiral ganglion populations in mice. Hear Res 304:153–158
5. Wu PZ et al (2020) Age-related hearing loss is dominated by damage to inner ear sensory cells, not the cellular battery that powers them. J Neurosci 40(33):6357–6366
6. Liberman MC, Kiang NY (1978) Acoustic trauma in cats. Cochlear pathology and auditory-nerve activity. Acta Otolaryngol Suppl 358:1–63
7. Johnsson LG, Hawkins Jr JE (1976) Degeneration patterns in human ears exposed to noise. Ann Otol Rhinol Laryngol 85(6 PT. 1):725–39
8. Chen MA et al (2006) Presbycusic neuritic degeneration within the osseous spiral lamina. Otol Neurotol 27(3):316–322
9. Wu PZ et al (2018) Primary neural degeneration in the human cochlea: evidence for hidden hearing loss in the aging ear. Neuroscience 407:8–20
10. Resnik J, Polley DB (2021) Cochlear neural degeneration disrupts hearing in background noise by increasing auditory cortex internal noise. Neuron 109(6):984–996.e4
11. Suzuki J, Corfas G, Liberman MC (2016) Round-window delivery of neurotrophin 3 regenerates cochlear synapses after acoustic overexposure. Sci Rep 6:24907
12. Khan AM et al (2005) Is word recognition correlated with the number of surviving spiral ganglion cells and electrode insertion depth in human subjects with cochlear implants? Laryngoscope 115(4):672–677
13. Merchant SN et al (2008) Human temporal bone consortium for research resource enhancement. J Assoc Res Otolaryngol 9(1):1–4
14. Merchant SN, Nadol JB (2010) Schuknecht's pathology of the ear, 3rd edition Shelton. People's Medical Publishing House - USA, CT, p 942
15. Greenwood DD (1990) A cochlear frequency-position function for several species–29 years later. J Acoust Soc Am 87(6):2592–2605
16. Schuknecht HF (1974) Pathology of the Ear Cambrideg. Harvard Universtiy Press, MA, p 574
17. Stakhovskaya O et al (2007) Frequency map for the human cochlear spiral ganglion: implications for cochlear implants. J Assoc Res Otolaryngol 8(2):220–233
18. Liberman MC, Oliver ME (1984) Morphometry of intracellularly labeled neurons of the auditory nerve: correlations with functional properties. J Comp Neurol 223(2):163–176
19. Spoendlin H, Schrott A (1988) The spiral ganglion and the innervation of the human organ of Corti. Acta Otolaryngol 105(5–6):403–410
20. Liu W et al (2015) The pre- and post-somatic segments of the human type I spiral ganglion neurons–structural and functional considerations related to cochlear implantation. Neuroscience 284:470–482
21. Davis H et al (1950) Temporary deafness following exposure to loud tones and noise. Acta Otolaryngol Suppl 88:1–56
22. Wu PZ et al (2021) Primary neural degeneration in noise-exposed human cochleas: correlations with outer hair cell loss and word-discrimination scores. J Neurosci 41(20):4439–4447
23. Wu PZ et al (2019) Assessing fractional hair cell survival in archival human temporal bones. Laryngoscope 130(2):487–495
24. Hinojosa R, Marion M (1983) Histopathology of profound sensorineural deafness. Ann N Y Acad Sci 405:459–484
25. Nadol JB Jr (1988) Quantification of human spiral ganglion cells by serial section reconstruction and segmental density estimates. Am J Otolaryngol 9(2):47–51
26. Abercrombie M (1946) Estimation of nuclear population from microtome sections. Anat Rec 94:239–247
27. Stamataki S et al (2006) Synaptic alterations at inner hair cells precede spiral ganglion cell loss in aging C57BL/6J mice. Hear Res 221(1–2):104–118
28. Leake PA et al (2011) Brain-derived neurotrophic factor promotes cochlear spiral ganglion cell survival and function in deafened, developing cats. J Comp Neurol 519(8):1526–1545
29. Guthrie OW (2017) Noise stress induces an epidermal growth factor receptor/xeroderma pigmentosum-a response in the auditory nerve. J Histochem Cytochem 65(3):173–184
30. Kaur C et al (2020) Age-related changes in the number of cresyl-violet-stained, parvalbumin and NMDAR 2B expressing neurons in the human spiral ganglion. Hear Res 388:107883
31. Tang Y, Lopez I, Ishiyama A (2002) Application of unbiased stereology on archival human temporal bone. Laryngoscope 112(3):526–533
32. Kiang NY et al (1982) Hair-cell innervation by spiral ganglion cells in adult cats. Science 217(4555):175–177
33. Nadol Jr JB, Burgess BJ, Reisser C (1990) Morphometric analysis of normal human spiral ganglion cells. Ann Otol Rhinol Laryngol 99(5 Pt 1):340–8
34. Kiang NY et al (1984) Afferent innervation of the mammalian cochlea. In: Bolis L, Keynes RD, Maddrell SHP (eds) Comparative physiology of sensory systems. Cambridge University Press, pp 143–161
35. Spoendlin H, Schrott A (1989) Analysis of the human auditory nerve. Hear Res 43(1):25–38
36. Nadol JB Jr et al (2001) Histopathology of cochlear implants in humans. Ann Otol Rhinol Laryngol 110(9):883–891
37. Khan AM et al (2005) Effect of cochlear implantation on residual spiral ganglion cell count as determined by comparison with the contralateral nonimplanted inner ear in humans. Ann Otol Rhinol Laryngol 114(5):381–385
38. Fayad JN, Linthicum FH Jr (2006) Multichannel cochlear implants: relation of histopathology to performance. Laryngoscope 116(8):1310–1320
39. Ishai R et al (2017) The pattern and degree of capsular fibrous sheaths surrounding cochlear electrode arrays. Hear Res 348:44–53
40. Jahn KN, Arenberg JG (2020) Identifying cochlear implant channels with relatively poor electrode-neuron interfaces using the electrically evoked compound action potential. Ear Hear 41(4):961–973
41. Hutson KA et al (2021) Light sheet microscopy of the gerbil cochlea. J Comp Neurol 529(4):757–785
42. Suzuka Y, Schuknecht HF (1988) Retrograde cochlear neuronal degeneration in human subjects. Acta Otolaryngol Suppl 450:1–20
43. Sugawara M, Corfas G, Liberman MC (2005) Influence of supporting cells on neuronal degeneration after hair cell loss. J Assoc Res Otolaryngol 6(2):136–147
44. Miller JM et al (1997) Neurotrophins can enhance spiral ganglion cell survival after inner hair cell loss. Int J Dev Neurosci 15(4–5):631–643
45. Stankovic K et al (2004) Survival of adult spiral ganglion neurons requires erbB receptor signaling in the inner ear. J Neurosci 24(40):8651–8661
46. Johnsson LG (1974) Sequence of degeneration of Corti's organ and its first-order neurons. Ann Otol Rhinol Laryngol 83(3):294–303
47. Kujawa SG, Liberman MC (2009) Adding insult to injury: cochlear nerve degeneration after "temporary" noise-induced hearing loss. J Neurosci 29(45):14077–14085
48. Fernandez KA et al (2015) Aging after noise exposure: acceleration of cochlear synaptopathy in "recovered" ears. J Neurosci 35(19):7509–7520

49. Wise AK et al (2005) Resprouting and survival of guinea pig cochlear neurons in response to the administration of the neurotrophins brain-derived neurotrophic factor and neurotrophin-3. J Comp Neurol 487(2):147–165
50. Cassinotti LR et al (2022) Cochlear neurotrophin-3 overexpression at mid-life prevents age-related inner hair cell synaptopathy and slows age-related hearing loss. Aging Cell 21(10):e13708