Хиймэл оюунд суурилсан тархины нийт эзлэхүүний хэмжилт ба радиологичийн тайланд бүртгэгдсэн атрофийн зэрэглэлийн хамаарлыг судалсан нь

(ЧШХШ2-той болон хяналтын бүлгийн тоон CPT судалгаа)

А.Цэрэнсугир[1,2], Б.Оюунтөгс[3], Д.Гончигсүрэн[4], Д.Түвшинжаргал[1], А.Товуудорж[5], Г.Мөнгөнбагана[6], Д.Мөнхбаатар[1], Х.Дэлгэрдалай[1]

[1]АШУҮИС, АУС, Радиологийн тэнхим
[2]АШУҮИС, Ахисан түвшний сургууль
[3]АШУҮИС, АУС, Дотоод шүүрлийн тэнхим
[4]Грандмед эмнэлэг, Радиологийн тасаг
[5]АШУҮИС, АУС, Мэдрэл судлалын тэнхим
[6]Бриллиант эмнэлэг, Радиологийн тасаг
Цахим шуудан: amd26f002@gt.mnums.edu.mn, Утас: 99703955

Түлхүүр үг:

Тархины атрофи
Автомат сегментаци
Тархины эзлэхүүн
Морфометр
Нейродүрслэл

Товч утга:

Үндэслэл: Тархины бор гадаргын ерөнхий атрофийн (Global Cortical Atrophy, GCA) үнэлгээ нь соронзон резонанст томографи (CPT)-д кортикал атрофийг хурдан үнэлэхэд өргөн хэрэглэгддэг боловч тодорхой хэмжээнд субъектив шинжтэй. Харин автомат тоон морфометрын үзүүлэлтүүдийг объектив байдлаар хэмжиж, харааны үнэлгээг нэмэлт мэдээллээр дэмжих боломж олгодог. Манай улсад тархины атрофийн харааны үнэлгээг автомат тоон эзлэхүүний хэмжилттэй харьцуулан судалсан мэдээлэл хязгаарлагдмал хэвээр байгаа нь энэхүү судалгааг хийх үндэслэл боллоо. **Зорилго:** Радиологичийн тайланд бүртгэгдсэн атрофийн зэрэглэл болон Neurophet AQUA-аар хэмжсэн тархины нийт эзлэхүүний хамаарал, ангиллын тохирлыг үнэлж, ЧШХШ2-той болон хяналтын бүлгийн эзлэхүүнийг харьцуулах. **Арга, аргачлал:** Судалгааг нэг төвт, аналитик судалгааны агшингийн загвараар хийж гүйцэтгэв.Судалгаанд ≥40 настай 367 оролцогч (ЧШХШ2, n=144; хяналт, n=223) хамрагдав. Neurophet AQUA v3.1 программаар 3D T1 дүрслэлийг боловсруулж, гавлын дотоод эзлэхүүнд (ICV) тохируулсан тархины нийт эзлэхүүнийг хэмжив. Радиологичийн тайланд бүртгэгдсэн атрофийг 0–3 зэрэглэлээр кодлов. Зэрэглэл хоорондын ялгааг Welch ANOVA, ялган ангилах чадварыг ROC, ангиллын тохирлыг stratified five-fold cross-validation болон Cohen’s κ-гаар үнэлэв. Бүлэг хоорондын эзлэхүүний ялгааг нас, хүйсээр тохируулсан ANCOVA-аар тодорхойлов. **Үр дүн:** Атрофийн 0–3-р зэрэглэлд тархины нийт эзлэхүүний дундаж тус тус 1122.12, 1068.50, 1034.76, 968.33 мл болж буурсан (Welch $F(3,60.04)=207.30$; $P<0.001$). Хөрш зэрэглэлүүдийн ROC AUC тус тус 0.892, 0.897, 1.000, босго утгууд 1099.86, 1056.17, 995.57 мл байв. Five-fold cross-validation-д бүрэн тохирол 78.7% (289/367), unweighted $\kappa=0.684$, quadratic-weighted $\kappa=0.856$ байлаа. ЧШХШ2 бүлгийн тархины нийт эзлэхүүн хяналтын бүлгээс 11.98 мл-ээр бага байсан (95% CI −22.67–−1.30; $P=0.028$). Харин атрофийн зэрэглэлтэй холбоотой олон ангилалт логистик регрессийн шинжилгээнд ЧШХШ2-ийн нийт нөлөө статистикийн ач холбогдолгүй байв ($P=0.059$). **Дүгнэлт:** Радиологичийн тайланд бүртгэгдсэн атрофийн зэрэглэл нэмэгдэх тусам тархины нийт эзлэхүүн буурч, эзлэхүүнд суурилсан ангилал өндөр жинлэсэн тохирол үзүүлэв (quadratic-weighted $\kappa=0.856$). ЧШХШ2 бүлгийн тархины нийт эзлэхүүн хяналтын бүлгээс бага байв. Автомат тоон морфометр нь радиологичийн харааны үнэлгээг тоон мэдээллээр дэмжих боломжтой.

Үндэслэл: Тархины атрофи нь насжилтын явцад тодорхой хэмжээгээр илэрч болох боловч нейродегенератив, судасны болон бодисын солилцооны олон эмгэгийн үед илүү тод ажиглагддаг бүтцийн өөрчлөлт юм. Соронзон резонанст томографийн (СРТ) шинжилгээнд тархины бор гадаргын ерөнхий атрофийг үнэлэхэд Global Cortical Atrophy (GCA) үнэлгээний системийг өргөн хэрэглэдэг. GCA нь тархины ховилын өргөсөл болон нугалааны эзлэхүүний бууралтад үндэслэн атрофийг 0–3 зэрэглэлээр үнэлдэг. [1−3] Энэхүү харааны үнэлгээ нь хэрэглэхэд хялбар, хурдан боловч радиологичийн туршлагаас тодорхой хэмжээгээр хамаарч, бага зэргийн бүтцийн өөрчлөлтийг ялган үнэлэхэд хязгаарлагдмал байж болдог талаар цөөнгүй судлаачид тэмдэглэсэн байна. [2,4]

Сүүлийн жилүүдэд хиймэл оюунд суурилсан автомат тоон СРТ шинжилгээ хөгжсөнөөр тархины эзлэхүүнийг объектив хэмжих боломж нэмэгдсэн.[5] Зарим судаачид харааны атрофийн зэрэглэл нэмэгдэх тусам автомат хэмжилтээр тодорхойлсон тархины эзлэхүүн бага байх хандлага ажиглагдсан болохыг дурдсан байна.[6−8] Тухайлбал, Persson нар GCA-frontal зэрэглэл ба духны дэлбэнгийн автомат эзлэхүүний хооронд урвуу хамаарал, Nassanga нар GCA зэрэглэл ба тархины нийт эзлэхүүний хооронд дунд зэргийн урвуу хамаарал байгааг мэдээлсэн байна.[6,7] Гэвч эдгээр хамаарлын хүч нь үнэлгээний арга, программ хангамж болон судалгааны хүн амаас шалтгаалан ялгаатай болохыг онцолжээ.[4,5]

Neurophet AQUA программын автомат сегментацийн нарийвчлал болон хэмжилтийн давтагдах чанарыг хэд хэдэн судалгаагаар үнэлсэн байна. [9,10] Гэсэн хэдий ч тус программаар хэмжсэн тархины нийт эзлэхүүн болон радиологичийн тайланд бүртгэгдсэн атрофийн зэрэглэлийн хоорондын хамаарлыг шууд үнэлсэн эмнэлзүйн судалгааны мэдээлэл хязгаарлагдмал байна. Түүнчлэн чихрийн шижин хэв шинж 2 (ЧШХШ2) нь тархины эзлэхүүний бууралттай холбоотой болохыг цөөнгүй судаачдын үр дүнгүүд харуулжээ. [11, 12] Иймээс манай улсад тархины атрофийг харааны үнэлгээ болон автомат тоон эзлэхүүний хэмжилтээр харьцуулан судалсан мэдээлэл хязгаарлагдмал хэвээр байгаа нь энэхүү судалгааг хийх үндэслэл боллоо.

Зорилго: Радиологичийн тайланд бүртгэгдсэн атрофийн зэрэглэл болон Neurophet AQUA-аар хэмжсэн тархины нийт эзлэхүүний хамаарал, ангиллын тохирлыг үнэлж, тархины нийт эзлэхүүнийг ЧШХШ2-той болон хяналтын бүлгийн хооронд харьцуулах.

Арга, аргачлал: Энэхүү судалгааг аналитик судалгааны агшингийн загвараар АШУҮИС-ийн Монгол–Японы эмнэлэгт 2019 оны 6 дугаар сарын 1-нээс 2026 оны 6 дугаар сарын 1-ний хооронд хийгдсэн тархины СРТ шинжилгээ болон холбогдох эмнэлзүйн мэдээллийг ашиглан хийж гүйцэтгэв. Судалгаанд 40 ба түүнээс дээш настай, 3D T1 болон T2-FLAIR дүрслэл бүрэн, автомат боловсруулалтын шаардлага хангасан оролцогчдыг хамруулав. ЧШХШ2 бүлгийг KARTE системд бүртгэгдсэн оношоор, хяналтын бүлгийг чихрийн шижингийн онош бүртгэгдээгүй байдлаар тодорхойлов. Тархины морфологид мэдэгдэхүйц нөлөөлөх эмгэг, дүрсийн чанар хангалтгүй болон ашиглах боломжтой автомат тоон үр дүн үүсээгүй тохиолдлуудыг хасав. Эцсийн шинжилгээнд 367 оролцогч (ЧШХШ2, n=144; хяналт, n=223) хамрагдав.

СРТ шинжилгээг Canon Vantage Elan 1.5 T төхөөрөмжөөр хийж, 3D T1 дүрслэлийг Neurophet AQUA v3.1 (Neurophet Inc., Seoul, Republic of Korea) программаар боловсруулав. Статистик шинжилгээнд гавлын дотоод эзлэхүүнд (ICV) тохируулсан тархины нийт эзлэхүүнийг ашиглав. [9,10] Радиологичдын өдөр тутмын тайланд бүртгэгдсэн атрофийн зэргийг 0 (атрофи тэмдэглээгүй), 1 (хөнгөн), 2 (дунд), 3 (хүнд) гэж кодлов.

Атрофийн зэрэглэлүүдийн хооронд тархины нийт эзлэхүүний ялгааг Welch ANOVA-аар, хөрш зэрэглэлүүдийг (0–1, 1–2, 2–3) ялгах чадварыг ROC шинжилгээгээр үнэлж, Youden’s index ашиглан босго утгуудыг тодорхойлов. Эзлэхүүнд суурилсан ангиллын тохирлыг stratified five-fold cross-validation-аар шалгаж, exact agreement, unweighted болон quadratic-weighted Cohen’s κ коэффициентийг тооцов. ЧШХШ2 болон хяналтын бүлгийн тархины нийт эзлэхүүний ялгааг нас, хүйсээр тохируулсан ANCOVA шинжилгээгээр үнэлэв. Статистикийн ач холбогдлын түвшинг $P<0.05$ гэж тогтоов.

Судалгааг АШУҮИС-ийн Судалгааны ёс зүйн хяналтын хорооны зөвшөөрлөөр гүйцэтгэв (№2024/3-07; 2024 оны 9 дүгээр сарын 27). Судалгааны мэдээллийг хувь хүнийг шууд таних боломжгүй байдлаар кодчилон боловсруулав.

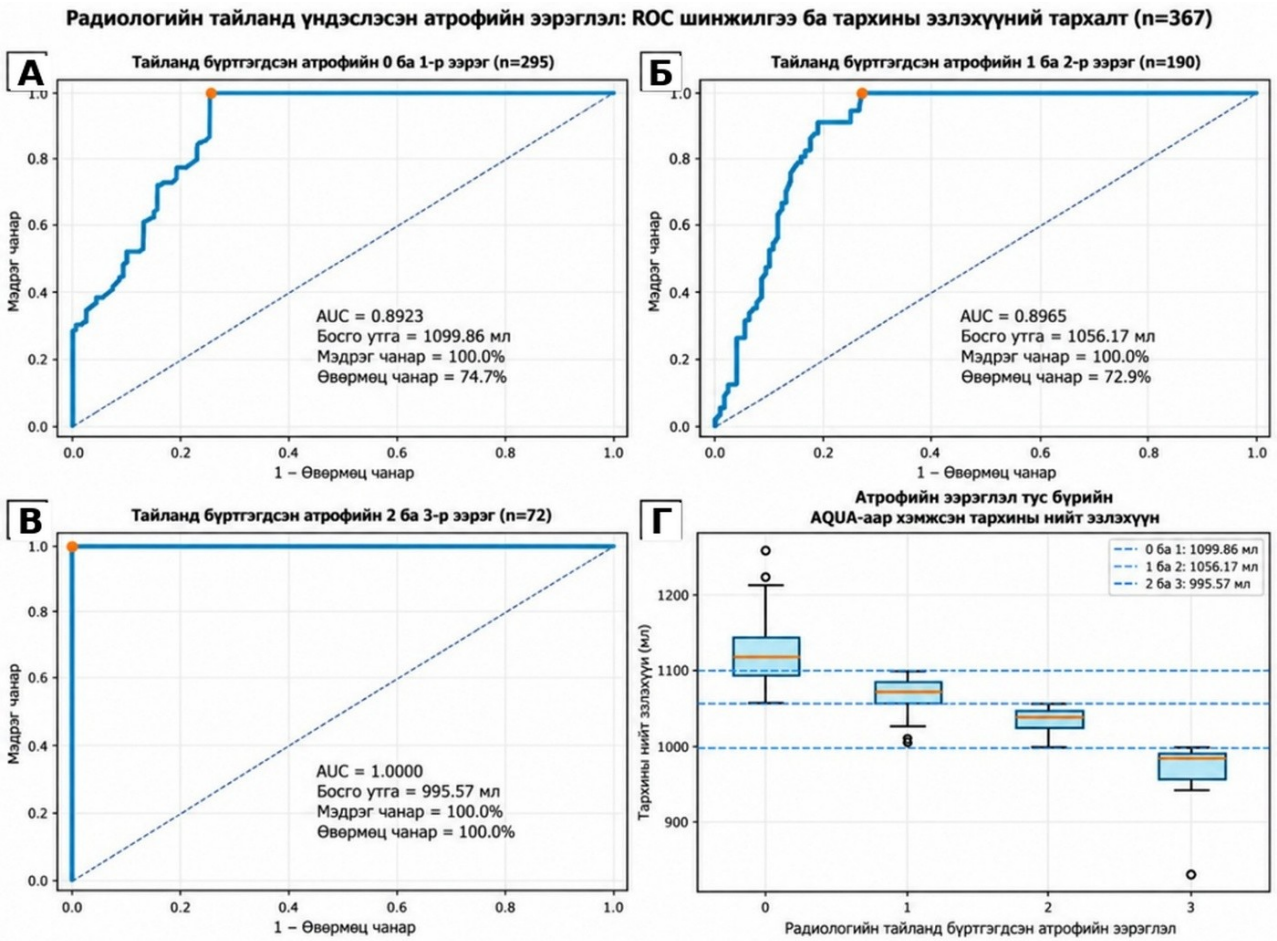
Үр дүн: Эцсийн шинжилгээнд 367 оролцогч хамрагдсанаас 144 (39.2%) нь ЧШХШ2, 223 (60.8%) нь хяналтын бүлэгт хамаарч байв. Радиологичийн тайланд бүртгэгдсэн атрофийн зэрэглэл нэмэгдэх тусам Neurophet AQUA-аар хэмжсэн, гавлын дотоод эзлэхүүнд (ICV) тохируулсан тархины нийт эзлэхүүн буурч байв. Атрофийн 0, 1, 2, 3-р зэрэглэлд харгалзах дундаж эзлэхүүн тус тус 1122.12, 1068.50, 1034.76, 968.33 мл байсан бөгөөд зэрэглэлүүдийн хооронд статистикийн ач холбогдол бүхий ялгаа илэрсэн Welch ANOVA: $F(3,60.04)=207.30$; $P<0.001$; (Хүснэгт 1).

Хүснэгт 1. Атрофийн зэрэглэлээр тархины нийт эзлэхүүн болон судалгааны бүлгүүдийн тархалтыг харьцуулсан үр дүн

Атрофийн зэрэглэл	Нийт, n (%)	Тархины нийт эзлэхүүн, мл	ЧШХШ2 бүлэг, n (%)	Хяналтын бүлэг, n (%)
0	162 (44.1)	1122.12±39.16	63 (43.8)	99 (44.4)
1	133 (36.2)	1068.50±21.75	46 (31.9)	87 (39.0)
2	57 (15.5)	1034.76±15.63	24 (16.7)	33 (14.8)
3	15 (4.1)	968.33±43.11	11 (7.6)	4 (1.8)
Нийт	367 (100.0)	-	144 (100.0)	223 (100.0)

Тайлбар: Тархины нийт эзлэхүүнийг дундаж±стандарт хазайлтаар, бусад үзүүлэлтийг n (%) хэлбэрээр илэрхийлэв.

Тархины нийт эзлэхүүний хөрш атрофийн зэрэглэлүүдийг ялгах чадварыг ROC шинжилгээгээр үнэлэхэд 0 ба 1-р зэрэглэлд AUC=0.892, мэдрэг чанар 100%, өвөрмөц чанар 74.7% (босго утга 1099.86 мл); 1 ба 2-р зэрэглэлд AUC=0.897, мэдрэг чанар 100%, өвөрмөц чанар 72.9% (1056.17 мл); 2 ба 3-р зэрэглэлд AUC=1.000, мэдрэг чанар 100%, өвөрмөц чанар 100% (995.57 мл) байв (Зураг 1).



Зураг 1. Радиологичийн тайланд бүртгэгдсэн атрофийн зэрэглэлийг ялгах ROC шинжилгээ болон ICV-д тохируулсан тархины нийт эзлэхүүний тархалтыг тодорхойлсон үр дүн.

Тайлбар: А–В: Атрофийн 0 ба 1, 1 ба 2, 2 ба 3-р зэрэглэлийн ROC муруй; Г: Атрофийн зэрэглэл тус бүрийн ICV-д тохируулсан тархины нийт эзлэхүүний boxplot. Тасархай хэвтээ шугамууд нь ROC шинжилгээгээр тухайн түүвэрт тодорхойлсон босго утгуудыг илэрхийлнэ. AUC=ROC муруйн доорх талбай; ICV=гавлын дотоод эзлэхүүн Stratified five-fold cross-validation шинжилгээнд тархины нийт эзлэхүүнд суурилсан ангилал радиологичийн тайланд бүртгэгдсэн атрофийн зэрэглэлтэй 78.7% (289/367) бүрэн тохирч, unweighted Cohen's κ=0.684, quadratic-weighted κ=0.856 байв.

Хэлцэмж: Энэхүү судалгаагаар радиологичийн тайланд бүртгэгдсэн атрофийн зэрэглэл нэмэгдэх тусам Neurophet AQUA-аар хэмжсэн, гавлын дотоод эзлэхүүнд (ICV) тохируулсан тархины нийт эзлэхүүн шат дараалан буурч байгааг тогтоов. Атрофийн 0–3-р зэрэглэлүүдийн хооронд тархины нийт эзлэхүүн статистик ач холбогдол бүхий ялгаатай байсан (Welch ANOVA: $F(3,60.04)=207.30$; $P<0.001$).

Атрофийн харааны үнэлгээ болон автомат эзлэхүүний хэмжилтийн хоорондын хамаарлыг хэд хэдэн судалгаанд үнэлсэн. Persson нар автомат аргаар хэмжсэн духны дэлбэнгийн эзлэхүүн болон GCA-frontal зэрэглэлийн хооронд урвуу хамаарал ($r=-0.62$), Nassanga нар GCA зэрэглэл болон автомат аргаар тодорхойлсон тархины нийт эзлэхүүний перцентилийн хооронд урвуу хамаарал ($\rho=-0.56$; $P<0.001$) илрүүлсэн.[6,7] Wittens, Min болон Persson нарын судалгаанд чамархайн дэлбэнгийн дотор

хэсгийн атрофийн (MTA) харааны үнэлгээ болон автомат эзлэхүүний хэмжилтийн хооронд хамаарал тогтоогдсон.[13-15] Эдгээр судалгаанд стандартчилсан GCA, MTA үнэлгээ ашигласан бол бидний судалгаанд өдөр тутмын радиологийн тайланд бүртгэгдсэн атрофийн зэрэглэлийг хэрэглэсэн тул арга зүйн хувьд бүрэн адил биш юм.

Тархины нийт эзлэхүүнээр атрофийн хөрш зэрэглэлүүдийг ялгах чадварыг ROC шинжилгээгээр үнэлэхэд 0–1, 1–2, 2–3-р зэрэглэлийн AUC тус тус 0.892, 0.897, 1.000, босго утгууд 1099.86, 1056.17, 995.57 мл байв. Харин cross-validation шинжилгээнд радиологичийн тайлангийн зэрэглэлтэй бүрэн тохирол 78.7%, unweighted Cohen's κ=0.684, quadratic-weighted κ=0.856 гарсан. ROC шинжилгээгээр зэрэглэлүүдийг ялгах чадварыг, дотоод баталгаажуулалтаар ангиллын тохирлыг тус тус үнэлсэн. Гуравдугаар зэргийн атрофитой оролцогчдын тоо цөөн (n=15) байсан тул AUC=1.000 гэсэн үр дүнгийн тогтвортой байдлыг бүрэн үнэлэх боломжгүй байв.

Харааны үнэлгээ нь радиологичийн туршлагаас, автомат эзлэхүүний хэмжилт нь CPT-ийн дүрсний чанар, дүрслэлийн протокол, төхөөрөмж болон сегментацийн алгоритмаас хамаарч болно.[2,4,5,8,16]Vernooij болон Pemberton нар радиологичийн харааны үнэлгээнд автомат тоон хэмжилтийг хавсран ашиглах асуудлыг судалсан.[17,18] Харин Yu нар Neurophet AQUA-ийн сегментацийн гүйцэтгэлийг FreeSurfer-тэй харьцуулж, хэмжилтийн давтагдах чанарыг үнэлсэн бол Lee нар өөр өөр соронзон орны хүч бүхий CPT төхөөрөмжүүдэд хэмжилтийн тогтвортой байдлыг судалсан.[9,10] Эдгээр судалгаанд сегментацийн гүйцэтгэл, хэмжилтийн давтагдах чанар болон тогтвортой байдлыг үнэлсэн бол бидний судалгаанд автомат хэмжилтээр атрофийн зэрэглэлүүдийг ялгах чадвар, радиологичийн тайлантай тохирох байдлыг үнэлсэн.

Бидний судалгаанд ЧШХШ2-той оролцогчдын ICV-д тохируулсан тархины нийт эзлэхүүн нас, хүйсээр тохируулсан ANCOVA шинжилгээгээр хяналтын бүлгийнхээс 11.98 мл-ээр бага байв (95% CI −22.67–−1.30; P=0.028). ЧШХШ2-той оролцогчдын тархины нийт эзлэхүүн хяналтын бүлгийнхээс бага байгааг хэд хэдэн судалгаанд мэдээлсэн.[11,12] Харин радиологичийн тайланд бүртгэгдсэн атрофийн зэрэглэлийг нас, хүйсээр тохируулсан эрэмбэлэгдсэн логистик регрессээр үнэлэхэд бүлгүүдийн хооронд статистикийн ач холбогдол бүхий ялгаа тогтоогдоогүй (P=0.059).

Энэхүү судалгаа нь дараах хязгаарлагдмал талуудтай. Нэгдүгээрт, нэг төвд хийгдсэн бөгөөд бүх оролцогчийн дүрсийг нэг CPT төхөөрөмжөөр авсан. Хоёрдугаарт, радиологийн тайланд атрофи дурдаагүй тохиолдлыг 0-р зэрэглэлд ангилсан нь атрофи байхгүй гэсэн стандартчилсан үнэлгээтэй адил биш юм. Гуравдугаарт, автомат сегментацийн давхарласан дүрслэл (segmentation overlay) байхгүй тул сегментацийн чанарыг харааны аргаар шалгах болон гараар засварлах боломжгүй байв. Дөрөвдүгээрт, нас, хүйсээр тохируулсан шинжилгээнд бусад эмнэлзүйн болон судасны эрсдэлт хүчин зүйлсийг хамруулаагүй.

Дүгнэлт: Радиологичийн тайланд бүртгэгдсэн атрофийн зэрэглэл нэмэгдэх тусам Neurophet AQUA-аар хэмжсэн ICV-д тохируулсан тархины нийт эзлэхүүн шат дараалан буурч, эзлэхүүнд суурилсан ангилал тайлангийн зэрэглэлтэй өндөр жинлэсэн тохирол (quadratic-weighted κ=0.856) үзүүлэв. Мөн ЧШХШ2-той оролцогчдын тархины нийт эзлэхүүн нас, хүйсээр тохируулсны дараа хяналтын бүлгийнхээс статистикийн ач холбогдолтой бага байв. Эдгээр үр дүн нь автомат тоон морфометрийг радиологичийн харааны үнэлгээг орлох бус, нэмэлт объектив тоон мэдээллээр дэмжих хэрэгсэл болгон ашиглах боломжтойг харуулж байна.

Ном зүй:

1. *Pasquier F, Leys D, Weerts JGE, Mounier-Vehier F, Barkhof F, Scheltens P. Inter- and intraobserver reproducibility of cerebral atrophy assessment on MRI scans with hemispheric infarcts. Eur Neurol. 1996;36(5):268-272. doi:10.1159/000117270*
2. *Harper L, Barkhof F, Fox NC, Schott JM. Using visual rating to diagnose dementia: a critical evaluation of MRI atrophy scales. J Neurol Neurosurg Psychiatry. 2015;86(11):1225-1233. doi:10.1136/jnnp-2014-310090*
3. *Park M, Moon WJ. Structural MR imaging in the diagnosis of Alzheimer's disease and other neurodegenerative dementia: current imaging approach and future perspectives. Korean J Radiol. 2016;17(6):827-845. doi:10.3348/kjr.2016.17.6.827*
4. *Fornari C, Fumagalli GG. Visual rating scales of brain atrophy in dementia: a systematic review of reliability and diagnostic accuracy. Rev Neurosci. 2026;37(1):1-18. doi:10.1515/revneuro-2025-0066*
5. *Pemberton HG, Zaki LAM, Goodkin O, et al. Technical and clinical validation of commercial automated volumetric MRI tools for dementia diagnosis—a systematic review. Neuroradiology. 2021;63(11):1773-1789. doi:10.1007/s00234-021-02746-3*
6. *Persson K, Barca ML, Edwin TH, et al. Regional MRI volumetry using NeuroQuant versus visual rating scales in patients with cognitive impairment and dementia. Brain Behav. 2024;14(2):e3397. doi:10.1002/brb3.3397*
7. *Nassanga R, Nakasujja N, Kaddumukasa M, et al. Visual MRI ratings and automated volumetry in Ugandan dementia care. Academia Mental Health and Well-Being. 2026;3(2). doi:10.20935/MHealthWellB8385*
8. *Ingala S, van Maurik IS, Altomare D, et al. Clinical applicability of quantitative atrophy measures on MRI in patients suspected of Alzheimer's disease. Eur Radiol. 2022;32(11):7789-7799. doi:10.1007/s00330-021-08503-7*
9. *Yu H, Lee H, Lee M, et al. Multi-center evaluation of Neurophet AQUA for brain MRI segmentation: T1 compared with FreeSurfer and T2-FLAIR compared with ground truth. Front Neurol. 2025;16:1672133. doi:10.3389/fneur.2025.1672133*
10. *Lee H, Kim HW, Lee M, et al. Evaluating brain volume segmentation accuracy and reliability of FreeSurfer and Neurophet AQUA at variations in MRI magnetic field strengths. Sci Rep. 2024;14(1):24513. doi:10.1038/s41598-024-74622-y*
11. *Roberts RO, Knopman DS, Przybelski SA, et al. Association of type 2 diabetes with brain atrophy and cognitive impairment. Neurology. 2014;82(13):1132-1141. doi:10.1212/WNL.0000000000000269*
12. *Zhang T, Shaw M, Cherbuin N. Association between type 2 diabetes mellitus and brain atrophy: a meta-analysis. Diabetes Metab J. 2022;46(5):781-802. doi:10.4093/dmj.2021.0189*
13. *Wittens MMJ, Allemeersch GJ, Sima DM, et al. Towards validation in clinical routine: a comparative analysis of visual MTA ratings versus the automated ratio between inferior lateral ventricle and hippocampal volumes in Alzheimer's disease diagnosis. Neuroradiology. 2024;66(4):487-506. doi:10.1007/s00234-024-03280-8*
14. *Min J, Moon WJ, Jeon JY, Choi JW, Moon YS, Han SH. Diagnostic efficacy of structural MRI in patients with mild-to-moderate*

Alzheimer disease: automated volumetric assessment versus visual assessment. AJR Am J Roentgenol. 2017;208(3):617-623. doi:10.2214/AJR.16.16894

15. *Persson K, Barca ML, Cavallin L, et al. Comparison of automated volumetry of the hippocampus using NeuroQuant and visual assessment of the medial temporal lobe in Alzheimer's disease. Acta Radiol. 2018;59(8):997-1001. doi:10.1177/0284185117743778*
16. *Vernooij MW, Jasperse B, Steketee R, et al. Automatic normative quantification of brain tissue volume to support the diagnosis of dementia: a clinical evaluation of diagnostic accuracy. Neuroimage Clin. 2018;20:374-379. doi:10.1016/j.nicl.2018.08.004*
17. *Pemberton HG, Goodkin O, Prados F, et al. Automated quantitative MRI volumetry reports support diagnostic interpretation in dementia: a multi-rater, clinical accuracy study. Eur Radiol. 2021;31(7):5312-5323. doi:10.1007/s00330-020-07455-8*
18. *Gebest M, Weiß C, Cho CG, et al. Longitudinal automated brain volumetry versus expert visual assessment of atrophy progression on MRI: an exploratory study. Sci Rep. 2025;15(1):14968. doi:10.1038/s41598-025-98360-x*

Association between AI-based whole-brain volume measurements and radiology report–based atrophy grades

(A Quantitative MRI Study in Adults with and Without Type 2 Diabetes)

Tserensugir A[1,2], Oyuntugs B[3], Gonshigsuren D[4], Tuvshinjargal D[1], Tovuudorj A[5], Mungunbagana G[6], Munkhbaatar D[1], Delgerdalai Kh[1]

[1]Department of Radiology, School of Medicine, MNUMS
[2]Graduate School, MNUMS
[3]Department of Endocrinology, School of Medicine, MNUMS
[4]Department of Radiology, Grandmed Hospital
[5]Department of Neurology, School of Medicine, MNUMS
[6]Department of Radiology, Brilliant Hospital
E-mail: amd26f002@gt.mnums.edu.mn, Tel: +976-99703955

Abstract:

Background: The Global Cortical Atrophy (GCA) scale is widely used for visual assessment of cortical atrophy on MRI but is subject to interobserver variability. Automated brain volumetry provides objective measurements; however, its association with visual atrophy grading remains understudied in Mongolia.

Aim: To assess the association and agreement between radiology report–based atrophy grades and AI-derived whole-brain volume, and to compare brain volumes in adults with and without type 2 diabetes mellitus (T2DM).

Materials and Methods: This retrospective, single-center study included 367 adults aged ≥40 years (T2DM, n=144; controls, n=223). Intracranial volume–adjusted whole-brain volumes were obtained from 3D T1-weighted MRI using Neurophet AQUA v3.1. Atrophy severity recorded in routine radiology reports was coded from 0 to 3. Welch's ANOVA assessed differences across grades, and ROC analysis evaluated discrimination between adjacent grades. Classification agreement was assessed using stratified five-fold cross-validation and Cohen's κ. Age- and sex-adjusted ANCOVA compared brain volumes between groups.

Results: Mean whole-brain volume decreased across atrophy grades 0–3 (1122.12, 1068.50, 1034.76, and 968.33 mL; Welch's $F(3, 60.04)=207.30$; $P<0.001$). ROC AUCs for adjacent grades were 0.892, 0.897, and 1.000, with corresponding thresholds of 1099.86, 1056.17, and 995.57 mL. Five-fold cross-validation yielded 78.7% exact agreement (289/367), unweighted $\kappa=0.684$, and quadratic-weighted $\kappa=0.856$. Adjusted whole-brain volume was 11.98 mL lower in T2DM than in controls (95% CI, −22.67 to −1.30; $P=0.028$), whereas the overall association between T2DM and atrophy grade was not statistically significant ($P=0.059$).

Conclusion: Higher report-based atrophy grades were associated with lower whole-brain volumes, with strong quadratic-weighted agreement between volumetric and report-based classifications. Adults with T2DM had lower adjusted brain volumes than controls. Automated volumetry may complement radiologists' assessments by providing objective quantitative information.

Keywords: brain atrophy, automated segmentation, brain volume, morphometry, neuroimaging