European Review for Medical and Pharmacological Sciences 2023; 27: 6496-6509

The identification of sex-specific biomarkers in peripheral blood mononuclear cells from elderly individuals with ischemic stroke

L.-H. QIN[1,2], L. CHEN[1], X. CAO[1], T.-J. HUANG[1], Z.-Y. LI[1], S. LI[2], G.-Z. WANG[3]

[1]School of Nursing, Hunan University of Chinese Medicine, Changsha, China
[2]Key Laboratory of Hunan Province for Prevention and Treatment of Integrated Traditional Chinese and Western Medicine on Cardiocerebral Diseases, Hunan University of Chinese Medicine, Changsha, China
[3]The Second Clinical School of Chinese Medicine, Hunan University of Chinese Medicine, Changsha, China

L. Chen, S. Li, and G.-Z. Wang contributed equally to this study

Abstract. – OBJECTIVE: The aim of this study was to identify sex-specific biomarkers for ischemic stroke (IS) prophylaxis in elderly individuals.

MATERIALS AND METHODS: The GSE22255 dataset for elderly individuals with IS was retrieved from the gene expression omnibus database. Thereafter, gene ontology and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed, as well as gene set enrichment analysis (GSEA). Furthermore, protein-protein interactions (PPIs) were explored using the STRING database, and to screen central genes from the Cytoscape PPI network, corresponding to peripheral blood samples from elderly individuals, we used the molecular complex detection plug-in and cytoHubba. Moreover, a Venn diagram was used to visualize the key genes common among elderly women and men with IS. Statistical analysis was also performed, and receiver operating characteristic (ROC) curves were constructed to evaluate the specificity and sensitivity of the prediction of IS in the elderly.

RESULTS: Compared with the healthy controls, in elderly women with IS, 511 biological process (BP) terms, 16 molecular function (MF) terms, and 34 KEGG terms were significantly enriched, whereas in the elderly men with IS, 681 BP terms, 12 MF terms, and 44 KEGG terms were enriched. The GSEA revealed 99 and 140 significantly enriched gene sets in elderly women and men with IS, respectively. Furthermore, in the PPI network, 10 hub genes for each sex with high specificity and sensitivity were identified using ROC curves.

CONCLUSIONS: Ten genes for each sex with significant differential expression were also identified in individuals with IS. The novel sex-specific gene targets may be promising diagnostic or prognostic markers and potential therapeutic targets for IS in the elderly.

Key Words:
Ischemic stroke, Differentially expressed genes, Sex-specific, Elderly.

Introduction

Ischemic stroke (IS) is the most common type of stroke (75-80%) and is the second leading cause of death and one of the leading causes of disability worldwide[1]. Furthermore, the incidence, prevalence, disability, mortality, and recurrence of stroke show a consistent annual increase within the aging population, thus creating a huge socioeconomic impact on society[2]. Currently, IS is diagnosed mainly through imaging examination. However, stroke mimics pose diagnostic challenges, making identifying a real stroke difficult and leading to a delay in diagnosis[3]. Currently, the only approved therapies for acute IS are thrombolytic therapy and endovascular thrombus removal[4]. Then the arrow time window for treatment efficiency and serious side effects of IS therapy limit the efficacy of such treatments[5]. Therefore, strategies for early diagnosis and appropriate treatments for IS are critical. Together, innovative diagnostic tools and biomarkers are urgently required to improve the accuracy of clinical IS diagnosis, monitor disease progression, and improve therapeutic efficacy.

In recent years, increasing evidence[6,7] has shown that the injury mechanisms, clinical characteristics, outcomes, treatments received, and response to IS treatment are associated with sex. Furthermore, the occurrence and prognosis of IS

Corresponding Author: Lihua Qin, MD; e-mail: 479157643@qq.com

are associated with age. Women have also been reported[8] to exhibit nontraditional symptoms at the beginning of a stroke that often contribute to a delay in diagnosis and treatment. The pathophysiological biomarkers of IS show partial efficacy, and it is still unclear how gender affects IS variability[9,10]. Therefore, both sex and age should be given considerable attention in studies related to IS diagnostics and therapeutics.

Given that IS is associated with the gender and age of individuals, identifying sex-specific blood-based biomarkers to predict stroke occurrence and progression has important clinical implications in IS prevention, diagnosis, and treatment. Bioinformatics has unveiled new avenues for identifying novel biomarkers for early disease detection[11]. Furthermore, using peripheral blood mononuclear cells (PBMCs) in this respect is advantageous owing to easy access to blood samples and increased detection sensitivity[12]. In this study, we hypothesized that using bioinformatic analysis, it is possible to identify the sex-specific hub genes in PBMCs from elderly patients with IS to predict the occurrence, development, and treatment response effects of elderly patients with IS. Therefore, this study aimed to identify potential sex-specific hub genes for the early diagnosis, prevention, and treatment of IS in elderly women and men. Therefore, we analyzed PBMC data corresponding to elderly individuals with IS in the gene expression omnibus (GEO) dataset GSE22255 using integrated bioinformatic analyses.

Materials and Methods

Dataset Analysis

The GSE22255 dataset, which contains PBMC gene expression results corresponding to patients with IS profiled on the Affymetrix Human Genome U133 Plus 2.0 Array (GEO GPL570), was downloaded from the GEO database (available at: https://www.ncbi.nlm.nih.gov/gds/?term=GSE22255) *via* the R package 'GEOquery' (version info: R 3.2.3, Biobase 2.30.0, GEOquery 2.40.0, limma 3.26.8). The GEO dataset included 20 patients with IS (10 men and 10 women) and 20 sex- and age-matched controls (10 men and 10 women). Finally, based on the study design, data corresponding to 10 elderly patients with IS[13] (5 men and 5 women aged ≥ 60 years), and 10 control subjects (5 men and 5 women aged ≥ 60 years) were included for further analysis. Patients with IS had only one stroke episode, at least 6 months before blood samples collection, and the control subjects had no family history of stroke[12].

Genomic Analysis of Differentially Expressed Genes (DEGs)

Probes without gene annotation or multiple matching gene symbols were removed. Subsequently, DEGs between IS and the control samples in the GSE22255 dataset were screened using the R package 'limma' (Univ Melbourne Dept Math & Stat, Parkville, Australia) with a cut-off criterion of $|\log_2$fold-change (FC)$| \geq 1$ and an adjusted p-value [Benjamini-Hochberg (BH) method] < 0.05. The R packages packages 'ggplot2' (version 3.3.3, RStudio, Copenhagen, Denmark) and 'ComplexHeatmap' (version 2.2.0, German Cancer Research Center, Heidelberg, Baden-Württemberg, Germany) were then used to analyze the DEGs, and to construct volcano plots and heat maps to visualize their distribution and expression[14].

Functional Enrichment Analysis

The biological significance of the DEGs was assessed *via* gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) pathway enrichment analyses[15,16]. Statistical analysis and visualization of GO/KEGG results were performed using the R software. Specifically, the R packages 'ggplot2' and 'clusterProfiler' (Southern Medical University, Guangzhou, Guangdong, China) were used to visualize the results of the GO/KEGG enrichment analysis of the selected data, respectively. Furthermore, the R package 'org.Hs.eg.db' (version 3.10.0, Fred Hutchinson Cancer Research Center, Seattle, WA, USA) was used for ID conversion, and statistical significance was established at an adjusted p-value (BH method) < 0.05.

Gene Set Enrichment Analysis (GSEA)

The R packages 'clusterProfiler' and 'ggplot2' were used to perform GSEA and data visualization, respectively[15,17]. Furthermore, c2.cp.v7.2. symbols. gmt was the reference gene dataset, whereas the molecular signatures database collection (available at: https://www.gsea-msigdb.org/gsea/msigdb/index.jsp) was the gene set database. A threshold for significant enrichment was established with an adjusted p-value (BH method) < 0.05 and a false discovery rate (FDR) value < 0.25.

Protein-Protein Interaction (PPI) Network Analysis

The STRING database, version 11.0 (available at: https://www.string-db.org/), evaluated

PPIs among DEG products[18]. Subsequently, the PPI network was constructed based on information in the STRING database with a confidence score of > 0.4 and visualized using Cytoscape 3.8.2 software (Institute of Systems Biology, Boston, MA, USA). The molecular complex detection (MCODE) plug-in and the cytoHubba function in Cytoscape were used to identify potential hub genes in the PPI network.

Venn Diagram Analysis

R package 'ggplot2' (version 3.3.3) was used for Venn diagram analysis and data visualization to identify the key genes responsible for the sex differences observed in elderly individuals with IS.

Statistical Analysis

An independent sample *t*-test or Wilcoxon rank sum test was performed using the R package 'ggplot2' (version 3.3.3) to determine differences in hub gene expression between IS samples and healthy controls. Furthermore, to evaluate the specificity and sensitivity of the hub gene values in predicting IS in elderly women and men, receiver operating characteristic (ROC) curves were generated using the 'pROC' (Swiss Institute of Bioinformatics, Basel, CH, Switzerland) and 'ggplot2' packages in R. In addition, the area under the curve (AUC) was calculated to evaluate the diagnostic value of the hub genes. Statistical significance was established at $p < 0.05$.

Results

DEGs Between Elderly Women and Men with IS and Their Respective Healthy Controls

Compared to the gene expression results of healthy female controls, 174 DEGs were identified in elderly women with IS, and among these, 25 DEGs were upregulated, and 149 were downregulated. Compared to the gene expression in healthy male controls, 172 DEGs were identified in elderly men with IS and 141 and 31 of these genes were upregulated and downregulated, respectively (Table I). As shown in the volcano plots in Figure 1a and 1b, these IS-specific DEGs showed a clear distinction between elderly female and male patients and the corresponding healthy controls. The top 10 upregulated and downregulated DEGs are listed in Table I and represented using a heat map in Figure 1c and 1d.

Enriched BPs, MFs, and KEGG Enrichment Pathways in Elderly Individuals With IS

After setting thresholds with an adjusted *p*-value < 0.05 and q-value < 0.2 for GO and KEGG enrichment analyses, 511 BP terms, 16 MF terms, and 34 KEGG pathways were identified as significantly enriched in elderly women with IS. Similarly, in elderly men with IS, 681 BP terms, 12 MF terms, and 44 KEGG pathways were identified as significantly enriched. The enriched BP and MF terms and KEGG pathways are shown in Table II and Figure 2 a-f.

Gene Sets in Elderly Women and Men with IS Identified from GSEA

The GSEA revealed 99 and 140 significantly different gene sets in elderly women and men with IS, respectively, compared to their respective healthy controls (both $p < 0.05$). Figure 3a and 3b show the gene sets identified, including reactome interleukin 4 (IL-4) and IL-13 signaling, IL-10, interleukins reactome signaling, reactome interleukin signaling KEGG cytokine, cytokine receptor interactions, the IL-18 signaling pathway, senescence, and autophagy in cancer.

DEGs in Elderly Men and Women Identified Using MCODE and Hub Genes Significantly Associated With IS

After screening 174 and 172 DEGs in elderly women and men with IS, respectively, a PPI network was constructed using Cytoscape, as shown in Figure 4a and 4b. Subsequently, MCODE was used to score and select the optimal parameters to obtain the best results and locate clusters within the network.

Five clusters were identified in elderly women with IS, with cluster 1 (102 edges and 16 node IDs) showing the highest score at 13.6 (Figure 4c). Furthermore, according to the MCODE score ranking, 16 DEGs possibly significantly associated with IS in elderly women were identified. These included: CXC motif chemokine ligand 2 (*CXCL2*), C-C motif chemokine receptor like 2 (*CCRL2*), prostaglandin-endoperoxide synthase 2 (*PTGS2*), C-C motif chemokine ligand 4 (*CCL4*), *IL-1A*, C-C motif chemokine receptor 2 (*CCR2*), *IL-6*, tumor necrosis factor (*TNF*) superfamily member 10 (*TNFSF10*), *IL1B*, interferon-gamma (*IFNG*), NLR family pyrin domain containing 3 (*NLRP3*), suppressor of cytokine signaling 3 (*SOCS3*), *TNFRSF1A, CCL20, CXCL3,* and intercellular adhesion molecule-1 (*ICAM-1*). Furthermore, when arranged in descending order

Table I. Top 10 upregulated and downregulated differentially expressed genes in elderly women and men with IS compared with their respective healthy controls.

Sex	Gene ID	Gene symbol	Adj. *p*-value	logFC	Up/Down
Elderly woman	225207_at	*PDK4*	0.017	3.102	Up
Elderly woman	206978_at	*CCR2*	0.025	2.241	Up
Elderly woman	220005_at	*P2RY13*	0.018	1.922	Up
Elderly woman	215933_s_at	*HHEX*	0.005	1.858	Up
Elderly woman	225731_at	*ANKRD50*	0.017	1.857	Up
Elderly woman	235306_at	*GIMAP8*	0.045	1.823	Up
Elderly woman	225283_at	*ARRDC4*	0.011	1.678	Up
Elderly woman	209930_s_at	*NFE2*	0.011	1.616	Up
Elderly woman	1556314_a_at	*RP11-389C8.2*	0.001	1.589	Up
Elderly woman	223377_x_at	*CISH*	0.032	1.585	Up
Elderly woman	210118_s_at	*IL-1A*	0.003	-5.493	Down
Elderly woman	205476_at	*CCL20*	0.005	-5.088	Down
Elderly woman	205207_at	*IL-6*	0.020	-5.050	Down
Elderly woman	207850_at	*CXCL3*	0.005	-4.829	Down
Elderly woman	204363_at	*F3*	0.015	-4.784	Down
Elderly woman	223484_at	*C15orf48*	0.008	-4.673	Down
Elderly woman	217996_at	*PHLDA1*	0.006	-4.061	Down
Elderly woman	36711_at	*MAFF*	0.002	-3.974	Down
Elderly woman	205767_at	*EREG*	0.011	-3.797	Down
Elderly woman	230380_at	*THAP2*	0.001	-3.220	Down
Elderly man	205476_at	*CCL20*	0.000	6.735	Up
Elderly man	207850_at	*CXCL3*	0.002	6.590	Up
Elderly man	210118_s_at	*IL-1A*	0.001	5.907	Up
Elderly man	209774_x_at	*CXCL2*	0.037	5.742	Up
Elderly man	205207_at	*IL-6*	0.024	5.333	Up
Elderly man	205067_at	*IL-1B*	0.001	5.174	Up
Elderly man	217996_at	*PHLDA1*	0.004	4.977	Up
Elderly man	205767_at	*EREG*	0.021	4.795	Up
Elderly man	204363_at	*F3*	0.023	4.763	Up
Elderly man	223484_at	*C15orf48*	0.023	4.403	Up
Elderly man	215933_s_at	*HHEX*	0.021	-2.075	Down
Elderly man	228170_at	*OLIG1*	0.023	-2.024	Down
Elderly man	204036_at	*LPAR1*	0.024	-1.944	Down
Elderly man	229934_at	*mir-223*	0.039	-1.879	Down
Elderly man	223583_at	*TNFAIP8L2*	0.048	-1.787	Down
Elderly man	219957_at	*RUFY2*	0.025	-1.776	Down
Elderly man	209930_s_at	*NFE2*	0.046	-1.739	Down
Elderly man	1556314_a_at	*RP11-389C8.2*	0.023	-1.699	Down
Elderly man	213418_at	*HSPA6*	0.046	-1.411	Down
Elderly man	207643_s_at	*TNFRSF1A*	0.048	-1.360	Down

according to their MCC scores obtained using cytoHubba, the top 10 genes in network sheet 1 included *IL-1B*, *IL-6*, *ICAM-1*, *CCL4*, *IL-1A*, *PTGS2*, *CXCL2*, *IFNG*, *CCR2*, and *TNFRSF1A* (Figure 4e).

For elderly men with IS, five clusters were identified. Cluster 1 (145 edges and 19 node IDs) showed the highest score (16.111) (Figure 4d). Based on the MCODE score ranking, 19 DEGs that may be significantly associated with IS in elderly men were identified. These included genes encoding *CD83*, *CXCL2*, cytotoxic T-lymphocyte associated protein 4 (*CTLA4*), *CCL4*, *CD69*, *IL-1A*, *CXCL3*, *CXCL1*, *NLRP3*, *IL-6*, *IL-1B*, *ICAM1,* interleukin 1 receptor antagonist (*IL1RN*), *CCRL2*, *CCL20*, TNF alpha-induced protein 3 (*TNFAIP3*),

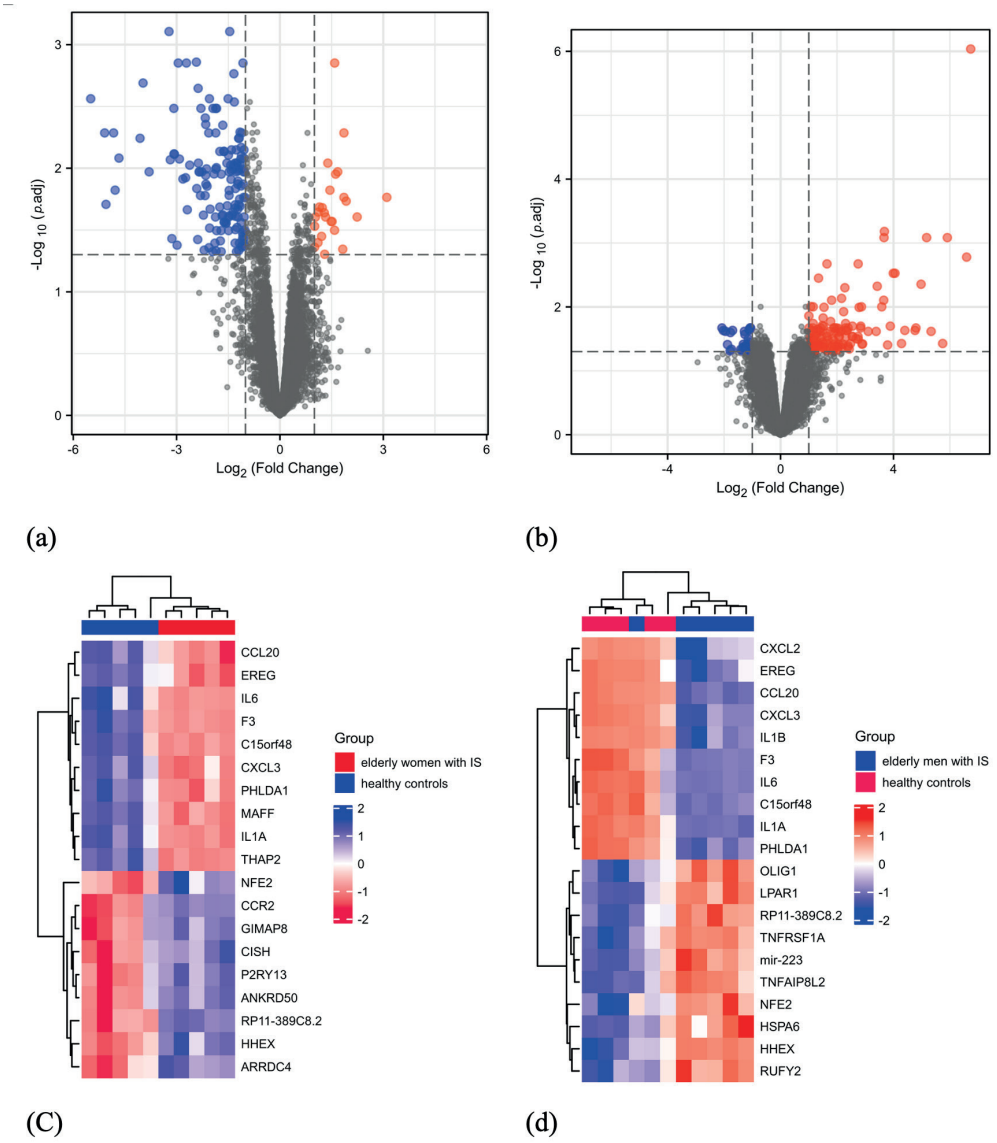

Figure 1. Volcano plot and heat map of differentially expressed genes (DEGs) in the GSE22255 dataset. The upregulated genes are indicated using red dots, whereas the downregulated genes are indicated using blue dots. The genes that did not show any significant changes are indicated using gray dots. **a**, Volcano plot showing DEGs between elderly women with IS and healthy controls. **b**, Volcano plot showing DEGs between elderly men with IS and healthy controls. **c**, Heat map of the significantly altered genes in elderly women with IS. **d**, Heat map of the significantly altered genes in elderly men with IS.

PTGS2, *TNF*, and *TNFRSF1A*. Furthermore, based on MCC scores obtained using cytoHubba, the top 10 genes on network sheet 1 were *TNF*, *IL-6*, *IL-1B*, *ICAM1*, *CCL4*, *IL-1A*, *CXCL1*, *CXCL2*, *TNFAIP3*, and *IL1RN* (Figure 4f).

Six Common Genes for Elderly Men and Women With IS

Figure 5 shows a Venn diagram of key genes in IS in elderly women and men. Six key genes (*IL-1B*, *IL-6*, *ICAM1*, *CCL4*, *IL-1A*, and *CXCL2*) were common to both elderly men and women with IS. Furthermore, the results indicated that four additional key genes (*PTGS2*, *IFNG*, *CCR2*, and *TNFRSF1A*) were exclusive to elderly women with IS, whereas four others (*TNF*, *CXCL1*, *TNFAIP3*, and *IL1RN*) were exclusive to elderly men with IS.

Predictive Value of 10 Hub Genes in Elderly Individuals With IS

The low expression levels of *IL-1B*, *IL-6*, *ICAM1*, *CCL4*, *IL-1A*, *CXCL2*, *PTGS2*, and *IFNG* and high expression levels of *CCR2* and *TNFRS-*

Table II. Biological process, molecular function, and KEGG enrichment analysis results for identified sex-specific differentially expresses genes.

GO	Sex	ID	Description	*p*-adjust
BP	Elderly woman	GO:1903039	positive regulation of leukocyte cell-cell adhesion	6.07e-07
BP	Elderly woman	GO:0022409	positive regulation of cell-cell adhesion	2.14e-06
BP	Elderly woman	GO:1903037	regulation of leukocyte cell-cell adhesion	2.14e-06
BP	Elderly woman	GO:0032743	positive regulation of interleukin-2 production	2.14e-06
BP	Elderly woman	GO:0050863	regulation of T cell activation	2.38e-06
BP	Elderly man	GO:1903039	positive regulation of leukocyte cell-cell adhesion	8.73e-11
BP	Elderly man	GO:1903037	regulation of leukocyte cell-cell adhesion	1.12e-10
BP	Elderly man	GO:0022407	regulation of cell-cell adhesion	1.71e-10
BP	Elderly man	GO:0042110	T cell activation	2.55e-10
BP	Elderly man	GO:0022409	positive regulation of cell-cell adhesion	2.55e-10
MF	Elderly woman	GO:0046935	1-phosphatidylinositol-3-kinase regulator activity	7.74e-04
MF	Elderly woman	GO:0005126	cytokine receptor binding	7.74e-04
MF	Elderly woman	GO:0035014	phosphatidylinositol 3-kinase regulator activity	0.001
MF	Elderly woman	GO:0005125	cytokine activity	0.001
MF	Elderly woman	GO:0042379	chemokine receptor binding	0.001
MF	Elderly man	GO:0005126	cytokine receptor binding	1.64e-06
MF	Elderly man	GO:0005125	cytokine activity	1.62e-04
MF	Elderly man	GO:0042379	chemokine receptor binding	0.001
MF	Elderly man	GO:0008009	chemokine activity	0.003
MF	Elderly man	GO:0002020	protease binding	0.004
KEGG	Elderly woman	hsa04668	TNF signaling pathway	1.19e-07
KEGG	Elderly woman	hsa04657	IL-17 signaling pathway	1.90e-06
KEGG	Elderly woman	hsa04064	NF-kappa B signaling pathway	4.06e-05
KEGG	Elderly woman	hsa05323	Rheumatoid arthritis	1.39e-04
KEGG	Elderly woman	hsa04061	Viral protein interaction with cytokine and cytokine receptor	1.72e-04
KEGG	Elderly man	hsa04668	TNF signaling pathway	1.23e-12
KEGG	Elderly man	hsa04657	IL-17 signaling pathway	7.66e-09
KEGG	Elderly man	hsa04064	NF-kappa B signaling pathway	1.71e-08
KEGG	Elderly man	hsa05323	Rheumatoid arthritis	8.84e-07
KEGG	Elderly man	hsa05134	Legionellosis	2.47e-06

F1A in elderly women with IS relative to healthy controls are shown in Figure 6a and 6b. Furthermore, the expression levels of *IL-1B*, *IL-6*, *ICAM1*, *CCL4*, *IL-1A*, *CXCL2*, *TNF*, *CXCL1*, *TNFAIP3*, and *IL1RN* were higher in elderly men with IS compared to healthy men, as shown in Figure 6c and 6d.

The 10 hub genes demonstrated strong discriminative ability, with an AUC a greater than 0.950. In addition, their sensitivity and specificity were 100% and 80-100%, respectively, in elderly women with IS (Table III and Figure 7 a-j). Similar observations were made for the hub genes corresponding to elderly men with IS, their sensitivity and specificity were 100% and 80-100%, respectively (Table III and Figure 7 k-t).

Discussion

It is necessary to unravel the hub genes associated with the onset and progression of IS, as such genes can function as diagnostic and therapeutic biomarkers for IS. However, studies that focus on exploring the association between the occurrence, prevalence, and impacts of IS and the sex of patients are limited. In this study, we performed bioinformatic analysis to identify sex-specific hub genes in elderly patients diagnosed with IS.

GO functional enrichment analysis showed that the identified DEGs in elderly women and elderly men with IS were mainly enriched in the leukocyte cell-cell adhesion and cytokine receptor functional categories. Furthermore, the enrichment analysis

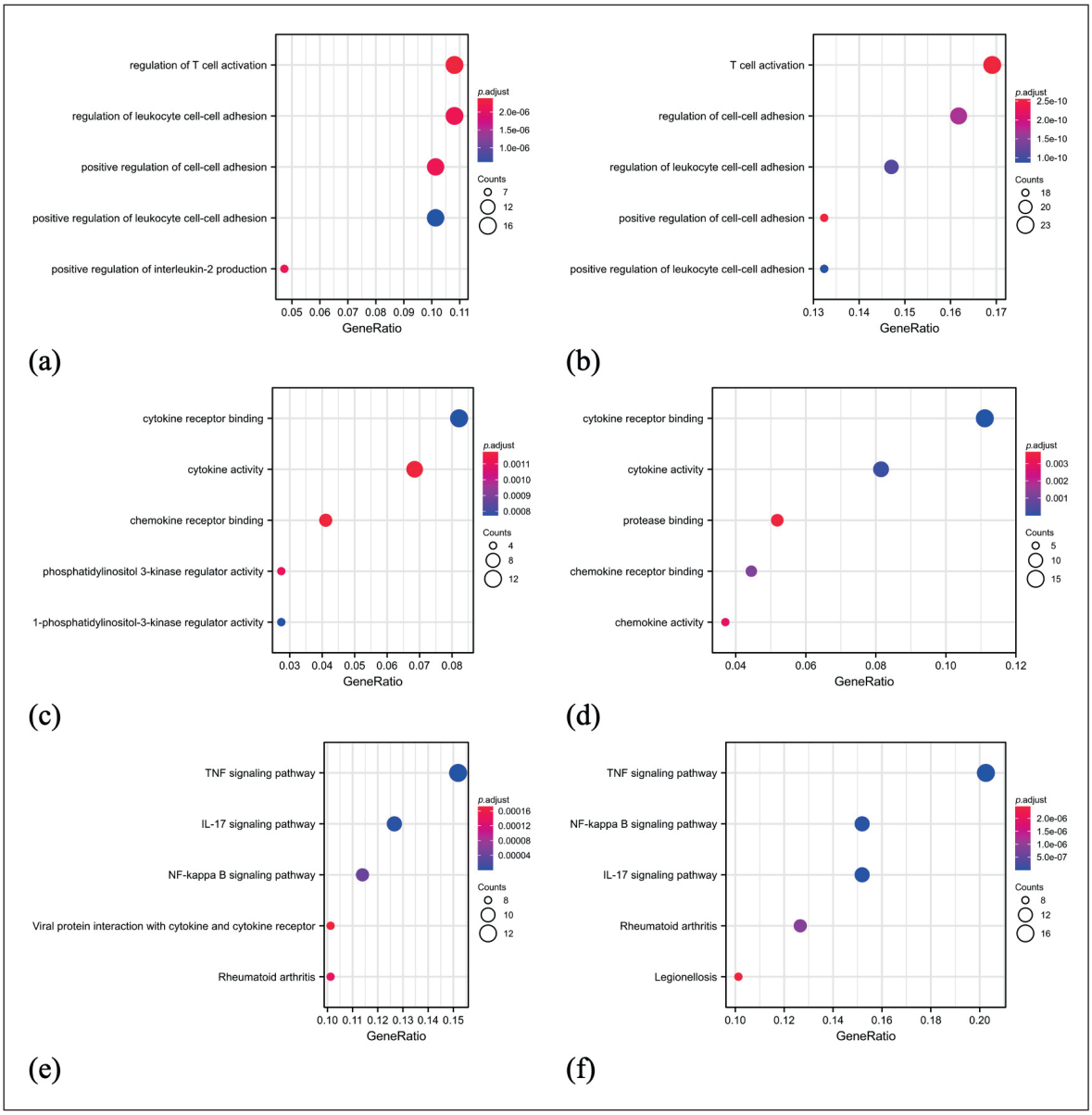

Figure 2. GO and KEGG enrichment analyses of DEGs in the GSE22255 dataset. **a**, **c**, and **e**, Top five enriched BP terms, MF terms, and KEGG pathways for elderly women with IS. **b**, **d**, and **f**, Top five enriched BP terms, MF terms, and KEGG pathways for elderly men with IS. The bubble sizes are consistent with the number of counts in the corresponding result record, which represents the total number of intersections between the input molecules and the molecules in the corresponding ID entry. The intensity of the bubble color reflects the *p*-adj value in the corresponding result record, which is the size of the *p*-value obtained after statistical inspection and correction. The abscissa after inversion represents the molecular proportion, which is consistent with the generated data in the corresponding result record.

of the KEGG pathway revealed that the DEGs in elderly women and men with IS were mainly enriched in inflammation-related pathways, such as the TNF, IL-17, and NF-κB signaling pathways. After a subsequent GSEA analysis, we identified the classes of genes differentially expressed between test groups. For example, genes involved in IL-4, 1L-13, and IL-10 reactome signaling by interleukins, KEGG cytokine and cytokine receptor interactions, IL-18 signaling pathway, senescence, and autophagy in cancer were significantly downregulated in elderly women with IS, but significantly upregulated in elderly men with IS. Reportedly, the expression of leukocyte adhesion molecules is significantly increased in patients with acute IS compared to healthy controls[19]. Cytokines, as the

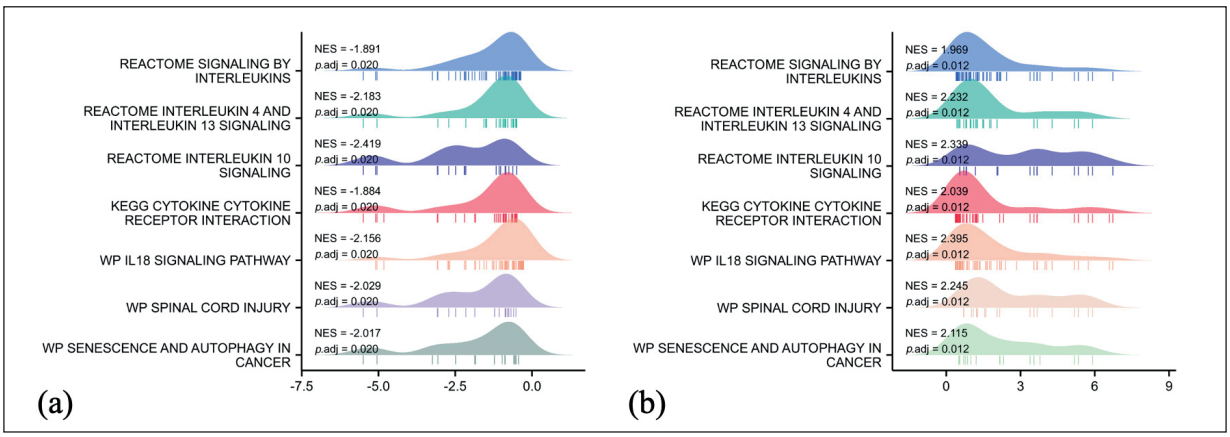

Figure 3. Gene set enrichment analysis. **a**, Gene set enrichment analysis for elderly women with IS. **b**, Gene set enrichment analysis for elderly men with IS. The Y-axis represents gene set, whereas the X-axis represents logFC of the distribution of the core molecules in each gene set. Each gene set corresponds to a peak and the shape of the peak represents logFC of the distribution of the core molecules in the gene set, and the position of the peak height indicates where the logFC values of most of the molecules in the group are concentrated at this position. When normalize enrichment score (NES) is negative, the summit of the gene set is generally on the left side of zero; however, when it is positive, the summit of the gene set is generally on the right side of zero.

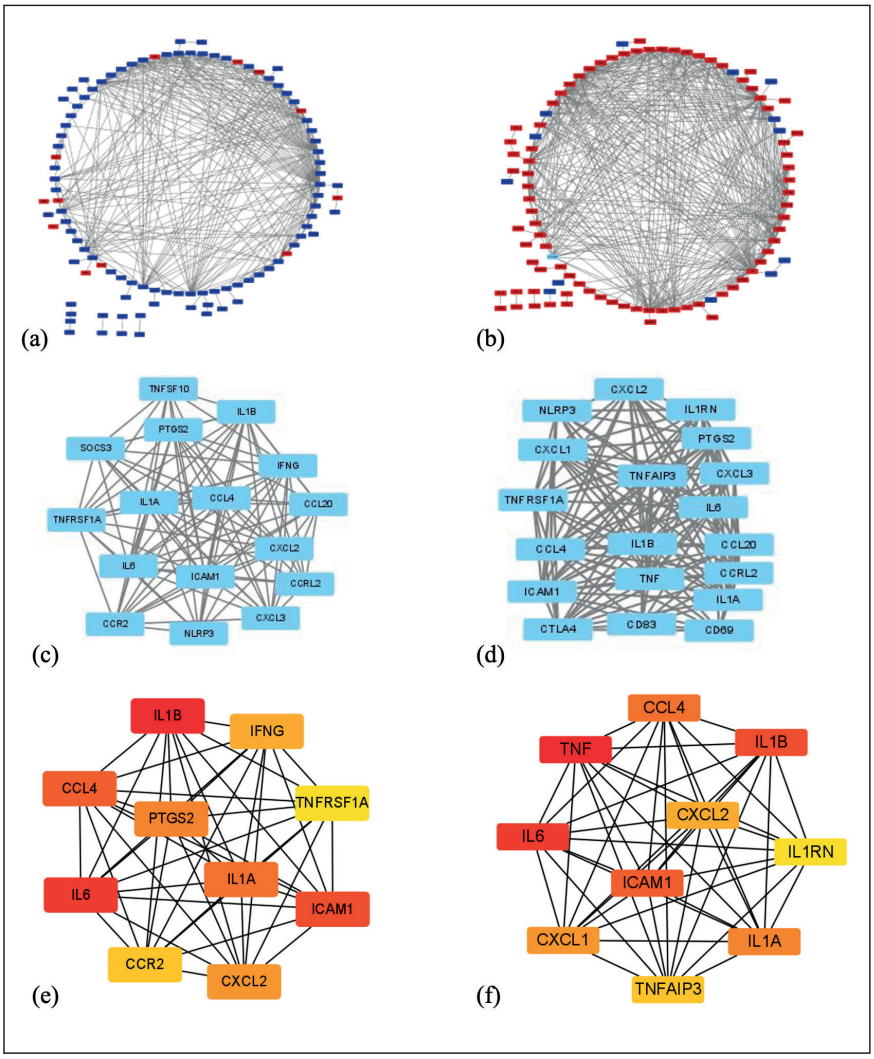

Figure 4. PPI network and hub genes of DEGs. **a**, PPI network for elderly women with IS. **b**, Results PPI network for elderly men with IS. Red represents upregulated genes and blue represents downregulated genes in the PPI networks. Each node represents a gene. The connection (edge) between nodes represents the interaction between these nodes in the PPI network. **c**, Hub genes and their interactions in elderly women with IS identified using MCODE. **d**, Hub genes and their interactions in elderly men with IS identified using MCODE. **e**, Hub genes and their interactions of elderly women with IS identified using the MCC method in cytoHubba. **f**, Hub genes and their interactions in elderly men with IS identified using the MCC method in cytoHubba. The more intense the red node color, the higher the gene score in cytoHubba.

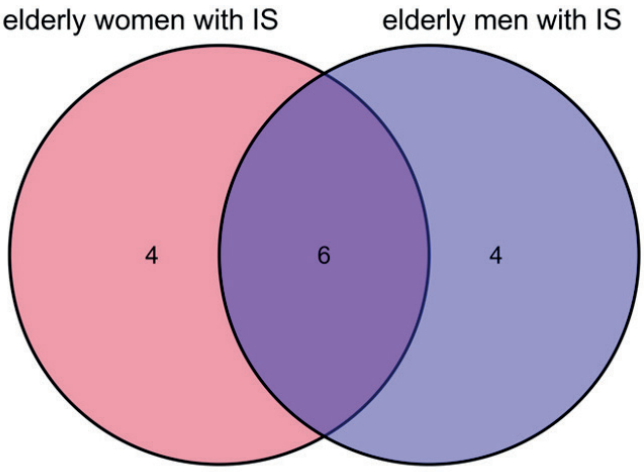

Figure 5. Venn diagram showing the hub genes common to elderly women and men with ischemic stroke.

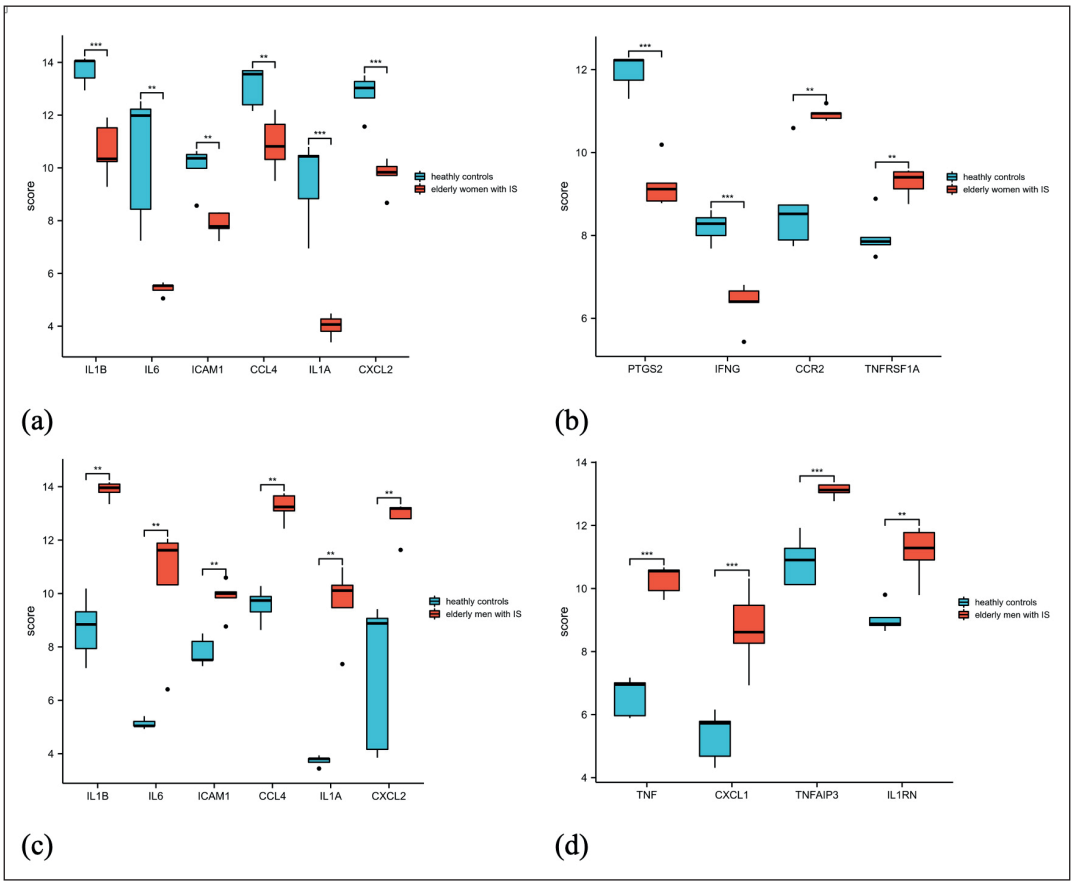

Figure 6. Hub genes in elderly patients with ischemic stroke (IS) and their healthy counterparts. **a-b**, Expression level of four hub genes in elderly women with IS and healthy controls. Red and blue indicate the expression of hub genes in elderly women with IS and healthy controls, respectively. **c-d**, Expression of four hub genes in elderly men with IS and healthy controls. Red and blue indicate the expression of hub genes in elderly men with IS and healthy controls, respectively. Significance identification: ns, $p \geq 0.05$; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$. The horizontal line in the middle of the box represents the median, whereas the upper and lower sides of the box represent the upper quartile (75th percentile) and the lower quartile (25th percentile) in box chart, respectively.

Table III. Receiver operating characteristics analysis results for elderly women and men with ischemic stroke.

	mRNA	Cut-off value	Se (%)	Sp (%)	AUC	95% CI
Elderly women with IS	IL-1B	12.423	1.000	1.000	1.000	1.000-1.000
Elderly women with IS	IL-6	6.448	1.000	1.000	1.000	1.000-1.000
Elderly women with IS	ICAM1	8.440	1.000	1.000	1.000	1.000-1.000
Elderly women with IS	CCL4	11.905	1.000	1.000	0.960	0.849-1.000
Elderly women with IS	IL-1A	5.714	1.000	1.000	1.000	1.000-1.000
Elderly women with IS	CXCL2	10.957	1.000	1.000	1.000	1.000-1.000
Elderly women with IS	PTGS2	10.743	1.000	1.000	1.000	1.000-1.000
Elderly women with IS	IFNG	7.247	1.000	1.000	1.000	1.000-1.000
Elderly women with IS	CCR2	10.680	1.000	1.000	1.000	1.000-1.000
Elderly women with IS	TNFRSF1A	8.355	1.000	0.800	0.960	0.849-1.000
Elderly men with IS	IL-1B	11.766	1.000	1.000	1.000	1.000-1.000
Elderly men with IS	IL-6	5.907	1.000	1.000	1.000	1.000-1.000
Elderly men with IS	ICAM1	8.633	1.000	1.000	1.000	1.000-1.000
Elderly men with IS	CCL4	11.354	1.000	1.000	1.000	1.000-1.000
Elderly men with IS	IL-1A	5.647	1.000	1.000	1.000	1.000-1.000
Elderly men with IS	CXCL2	10.523	1.000	1.000	1.000	1.000-1.000
Elderly men with IS	TNF	8.405	1.000	1.000	1.000	1.000-1.000
Elderly men with IS	CXCL1	6.544	1.000	1.000	1.000	1.000-1.000
Elderly men with IS	TNFAIP3	12.347	1.000	1.000	1.000	1.000-1.000
Elderly men with IS	IL-1RN	9.436	1.000	0.800	0.960	0.849-1.000

Se, sensitivity; Sp, specificity; AUC, area under the curve; CI, confidence interval.

key factors in the inflammatory mechanism, can contribute to the progression of ischemic damage in IS[20]. Inflammation has also been reported[21,22] to contribute to the onset and progression of IS by causing thrombosis, infiltration of immune cell and solute infiltration into the brain parenchyma, and destruction of the blood-brain barrier.

Our findings indicated that sex differences in elderly patients with IS were significantly associated with proinflammatory cytokines and their downstream signaling pathways. These results provide valuable information on the pathophysiological mechanisms of IS in elderly patients.

The hub genes, *IL-1B, IL-6, ICAM1, CCL4, IL-1A,* and *CXCL2*, were downregulated in elderly women with IS, but upregulated in elderly men with IS. Furthermore, *PTGS2*, *IFNG*, *CCR2*, and *TNFRSF1A* were expressed exclusively in elderly women with IS, whereas *TNF*, *CXCL1*, *TNFAIP3*, and *IL1RN* were exclusively expressed in elderly men with IS. ROC curves were constructed to further determine the diagnostic efficiency of the biomarkers. All hub genes showed acceptable diagnostic values (AUC > 0.95), indicating their potential as relevant diagnostic markers. Therefore, the hub genes identified in this study possibly play an important role in several aspects related to the initiation and progression of IS.

The IL-1 cytokine family includes IL-1A, IL-1B, and the interleukin 1 receptor antagonist (IL-1Ra). IL-1A has been shown[23] to increase neuroprotection against stroke in a male mouse model, whereas IL-1B showed[24] increased mRNA levels in rats with ischemia/reperfusion (I/R) injury. Furthermore, IL-1Ra, encoded by *IL1RN,* can competitively bind to IL1R and block IL-1A and IL-1B binding[25]. Some[26] have also observed that in male patients with IS, the concentration of IL-6 is significantly correlated with brain infarct volume, clinical outcome, and stroke severity. *TNFRSF1A*, previously called *TNFR1*, encodes a transmembrane receptor for TNF, and increased levels of *TNFR1* expression have also been reported[27] in patients with IS. *TNFAIP3*, also known as *A20*, is an important mediator of the anti-inflammatory response and a suppressor of the NF-κB pathway. The expression levels of A20, TNF, and PTGS2 have been reported[28,29] to be upregulated in mice with middle cerebral artery (MCAO). Furthermore, rats with I/R injury also showed significantly increased TNF-α and IFNG-γ mRNA levels[30]. In contrast, in a clinical

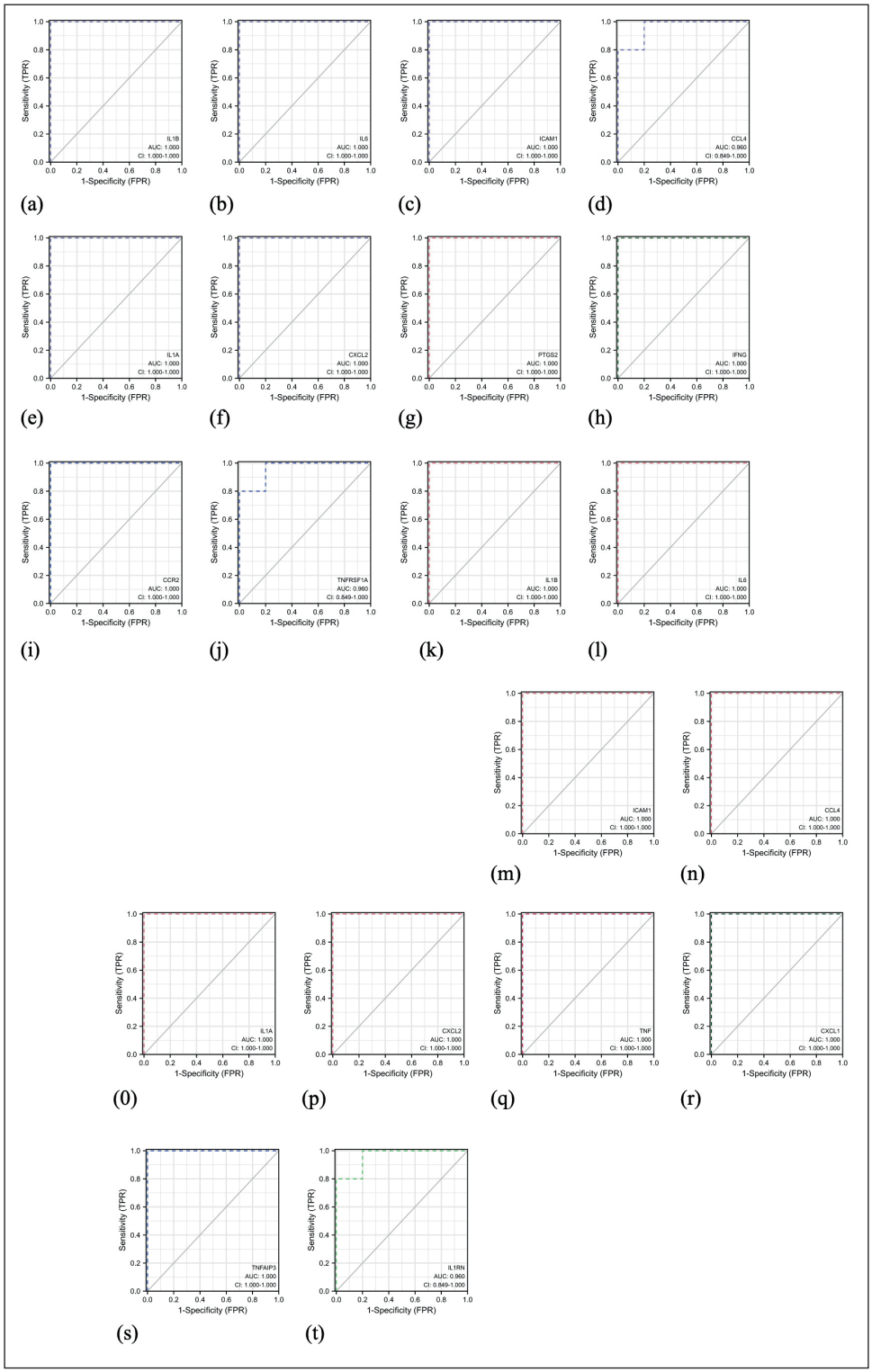

Figure 7. Receiver operating characteristic (ROC) curves for the hub genes corresponding to elderly women and men for predicting IS. **a-j**, ROC curves for the expression levels of *IL1B*, *IL6*, *ICAM1*, *CCL4*, *IL1A*, *CXCL2*, *PTGS2*, and *IFNG* (lower relative to the control) and those of *CCR2* and *TNFRSF1A* (higher relative to the control) respectively, for predicting IS in elderly women. **k-t**, ROC curves for the expression of *IL1B*, *IL6*, *ICAM1*, *CCL4*, *IL1A*, *CXCL2*, *TNF*, *CXCL1*, *TNFAIP3*, and *IL-1RN* (higher relative to the control), respectively for predicting IS in elderly men. The abscissa represents 1-specificity (FPR), whereas the ordinate represents sensitivity (TRP). The area value under the ROC curve generally varied between 0.5 and 1. The closer the AUC is to 1, the better the diagnostic effect and the higher the diagnostic accuracy.

study[31], no significant changes were observed in the expression levels of *TNF*, *IL-1A*, *IL-1B*, and *IL-1Ra* in patients with IS compared with those in the control group. In the aforementioned clinical study[31], patients were aged over 18 years of age and admitted within 48 h of symptom appearance.

Furthermore, patients underwent treatment for thrombosis. In another clinical study[7], low expression levels of *IL-1A*, *IL-1B*, *IL-6*, and *TNF* were observed in women, however, these genes were overexpressed in men. In this study, the results for *IL-1A*, *IL-1B*, and *IL-6*, but not for *TNF*, were consistent with these findings[7]. Moreover, inconsistencies in the results could be attributed to differences in the study design (pre-clinical versus clinical), sex, and age of the research subjects. However, our results are like those of a previous study[7], in which the results of the stratified analysis were interpreted considering age and sex into consideration.

ICAM-1 is a type I transmembrane protein expressed on the cell surface of numerous cell types[32]. It contributes to the recruitment of leukocytes from circulation to the inflammation site[33]. Higher serum ICAM-1 levels have been detected in men with acute IS exhibiting poor outcomes than in those exhibiting good outcomes[34], which is consistent with the results reported in our study.

Inflammatory responses in atherosclerotic plaque vulnerability and cerebral infarction are orchestrated by CC and CXC chemokines[35]. In particular, CXCL2 is involved in regulating the inflammatory response process[36], whereas CXCL1 attracts neutrophils. It has also been reported[37] that CXCL1 expression is markedly increased in an adult male C57BL/6 mouse model of thromboembolic stroke and significantly increased serum CCR2 serum levels have been observed[27] in patients with IS. CCL4, also known as macrophage inflammatory protein-1β, is located on chromosome 17 and potentially exerts chemoattraction[38]. It has also been reported[39] that the expression of CCL4 and TNF-α can predict infarction volume in acute IS.

The discrepancies observed between the results of our work and those of previous studies could be explained by the differences in the inclusion criteria, ages of the samples included, interpretation of the results based on the patients' sex, source data for patients with different ethnicities and the races, and control groups comprising healthy relatives of patients with IS. In animal experiments, young male animals are usually selected as MCAO models, whereas patients with IS in clinical practice are predominantly the elderly.

Although the underlying mechanisms involved in discriminating gene expression levels between patients of different sex are not fully understood, in recent years, several attempts have been made to bridge this knowledge gap. Sex-specific differences have been observed[40] notably in pathological samples obtained from male and female patients with IS. Thrombotic strokes are more common in men, whereas embolic strokes are more common in women[41]. A close association between sex and the location has been observed[42], wherein women are more likely to experience a stroke involving the anterior circulation, whereas men are more likely to experience lacunar strokes. IS pathophysiology is also affected by the sex of the patients[43]. For example, cell death mechanisms may differ depending on the sex of the patient[40]. In another study[43], sex differences were observed regarding autophagy and autophagy regulators in IS. However, opposing trends, such as increased Beclin1 levels in male mice with stroke and decreased Beclin1 levels in female mice with stroke, have been observed[43]. Genetic and epigenetic factors, differential activation of the cell death program, intercellular signaling pathways, and systemic immune responses have sex-specific implications in IS prevalence and progression[44].

Limitations

Although we identified some sex-specific hub genes that showed significant associations with IS in elderly patients, our study had several limitations. First, the sample size was relatively small. These gene hubs should be validated in a larger cohort to further evaluate their diagnostic and/or prognostic potential. Second, the hub genes were not further confirmed using quantitative or qualitative gene/protein expression validation techniques. The expression of the hub genes can be validated using blood samples from patients with IS. Furthermore, cellular and animal experiments should be conducted to further assess their expression levels in elderly patients with IS of different sexes, and elucidate the direct effect of the genes in elderly individuals with IS.

Conclusions

We used the GSE22255 dataset to identify sex-specific DEGs in elderly patients with IS. The top 10 sex-specific hub genes that could be used as novel markers, specifically for elderly female

and male patients with IS, were identified. These genes have the potential for clinical applications as sex-specific markers associated with IS diagnosis, treatment, and prognosis.

Ethics Approval
GEO belongs to public databases. The data from these patients have obtained ethical approval and are in the database. Our study is based on open-source data; therefore, there are no ethical issues or other conflicts of interest.

Informed Consent
Not applicable.

Data Availability
The GSE22255 data used in the study are available via GEO (https://www.ncbi.nlm.nih.gov/gds/?term=GSE22255). All the data generated or analyzed during the study are included in this published article.

Conflicts of Interest
All the authors declare that there is no conflict of interest regarding the publication of this article.

Funding
This work was supported by the National Natural Science Foundation of China [grant number 81904180]; Scientific Research Project of Hunan Provincial Health Commission [grant number B202319018677]; Scientific Research Fund of Hunan Provincial Education Department [grant number 22A0266 and 20B429]; Hunan Provincial Natural Science Foundation of China [grant numbers 2023JJ50035]; and 2022 "Disciplinary Reveal System" project of Hunan University of Chinese Medicine [grant number 22JBZ041].

Authors' Contributions
LHQ, SL and GZW contributed to the study design and critical review of the manuscript for important intellectual content. XC, JTH, LC and ZYL performed data analysis and drafted the manuscript. SL and GZW contributed to data collection and interpretation. All authors contributed to the revision of the manuscript and have read and approved the final version of the manuscript.

Acknowledgements
The authors thank all individuals who participated in this study and donated samples.

ORCID ID
Lihua Qin: 0000-0003-3458-1520
Lei Chen: 0009-0009-2477-8400
Cao Xi: 0009-0001-5416-0531
Tengjia Huang: 0009-0004-9577-6465
Ziying Li: 0000-0001-6674-4801
Sheng Li: 0000-0003-4827-8856
Guozuo Wang: 0000-0002-3741-6478

References

1) Xu S, Lu J, Shao A, Zhang JH, Zhang J. Glial Cells: Role of the Immune Response in Ischemic Stroke. Front Immunol 2020; 11: 294.
2) Ding Q, Liu S, Yao Y, Liu H, Cai T, Han L. Global, Regional, and National Burden of Ischemic Stroke, 1990-2019. Neurology 2022; 98: e279-e290.
3) Musuka TD, Wilton SB, Traboulsi M, Hill MD. Diagnosis and management of acute ischemic stroke: speed is critical. CMAJ 2015; 187: 887-893.
4) Fang MC, Cutler DM, Rosen AB. Trends in thrombolytic use for ischemic stroke in the United States. J Hosp Med 2010; 5: 406-409.
5) Han B, Zhang Y, Zhang Y, Bai Y, Chen X, Huang R, Wu F, Leng S, Chao J, Zhang JH, Hu G, Yao H. Novel insight into circular RNA HECTD1 in astrocyte activation via autophagy by targeting MIR142-TIPARP: implications for cerebral ischemic stroke. Autophagy 2018; 14: 1164-1184.
6) Gibson CL. Cerebral ischemic stroke: is gender important? J Cereb Blood Flow Metab 2013; 33: 1355-1361.
7) Li WX, Dai SX, Wang Q, Guo YC, Hong Y, Zheng JJ, Liu JQ, Liu D, Li GH, Huang JF. Integrated analysis of ischemic stroke datasets revealed sex and age difference in anti-stroke targets. PeerJ 2016; 4: e2470.
8) Koh S, Gordon AS, Wienberg C, Sood SO, Morley S, Burke DM. Stroke experiences in weblogs: a feasibility study of sex differences. J Med Internet Res 2014; 16: e84.
9) Zhu W, Nan Y, Wang S, Liu W. Bioinformatics Analysis of Gene Expression Profiles of Sex Differences in Ischemic Stroke. Biomed Res Int 2019; 2019: 2478453.
10) Simpkins AN, Janowski M, Oz HS, Roberts J, Bix G, Doré S, Stowe AM. Biomarker Application for Precision Medicine in Stroke. Transl Stroke Res 2020; 11: 615-627.
11) Huang WT, Lin TS, Wu JY, Hong JM, Chen YL, Qiu FN. Evaluation of miR-429 as a novel serum biomarker for pancreatic ductal adenocarcinoma and analysis its tumor suppressor function and target genes. Eur Rev Med Pharmacol Sci 2022; 26: 4638-4653.
12) Krug T, Gabriel JP, Taipa R, Fonseca BV, Domingues-Montanari S, Fernandez-Cadenas I, Manso H, Gouveia LO, Sobral J, Albergaria I, Gaspar G, Jiménez-Conde J, Rabionet R, Ferro JM, Montaner J, Vicente AM, Silva MR, Matos I, Lopes G, Oliveira SA. TTC7B emerges as a novel risk factor for ischemic stroke through the convergence of several genome-wide approaches. J Cereb Blood Flow Metab 2012; 32: 1061-1072.
13) Campbell-Furtick M, Moore BJ, Overton TL, Laureano Phillips J, Simon KJ, Gandhi RR, Duane TM, Shafi S. Post-trauma mortality increase at age 60: a cutoff for defining elderly? Am J Surg 2016; 212: 781-785.

14) Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 2016; 32: 2847-2849.

15) Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012; 16: 284-287.

16) Walter W, Sánchez-Cabo F, Ricote M. GOplot: an R package for visually combining expression data with functional analysis. Bioinformatics 2015; 31: 2912-2914.

17) Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005; 102: 15545-15550.

18) Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ, Mering CV. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 2019; 47: D607-D613.

19) Tsai NW, Chang WN, Shaw CF, Jan CR, Huang CR, Chen SD, Chuang YC, Lee LH, Lu CH. The value of leukocyte adhesion molecules in patients after ischemic stroke. J Neurol 2009; 256: 1296-1302.

20) Doll DN, Barr TL, Simpkins JW. Cytokines: their role in stroke and potential use as biomarkers and therapeutic targets. Aging Dis 2014; 5: 294-306.

21) Pei J, You X, Fu Q. Inflammation in the pathogenesis of ischemic stroke. Front Biosci (Landmark Ed) 2015; 20: 772-783.

22) Shekhar S, Cunningham MW, Pabbidi MR, Wang S, Booz GW, Fan F. Targeting vascular inflammation in ischemic stroke: Recent developments on novel immunomodulatory approaches. Eur J Pharmacol 2018; 833: 531-544.

23) Salmeron KE, Maniskas ME, Edwards DN, Wong R, Rajkovic I, Trout A, Rahman AA, Hamilton S, Fraser JF, Pinteaux E, Bix GJ. Interleukin 1 alpha administration is neuroprotective and neuro-restorative following experimental ischemic stroke. J Neuroinflammation 2019; 16: 222.

24) Yang T, Feng C, Wang D, Qu Y, Yang Y, Wang Y, Sun Z. Neuroprotective and Anti-inflammatory Effect of Tangeretin Against Cerebral Ischemia-Reperfusion Injury in Rats. Inflammation 2020; 43: 2332-2343.

25) Dinarello CA. Immunological and inflammatory functions of the interleukin-1 family. Annu Rev Immunol 2009; 27: 519-550.

26) Smith CJ, Emsley HC, Gavin CM, Georgiou RF, Vail A, Barberan EM, del Zoppo GJ, Hallenbeck JM, Rothwell NJ, Hopkins SJ, Tyrrell PJ. Peak plasma interleukin-6 and other peripheral markers of inflammation in the first week of ischaemic stroke correlate with brain infarct volume, stroke severity and long-term outcome. BMC Neurol 2004; 4: 2.

27) Hansen RB, Laursen CCH, Nawaz N, Madsen JS, Nielsen HH, Kruuse C, Møller A, Degn M, Lambertsen KL. Leukocyte TNFR1 and TNFR2 Expression Contributes to the Peripheral Immune Response in Cases with Ischemic Stroke. Cells 2021; 10: 861.

28) Mc Guire C, Rahman M, Schwaninger M, Beyaert R, van Loo G. The ubiquitin editing enzyme A20 (TNFAIP3) is upregulated during permanent middle cerebral artery occlusion but does not influence disease outcome. Cell Death Dis 2013; 4: e531.

29) Zhou Z, Lu C, Meng S, Dun L, Yin N, An H, Xu H, Liu G, Cai Y. Silencing of PTGS2 exerts promoting effects on angiogenesis endothelial progenitor cells in mice with ischemic stroke via repression of the NF-κB signaling pathway. J Cell Physiol 2019; 234: 23448-23460.

30) Yang T, Feng C, Wang D, Qu Y, Yang Y, Wang Y, Sun Z. Neuroprotective and Anti-inflammatory Effect of Tangeretin Against Cerebral Ischemia-Reperfusion Injury in Rats. Inflammation 2020; 43: 2332-2343.

31) Clausen BH, Wirenfeldt M, Høgedal SS, Frich LH, Nielsen HH, Schrøder HD, Østergaard K, Finsen B, Kristensen BW, Lambertsen KL. Characterization of the TNF and IL-1 systems in human brain and blood after ischemic stroke. Acta Neuropathol Commun 2020; 8: 81.

32) Lawson C, Wolf S. ICAM-1 signaling in endothelial cells. Pharmacol Rep 2009; 61: 22-32.

33) Yang RC, Qu XY, Xiao SY, Li L, Xu BJ, Fu JY, Lv YJ, Amjad N, Tan C, Kim KS, Chen HC, Wang XR. Meningitic Escherichia coli-induced upregulation of PDGF-B and ICAM-1 aggravates blood-brain barrier disruption and neuroinflammatory response. J Neuroinflammation 2019; 16: 101.

34) Wang L, Chen Y, Feng D, Wang X. Serum ICAM-1 as a Predictor of Prognosis in Patients with Acute Ischemic Stroke. Biomed Res Int 2021; 2021: 5539304.

35) Mirabelli-Badenier M, Braunersreuther V, Viviani GL, Dallegri F, Quercioli A, Veneselli E, Mach F, Montecucco F. CC and CXC chemokines are pivotal mediators of cerebral injury in ischaemic stroke. Thromb Haemost 2011; 105: 409-420.

36) Zhong T, Guo L, Jiang CY, Peng XY, Zou J. Role of CXCL2 gene in LPS-induced inflammation of bronchial epithelial cells. Chin J Immunol 2021; 37: 1804-1808.

37) Stamatovic SM, Phillips CM, Keep RF, Andjelkovic AV. A novel approach to treatment of thromboembolic stroke in mice: Redirecting neutrophils toward a peripherally implanted CXCL1-soaked sponge. Exp Neurol 2020; 330: 113336.

38) Kobayashi Y, Konno Y, Kanda A, Yamada Y, Yasuba H, Sakata Y, Fukuchi M, Tomoda K, Iwai H, Ueki S. Critical role of CCL4 in eosinophil recruitment into the airway. Clin Exp Allergy 2019; 49: 853-860.

39) Martha SR, Cheng Q, Fraser JF, Gong L, Collier LA, Davis SM, Lukins D, Alhajeri A, Grupke S,

Pennypacker KR. Expression of Cytokines and Chemokines as Predictors of Stroke Outcomes in Acute Ischemic Stroke. Front Neurol 2020; 10: 1391.

40) Gibson CL, Attwood L. The impact of gender on stroke pathology and treatment. Neurosci Biobehav Rev 2016; 67: 119-124.

41) Kolominsky-Rabas PL, Weber M, Gefeller O, Neundoerfer B, Heuschmann PU. Epidemiology of ischemic stroke subtypes according to TOAST criteria: incidence, recurrence, and long-term survival in ischemic stroke subtypes: a population-based study. Stroke 2001; 32: 2735-2740.

42) Wiszniewska M, Niewada M, Czlonkowska A. Sex differences in risk factor distribution, severity, and outcome of ischemic stroke. Acta Clin Croat 2011; 50: 21-28.

43) Patrizz AN, Moruno-Manchon JF, O'Keefe LM, Doran SJ, Patel AR, Venna VR, Tsvetkov AS, Li J, McCullough LD. Sex-Specific Differences in Autophagic Responses to Experimental Ischemic Stroke. Cells 2021; 10: 1825.

44) Spychala MS, Honarpisheh P, McCullough LD. Sex differences in neuroinflammation and neuroprotection in ischemic stroke. J Neurosci Res 2017; 95: 462-471.