RE: confirming MRI stats

From: **Patrick Lyden** | plyden@usc.edu Monday, May 3, 3:50 PM

To: **Ryan Cabeen** | Ryan.Cabeen@loni.usc.edu

Cc: **Jessica Lamb** | lambj@usc.edu, **Diniz, Marcio A** | Marcio.Diniz@cshs.org, **'Andre Rogatko (Andre.Rogatko@cshs.org)'** | Andre.Rogatko@cshs.org

Ryan,

I am working through Marcio's report of the MRI volumetry. I have a few questions:

1. Of 710 subjects who started surgery, there are 13 Day 2 scans and 142 Day 30 scans that have no data. Some of these are no doubt lost due to subject death, or for whatever reason were never scanned. But can you tell me how many scans were unreadable or unusable to the pipeline?
2. How do I interpret the following variables: adc_rate_mean_csf, adc-rate_mean_tisse, adc_rate_mean_lesion. Also t2_rate_mean_(tissue, csf, lesion). Most importantly, adc_qa_snr and t2_qa_snr.
3. If the adc_rate variables are important, they need to converted to larger numbers because Marcio's output is limited to 00.00 format. He can do this in R for us, but I need to know if its worth his time.
4. What are the units for the volume measures, milliliters?
5. Tissue volume was the same Day 2 and Day 30 at 340. This is not possible if we had reasonable strokes. Or this indicates a significant 'graveyard effect' and the severe strokes have died before scanning. Any ideas?
6. Your reported volume of CSF would include only ventricles right? The sudural space is stripped off with the skull stripping step?
7. We would like to have estimates of ipsilateral and contralateral (right and left) hemispheres. Would the midline shift right and left variables approximate these?
8. Looking at the QA " rate" variables, Yale is consistently different than the other sites. Looking at the QA "SNR" variables, Augusta seems quite different. Are any of these differences meaningful?

Many thanks Ryan,

PL

From: **Ryan Cabeen** | Ryan.Cabeen@loni.usc.edu Monday, May 3, 11:45 PM

To: **Patrick Lyden** | plyden@usc.edu

Cc: **Jessica Lamb** | lambj@usc.edu, **Diniz, Marcio A** | Marcio.Diniz@cshs.org, **'Andre Rogatko (Andre.Rogatko@cshs.org)'** | Andre.Rogatko@cshs.org

Hi Dr Lyden,

Glad to help figure these out, I'll make comments inline:

> From:
>
> 1. Of 710 subjects who started surgery, there are 13 Day 2 scans and 142 Day 30 scans that have no data. Some of these are no doubt lost due to subject death, or for whatever reason were never scanned. But can you tell me how many scans were unreadable or unusable to the pipeline?

Considering both early and late timepoints, we had one scan that failed conversion, six scans that had a missing/incomplete ADC

volume, and eight that had severe motion artifact. That accounts for everything I found on the IDA, so any other missing cases didn't make it to the server, from what I gather. Here's a summary of those specific scans that were excluded:

Failed conversion:
late KW9444

Incomplete ADC:
early KW9346
early QC3037
late PH1100
late PH1500
late VH2011
late KX0055

Severe motion artifact:
early FR8256
late VH5214
early QC3839
late VH1921
late AM4572
early VH1889
late VH1987
early KX0028

> From:
>
> 2. How do I interpret the following variables: adc_rate_mean_csf, adc-rate_mean_tisse, adc_rate_mean_lesion. Also t2_rate_mean_(tissue, csf, lesion). Most importantly, adc_qa_snr and t2_qa_snr.

Those first six variables are the average parameter values of ADC and R2 in CSF, tissue, and lesion. I now remember that you said this wasn't relevant, so perhaps I should have excluded them.

The "snr" variables are the ration of the average foreground signal to the average background signal, so a rough measure of image quality. I've attached a data dictionary that explains each variable in more detail, in case it's helpful.

> From:
>
> 3. If the adc_rate variables are important, they need to converted to larger numbers because Marcio's output is limited to 00.00 format. He can do this in R for us, but I need to know if its worth his time.

That makes sense to me. That variable is in units of mm2/s, so if you multiple by 1000, then they could be easily reported in mm2/ms. If that loses some precision, multiplying by 10000 makes sense to me too, as long as we report the units accordingly.

> From:
>
> 4. What are the units for the volume measures, milliliters?

Yes, that's what I understand, going by the coordinates reported by the scanner.

> From:
>
> 5. Tissue volume was the same Day 2 and Day 30 at 340. This is not possible if we had reasonable strokes. Or this indicates a significant 'graveyard effect' and the severe strokes have died before scanning. Any ideas?

This seems odd to me as well, and I noticed it in looking at the data with Dr Ayata. I'm not sure why this is happening, but I can help look

into it more. One more observation, this was reflected in the systematic differences between sites as well, e.g. Iowa was similarly lower in volume in both early and late timepoints (please find plot attached).

> From:
>
> 6. Your reported volume of CSF would include only ventricles right? The sudural space is stripped off with the skull stripping step?

I believe so, the skull stripping appears to exclude subdural space. I should look again carefully to be sure that it isn't being misclassified as tissue for some reasons, given #5 though.

> From:
>
> 7. We would like to have estimates of ipsilateral and contralateral (right and left) hemispheres. Would the midline shift right and left variables approximate these?

Yes, actually the "midline_tissue_volume_left" and "midline_tissue_volume_right" variables report the hemispheric normal-appearing tissue volumes. Does that sound right? Or do you need hemispheric CSF volume as well?

> From:
>
> 8. Looking at the QA " rate" variables, Yale is consistently different than the other sites. Looking at the QA "SNR" variables, Augusta seems quite different. Are any of these differences meaningful?

I believe they are meaningful, for example, looking at the ADC scans from Augusta, it is visually evident that the background noise is higher than other sites. We have a denoising step that mitigates this to some extent, so I wouldn't expect it to have a big impact on the resulting lesion volumes. Maybe that's something to test statistically though...

Hope that helps, and please let me know if I can help with anything else. I'll plan to look more into the early vs late tissue volumes to see if I can make any sense of that.

Cheers,
Ryan

Ryan P. Cabeen, PhD
Chan Zuckerberg Imaging Scientist
Assistant Professor of Research Neurology
Laboratory of Neuro Imaging
USC Stevens Neuroimaging and Informatics Institute
Keck School of Medicine of USC
University of Southern California
2025 Zonal Ave.
Los Angeles, CA 90033
Tel: (323) 44-BRAIN
Email: rcabeen@loni.usc.edu
Web: cabeen.io
www.ini.usc.edu

> From: **Patrick Lyden** | plyden@usc.edu Monday, May 3, 3:50 PM
>
> Ryan,
>
> I am working through Marcio's report of the MRI volumetry. I have a few questions:

late PH1500
late VH2011
late KX0055

Severe motion artifact:
early FR8256
late VH5214
early QC3839
late VH1921
late AM4572
early VH1889
late VH1987
early KX0028

2. How do I interpret the following variables: adc_rate_mean_csf, adc-rate_mean_tisse, adc_rate_mean_lesion. Also t2_rate_mean_(tissue, csf, lesion). Most importantly, adc_qa_snr and t2_qa_snr.

Those first six variables are the average parameter values of ADC and R2 in CSF, tissue, and lesion. I now remember that you said this wasn't relevant, so perhaps I should have excluded them.

The "snr" variables are the ration of the average foreground signal to the average background signal, so a rough measure of image quality. I've attached a data dictionary that explains each variable in more detail, in case it's helpful.

3. If the adc_rate variables are important, they need to converted to larger numbers because Marcio’s output is limited to 00.00 format. He can do this in R for us, but I need to know if its worth his time.

That makes sense to me. That variable is in units of mm2/s, so if you multiple by 1000, then they could be easily reported in mm2/ms. If that loses some precision, multiplying by 10000 makes sense to me too, as long as we report the units accordingly.

4. What are the units for the volume measures, milliliters?

Yes, that's what I understand, going by the coordinates reported by the scanner.

5. Tissue volume was the same Day 2 and Day 30 at 340. This is not possible if we had reasonable strokes. Or this indicates a significant ‘graveyard effect’ and the severe strokes have died before scanning. Any ideas?

This seems odd to me as well, and I noticed it in looking at the data with Dr Ayata. I'm not sure why this is happening, but I can help look into it more. One more observation, this was reflected in the systematic differences between sites as well, e.g. Iowa was similarly lower in volume in both early and late timepoints (please find plot attached).

6. Your reported volume of CSF would include only ventricles right? The sudural space is stripped off with the skull stripping step?

I believe so, the skull stripping appears to exclude subdural space. I should look again carefully to be sure that it isn't being misclassified as tissue for some reasons, given #5 though.

7. We would like to have estimates of ipsilateral and contralateral (right and left) hemispheres. Would the midline shift right and left variables approximate these?

Yes, actually the "midline_tissue_volume_left" and "midline_tissue_volume_right" variables report the hemispheric normal-appearing tissue volumes. Does that sound right? Or do you need hemispheric CSF volume as well?

8. Looking at the QA “ rate” variables, Yale is consistently different than the other sites. Looking at the QA “SNR” variables, Augusta seems quite different. Are any of these differences meaningful?

I believe they are meaningful, for example, looking at the ADC scans from Augusta, it is visually evident that the background noise is higher than other sites. We have a denoising step that mitigates this to some extent, so I wouldn't expect it to have a big impact on the resulting lesion volumes. Maybe that's something to test statistically though...

Hope that helps, and please let me know if I can help with anything else. I'll plan to look more into the early vs late tissue volumes to see if I can make any sense of that.

Cheers,
Ryan

Ryan P. Cabeen, PhD
Chan Zuckerberg Imaging Scientist
Assistant Professor of Research Neurology
Laboratory of Neuro Imaging
USC Stevens Neuroimaging and Informatics Institute
Keck School of Medicine of USC
University of Southern California
2025 Zonal Ave.
Los Angeles, CA 90033
Tel: (323) 44-BRAIN
Email: rcabeen@loni.usc.edu
Web: cabeen.io
www.ini.usc.edu

> From: **Patrick Lyden** | plyden@usc.edu Monday, May 3, 3:50 PM
>
> Ryan,
>
> I am working through Marcio's report of the MRI volumetry. I have a few questions:
>
> 1. Of 710 subjects who started surgery, there are 13 Day 2 scans and 142 Day 30 scans that have no data. Some of these are no doubt lost due to subject death, or for whatever reason were never scanned. But can you tell me how many scans were unreadable or unusable to the pipeline?
> 2. How do I interpret the following variables: adc_rate_mean_csf, adc-rate_mean_tisse, adc_rate_mean_lesion. Also t2_rate_mean_(tissue, csf, lesion). Most importantly, adc_qa_snr and t2_qa_snr.
> 3. If the adc_rate variables are important, they need to converted to larger numbers because Marcio's output is limited to 00.00 format. He can do this in R for us, but I need to know if its worth his time.
> 4. What are the units for the volume measures, milliliters?
> 5. Tissue volume was the same Day 2 and Day 30 at 340. This is not possible if we had reasonable strokes. Or this indicates a significant 'graveyard effect' and the severe strokes have died before scanning. Any ideas?
> 6. Your reported volume of CSF would include only ventricles right? The sudural space is stripped off with the skull stripping step?
> 7. We would like to have estimates of ipsilateral and contralateral (right and left) hemispheres. Would the midline shift right and

late VH2011
late KX0055

Severe motion artifact:
early FR8256
late VH5214
early QC3839
late VH1921
late AM4572
early VH1889
late VH1987
early KX0028

2. How do I interpret the following variables: adc_rate_mean_csf, adc-rate_mean_tisse, adc_rate_mean_lesion. Also t2_rate_mean_(tissue, csf, lesion). Most importantly, adc_qa_snr and t2_qa_snr.

Those first six variables are the average parameter values of ADC and R2 in CSF, tissue, and lesion. I now remember that you said this wasn't relevant, so perhaps I should have excluded them.

The "snr" variables are the ration of the average foreground signal to the average background signal, so a rough measure of image quality. I've attached a data dictionary that explains each variable in more detail, in case it's helpful.

3. If the adc_rate variables are important, they need to converted to larger numbers because Marcio’s output is limited to 00.00 format. He can do this in R for us, but I need to know if its worth his time.

That makes sense to me. That variable is in units of mm2/s, so if you multiple by 1000, then they could be easily reported in mm2/ms. If that loses some precision, multiplying by 10000 makes sense to me too, as long as we report the units accordingly.

4. What are the units for the volume measures, milliliters?

Yes, that's what I understand, going by the coordinates reported by the scanner.

5. Tissue volume was the same Day 2 and Day 30 at 340. This is not possible if we had reasonable strokes. Or this indicates a significant ‘graveyard effect’ and the severe strokes have died before scanning. Any ideas?

This seems odd to me as well, and I noticed it in looking at the data with Dr Ayata. I'm not sure why this is happening, but I can help look into it more. One more observation, this was reflected in the systematic differences between sites as well, e.g. Iowa was similarly lower in volume in both early and late timepoints (please find plot attached).

6. Your reported volume of CSF would include only ventricles right? The sudural space is stripped off with the skull stripping step?

I believe so, the skull stripping appears to exclude subdural space. I should look again carefully to be sure that it isn't being misclassified as tissue for some reasons, given #5 though.

7. We would like to have estimates of ipsilateral and contralateral (right and left) hemispheres. Would the midline shift right and left variables approximate these?

Yes, actually the "midline_tissue_volume_left" and "midline_tissue_volume_right" variables report the hemispheric normal-appearing tissue volumes. Does that sound right? Or do you need hemispheric CSF volume as well?

8. Looking at the QA “ rate” variables, Yale is consistently different than the other sites. Looking at the QA “SNR” variables, Augusta seems quite different. Are any of these differences meaningful?

I believe they are meaningful, for example, looking at the ADC scans from Augusta, it is visually evident that the background noise is higher than other sites. We have a denoising step that mitigates this to some extent, so I wouldn't expect it to have a big impact on the resulting lesion volumes. Maybe that's something to test statistically though...

Hope that helps, and please let me know if I can help with anything else. I'll plan to look more into the early vs late tissue volumes to see if I can make any sense of that.

Cheers,
Ryan

Ryan P. Cabeen, PhD
Chan Zuckerberg Imaging Scientist
Assistant Professor of Research Neurology
Laboratory of Neuro Imaging
USC Stevens Neuroimaging and Informatics Institute
Keck School of Medicine of USC
University of Southern California
2025 Zonal Ave.
Los Angeles, CA 90033
Tel: (323) 44-BRAIN
Email: rcabeen@loni.usc.edu
Web: cabeen.io
www.ini.usc.edu

From: **Patrick Lyden** | plyden@usc.edu Monday, May 3, 3:50 PM

Ryan,

I am working through Marcio's report of the MRI volumetry. I have a few questions:

1. Of 710 subjects who started surgery, there are 13 Day 2 scans and 142 Day 30 scans that have no data. Some of these are no doubt lost due to subject death, or for whatever reason were never scanned. But can you tell me how many scans were unreadable or unusable to the pipeline?
2. How do I interpret the following variables: adc_rate_mean_csf, adc-rate_mean_tisse, adc_rate_mean_lesion. Also t2_rate_mean_(tissue, csf, lesion). Most importantly, adc_qa_snr and t2_qa_snr.
3. If the adc_rate variables are important, they need to converted to larger numbers because Marcio's output is limited to 00.00 format. He can do this in R for us, but I need to know if its worth his time.
4. What are the units for the volume measures, milliliters?
5. Tissue volume was the same Day 2 and Day 30 at 340. This is not possible if we had reasonable strokes. Or this indicates a significant 'graveyard effect' and the severe strokes have died before scanning. Any ideas?
6. Your reported volume of CSF would include only ventricles right? The sudural space is stripped off with the skull stripping step?
7. We would like to have estimates of ipsilateral and contralateral (right and left) hemispheres. Would the midline shift right and left variables approximate these?

> 8. Looking at the QA “ rate” variables, Yale is consistently different than the other sites. Looking at the QA “SNR” variables, Augusta seems quite different. Are any of these differences meaningful?
>
> Many thanks Ryan,
>
> PL

From: **Ryan Cabeen** | ryan.cabeen@loni.usc.edu — Tuesday, Jul 6, 6:47 PM

To: **Patrick Lyden** | plyden@usc.edu

For point #4, that’s correct, the units are reported in mm^3 — my apologies for my lapse in recalling unit conversion!

For point #5, yes, I’ll create additional plots that show left and right volumes as well. Should be interesting!

Ryan P. Cabeen, PhD
Chan Zuckerberg Imaging Scientist
Assistant Professor of Research Neurology
Laboratory of Neuro Imaging
USC Stevens Neuroimaging and Informatics Institute
Keck School of Medicine of USC
University of Southern California
2025 Zonal Ave.
Los Angeles, CA 90033
Tel: (323) 44-BRAIN
Email: rcabeen@loni.usc.edu
Web: cabeen.io
www.ini.usc.edu

> From: **Patrick Lyden** | plyden@usc.edu — Tuesday, Jul 6, 6:39 AM
>
> Two more follow up points:
>
> Point #4 below, what are the units. You agreed with me that the volume units are in ml. Published data used units of mm3. I don’t think these are the same thing: ml = cc3 = 1000mm3. Can you confirm that the linear units are mm, areal units mm2, and volume units mm3?
>
> Point #5 below, Tissue volumes the same Days 2 and 30. Thanks for the graphic, it tells the tale. Because most sites had smaller lesions, and because lesion is not counted as tissue at Day 2, the mean tissue volumes tended lower at Day 2 (Counter intuitive). Could you regenerate the same graph, but split tissue into left and right? That will show the issue best, I think.
>
> Many thanks
>
> P
>
> > From: **Ryan Cabeen** | Ryan.Cabeen@loni.usc.edu — To: **Patrick Lyden** | plyden@usc.edu — Tuesday, May 4, 2:45 AM
> >
> > Hi Dr Lyden,
> >
> > Glad to help figure these out, I'll make comments

inline:

1. Of 710 subjects who started surgery, there are 13 Day 2 scans and 142 Day 30 scans that have no data. Some of these are no doubt lost due to subject death, or for whatever reason were never scanned. But can you tell me how many scans were unreadable or unusable to the pipeline?

Considering both early and late timepoints, we had one scan that failed conversion, six scans that had a missing/incomplete ADC volume, and eight that had severe motion artifact. That accounts for everything I found on the IDA, so any other missing cases didn't make it to the server, from what I gather. Here's a summary of those specific scans that were excluded:

Failed conversion:
late KW9444

Incomplete ADC:
early KW9346
early QC3037
late PH1100
late PH1500
late VH2011
late KX0055

Severe motion artifact:
early FR8256
late VH5214
early QC3839
late VH1921
late AM4572
early VH1889
late VH1987
early KX0028

2. How do I interpret the following variables: adc_rate_mean_csf, adc-rate_mean_tisse, adc_rate_mean_lesion. Also t2_rate_mean_(tissue, csf, lesion). Most importantly, adc_qa_snr and t2_qa_snr.

Those first six variables are the average parameter values of ADC and R2 in CSF, tissue, and lesion. I now remember that you said this wasn't relevant, so perhaps I should have excluded them.

The "snr" variables are the ration of the average foreground signal to the average background signal, so a rough measure of image quality. I've attached a data dictionary that explains each variable in more detail, in case it's helpful.

3. If the adc_rate variables are important, they need to converted to larger numbers because Marcio’s output is limited to 00.00 format. He can do this in R for us, but I need to know if its worth his time.

That makes sense to me. That variable is in units of mm2/s, so if you multiple by 1000, then they could be easily reported in mm2/ms. If that loses some precision, multiplying by 10000 makes sense to me too, as long as we report the units accordingly.

4. What are the units for the volume measures, milliliters?

Yes, that's what I understand, going by the coordinates reported by the scanner.

5. Tissue volume was the same Day 2 and Day 30 at 340. This is not possible if we had reasonable strokes. Or this indicates a significant ‘graveyard effect’ and the severe strokes have died before scanning. Any ideas?

This seems odd to me as well, and I noticed it in looking at the data with Dr Ayata. I'm not sure why this is happening, but I can help look into it more. One more observation, this was reflected in the systematic differences between sites as well, e.g. Iowa was similarly lower in volume in both early and late timepoints (please find plot attached).

6. Your reported volume of CSF would include only ventricles right? The sudural space is stripped off with the skull stripping step?

I believe so, the skull stripping appears to exclude subdural space. I should look again carefully to be sure that it isn't being misclassified as tissue for some reasons, given #5 though.

7. We would like to have estimates of ipsilateral and contralateral (right and left) hemispheres. Would the midline shift right and left variables approximate these?

Yes, actually the "midline_tissue_volume_left" and "midline_tissue_volume_right" variables report the hemispheric normal-appearing tissue volumes. Does that sound right? Or do you need hemispheric CSF volume as well?

8. Looking at the QA “ rate” variables, Yale is consistently different than the other sites. Looking at the QA “SNR” variables, Augusta seems quite different. Are any of these differences meaningful?

I believe they are meaningful, for example, looking at the ADC scans from Augusta, it is visually evident that the background noise is higher than other sites. We have a denoising step that mitigates this to some extent, so I wouldn't expect it to have a big impact on the resulting lesion volumes. Maybe that's something to test statistically though...

Hope that helps, and please let me know if I can help with anything else. I'll plan to look more into the early vs late tissue volumes to see if I can make any sense of that.

Cheers,
Ryan

Ryan P. Cabeen, PhD
Chan Zuckerberg Imaging Scientist
Assistant Professor of Research Neurology
Laboratory of Neuro Imaging
USC Stevens Neuroimaging and Informatics Institute
Keck School of Medicine of USC
University of Southern California
2025 Zonal Ave.
Los Angeles, CA 90033
Tel: (323) 44-BRAIN
Email: rcabeen@loni.usc.edu
Web: cabeen.io
www.ini.usc.edu

From: **Patrick Lyden** | plyden@usc.edu Monday, May 3, 3:50 PM

Ryan,

I am working through Marcio's report of the MRI volumetry. I have a few questions:

1. Of 710 subjects who started surgery, there are 13 Day 2 scans and 142 Day 30 scans that have no data. Some of these are no doubt lost due to subject death, or for whatever reason were never scanned. But can you tell me how many scans were unreadable or unusable to the pipeline?
2. How do I interpret the following variables: adc_rate_mean_csf, adc-rate_mean_tisse, adc_rate_mean_lesion. Also t2_rate_mean_(tissue, csf, lesion). Most importantly, adc_qa_snr and t2_qa_snr.
3. If the adc_rate variables are important, they need to converted to larger numbers because Marcio's output is limited to 00.00 format. He can do this in R for us, but I need to know if its worth his time.
4. What are the units for the volume measures, milliliters?
5. Tissue volume was the same Day 2 and Day 30 at 340. This is not possible if we had reasonable strokes. Or this indicates a significant 'graveyard effect' and the severe strokes have died before scanning. Any ideas?
6. Your reported volume of CSF would include only ventricles right? The sudural space is stripped off with the skull stripping step?
7. We would like to have estimates of ipsilateral and contralateral (right and left) hemispheres. Would the midline shift right and left variables approximate these?
8. Looking at the QA " rate" variables, Yale is consistently different than the other sites. Looking at the QA "SNR" variables, Augusta seems quite different. Are any of these differences meaningful?

Many thanks Ryan,

PL

From: **Patrick Lyden** | plyden@usc.edu Wednesday, Jul 7, 10:58 AM

To: **Ryan Cabeen** | Ryan.Cabeen@loni.usc.edu

Thank you.

From: **Ryan Cabeen** | Ryan.Cabeen@loni.usc.edu To: **Patrick Lyden** | plyden@usc.edu Tuesday, Jul 6, 9:48 PM

For point #4, that's correct, the units are reported in mm^3 — my apologies for my lapse in recalling unit conversion!

For point #5, yes, I'll create additional plots that show left and right volumes as well. Should be interesting!

Ryan P. Cabeen, PhD
Chan Zuckerberg Imaging Scientist
Assistant Professor of Research Neurology
Laboratory of Neuro Imaging
USC Stevens Neuroimaging and Informatics Institute
Keck School of Medicine of USC
University of Southern California
2025 Zonal Ave.
Los Angeles, CA 90033
Tel: (323) 44-BRAIN

Email: rcabeen@loni.usc.edu
Web: cabeen.io
www.ini.usc.edu

From: **Patrick Lyden** | plyden@usc.edu Tuesday, Jul 6, 6:39 AM

Two more follow up points:

Point #4 below, what are the units. You agreed with me that the volume units are in ml. Published data used units of mm3. I don't think these are the same thing: ml = cc3 = 1000mm3. Can you confirm that the linear units are mm, areal units mm2, and volume units mm3?

Point #5 below, Tissue volumes the same Days 2 and 30. Thanks for the graphic, it tells the tale. Because most sites had smaller lesions, and because lesion is not counted as tissue at Day 2, the mean tissue volumes tended lower at Day 2 (Counter intuitive). Could you regenerate the same graph, but split tissue into left and right? That will show the issue best, I think.

Many thanks

P

From: **Ryan Cabeen** | Ryan.Cabeen@loni.usc.edu To: **Patrick Lyden** | plyden@usc.edu Tuesday, May 4, 2:45 AM

Hi Dr Lyden,

Glad to help figure these out, I&

late AM4572
early VH1889
late VH1987
early KX0028

2. How do I interpret the following variables: adc_rate_mean_csf, adc-rate_mean_tisse, adc_rate_mean_lesion. Also t2_rate_mean_(tissue, csf, lesion). Most importantly, adc_qa_snr and t2_qa_snr.

Those first six variables are the average parameter values of ADC and R2 in CSF, tissue, and lesion. I now remember that you said this wasn't relevant, so perhaps I should have excluded them.

The "snr" variables are the ration of the average foreground signal to the average background signal, so a rough measure of image quality. I've attached a data dictionary that explains each variable in more detail, in case it's helpful.

3. If the adc_rate variables are important, they need to converted to larger numbers because Marcio’s output is limited to 00.00 format. He can do this in R for us, but I need to know if its worth his time.

That makes sense to me. That variable is in units of mm2/s, so if you multiple by 1000, then they could be easily reported in mm2/ms. If that loses some precision, multiplying by 10000 makes sense to me too, as long as we report the units accordingly.

4. What are the units for the volume measures, milliliters?

Yes, that's what I understand, going by the coordinates reported by the scanner.

5. Tissue volume was the same Day 2 and Day 30 at 340. This is not possible if we had reasonable strokes. Or this indicates a significant ‘graveyard effect’ and the severe strokes have died before scanning. Any ideas?

This seems odd to me as well, and I noticed it in looking at the data with Dr Ayata. I'm not sure why this is happening, but I can help look into it more. One more observation, this was reflected in the systematic differences between sites as well, e.g. Iowa was similarly lower in volume in both early and late timepoints (please find plot attached).

6. Your reported volume of CSF would include only ventricles right? The sudural space is stripped off with the skull stripping step?

I believe so, the skull stripping appears to exclude subdural space. I should look again carefully to be sure that it isn't being misclassified as tissue for some reasons, given #5 though.

7. We would like to have estimates of ipsilateral and contralateral (right and left) hemispheres. Would the midline shift right and left variables approximate these?

Yes, actually the "midline_tissue_volume_left" and "midline_tissue_volume_right" variables report the hemispheric normal-appearing tissue volumes. Does that sound right? Or do you need hemispheric CSF volume as well?

8. Looking at the QA “ rate” variables, Yale is consistently different than the other sites. Looking at the QA “SNR” variables, Augusta seems quite different. Are any of these differences meaningful?

I believe they are meaningful, for example, looking at the ADC scans from Augusta, it is visually evident that the background noise is higher than other sites. We have a denoising step that mitigates this to some extent, so I wouldn't expect it to have a big impact on the resulting lesion volumes. Maybe that's something to test statistically though...

Hope that helps, and please let me know if I can help with anything else. I'll plan to look more into the early vs late tissue volumes to see if I can make any sense of that.

Cheers,
Ryan

Ryan P. Cabeen, PhD
Chan Zuckerberg Imaging Scientist
Assistant Professor of Research Neurology
Laboratory of Neuro Imaging
USC Stevens Neuroimaging and Informatics Institute
Keck School of Medicine of USC
University of Southern California
2025 Zonal Ave.
Los Angeles, CA 90033
Tel: (323) 44-BRAIN
Email: rcabeen@loni.usc.edu
Web: http://cabeen.io
www.ini.usc.edu

From: **Patrick Lyden** | plyden@usc.edu Monday, May 3, 3:50 PM

Ryan,

I am working through Marcio's report of the MRI volumetry. I have a few questions:

1. Of 710 subjects who started surgery, there are 13 Day 2 scans and 142 Day 30 scans that have no data. Some of these are no doubt lost due to subject death, or for whatever reason were never scanned. But can you tell me how many scans were unreadable or unusable to the pipeline?
2. How do I interpret the following variables: adc_rate_mean_csf, adc-rate_mean_tisse, adc_rate_mean_lesion. Also t2_rate_mean_(tissue, csf, lesion). Most importantly, adc_qa_snr and t2_qa_snr.
3. If the adc_rate variables are important, they need to converted to larger numbers because Marcio's output is limited to 00.00 format. He can do this in R for us, but I need to know if its worth his time.
4. What are the units for the volume measures, milliliters?
5. Tissue volume was the same Day 2 and Day 30 at 340. This is not possible if we had reasonable strokes. Or this indicates a significant 'graveyard effect' and the severe strokes have died before scanning. Any ideas?
6. Your reported volume of CSF would include only ventricles right? The sudural space is stripped off with the skull stripping step?
7. We would like to have estimates of ipsilateral and contralateral (right and left) hemispheres. Would the midline shift right and left variables approximate these?
8. Looking at the QA " rate" variables, Yale is consistently different than the other sites. Looking at the QA "SNR" variables, Augusta seems quite different. Are any of these differences meaningful?

Many thanks Ryan,

PL

From: **Ryan Cabeen** | ryan.cabeen@loni.usc.edu Thursday, Jul 8, 12:49 AM

To: **Patrick Lyden** | plyden@usc.edu

Here you go, the attached plots show the per-hemisphere tissue volumes by site and timepoint. Let me know if anything else would be helpful

Ryan P. Cabeen, PhD
Chan Zuckerberg Imaging Scientist
Assistant Professor of Research Neurology
Laboratory of Neuro Imaging
USC Stevens Neuroimaging and Informatics Institute
Keck School of Medicine of USC
University of Southern California
2025 Zonal Ave.
Los Angeles, CA 90033
Tel: (323) 44-BRAIN
Email: rcabeen@loni.usc.edu
Web: cabeen.io
www.ini.usc.edu

> From: **Patrick Lyden** | plyden@usc.edu Wednesday, Jul 7, 10:58 AM
>
> Thank you.
>
> > From: **Ryan Cabeen** | Ryan.Cabeen@loni.usc.edu To: **Patrick Lyden** | plyden@usc.edu Tuesday, Jul 6, 9:48 PM
> >
> > For point #4, that's correct, the units are reported in mm^3 — my apologies for my lapse in recalling unit conversion!
> >
> > For point #5, yes, I'll create additional plots that show left and right volumes as well. Should be interesting!
> >
> > Ryan P. Cabeen, PhD
> > Chan Zuckerberg Imaging Scientist
> > Assistant Professor of Research Neurology
> > Laboratory of Neuro Imaging
> > USC Stevens Neuroimaging and Informatics Institute
> > Keck School of Medicine of USC
> > University of Southern California
> > 2025 Zonal Ave.
> > Los Angeles, CA 90033
> > Tel: (323) 44-BRAIN
> > Email: rcabeen@loni.usc.edu
> > Web: cabeen.io
> > www.ini.usc.edu
> >
> > > From: **Patrick Lyden** | plyden@usc.edu Tuesday, Jul 6, 6:39 AM
> > >
> > > Two more follow up points:
> > >
> > > Point #4 below, what are the units. You agreed with me that the volume units are in ml. Published data used units of mm3. I don't think these are the same thing: ml = cc3 = 1000mm3. Can you confirm that the linear units are mm, areal units mm2, and volume

said this wasn't relevant, so perhaps I should have excluded them.

The
"snr" variables are the ration of the average foreground signal to the average background signal, so a rough measure of image quality. I've attached a data dictionary that explains each variable in more detail, in case it's helpful.

3. If the adc_rate variables are important, they need to converted to larger numbers because Marcio’s output is limited to 00.00 format. He can do this in R for us, but I need to know if its worth his time.

That makes sense to me. That variable is in units of mm2/s, so if you multiple by 1000, then they could be easily reported in mm2/ms. If that loses some precision, multiplying by 10000 makes sense to me too, as long as we report the units accordingly.

4. What are the units for the volume measures, milliliters?

Yes, that's what I understand, going by the coordinates reported by the scanner.

5. Tissue volume was the same Day 2 and Day 30 at 340. This is not possible if we had reasonable strokes. Or this indicates a significant ‘graveyard effect’ and the severe strokes have died before scanning. Any ideas?

This seems odd to me as well, and I noticed it in looking at the data with Dr Ayata. I'm not sure why this is happening, but I can help look into it more. One more observation, this was reflected in the systematic differences between sites as well, e.g. Iowa was similarly lower in volume in both early and late timepoints (please find plot attached).

6. Your reported volume of CSF would include only ventricles right? The sudural space is stripped off with the skull stripping step?

I believe so, the skull stripping appears to exclude subdural space. I should look again carefully to be sure that it isn't being misclassified as tissue for some reasons, given #5 though.

7. We would like to have estimates of ipsilateral and contralateral (right and left) hemispheres. Would the midline shift right and left variables approximate these?

Yes, actually the "midline_tissue_volume_left" and "midline_tissue_volume_right" variables report the hemispheric normal-appearing tissue volumes. Does that sound right? Or do you need hemispheric CSF volume as well?

8. Looking at the QA “ rate” variables, Yale is consistently different than the other sites. Looking at the QA “SNR” variables, Augusta seems quite different. Are any of these differences meaningful?

I believe they are meaningful, for example, looking at the ADC scans from Augusta, it is visually evident that the background noise is higher than other sites. We have a denoising step that mitigates this to some extent, so I wouldn't expect it to have a big impact on the resulting lesion volumes. Maybe that's something to test statistically though...

Hope that helps, and please let me know if I can help with anything else. I'll plan to look more into the early vs late tissue volumes to see if I can make any sense of that.

Cheers,

Ryan

Ryan P. Cabeen, PhD
Chan Zuckerberg Imaging Scientist
Assistant Professor of Research Neurology
Laboratory of Neuro Imaging
USC Stevens Neuroimaging and Informatics Institute
Keck School of Medicine of USC
University of Southern California
2025 Zonal Ave.
Los Angeles, CA 90033
Tel: (323) 44-BRAIN
Email: rcabeen@loni.usc.edu
Web: http://cabeen.io
www.ini.usc.edu

From: **Patrick Lyden** | plyden@usc.edu Monday, May 3, 3:50 PM

Ryan,

I am working through Marcio's report of the MRI volumetry. I have a few questions:

1. Of 710 subjects who started surgery, there are 13 Day 2 scans and 142 Day 30 scans that have no data. Some of these are no doubt lost due to subject death, or for whatever reason were never scanned. But can you tell me how many scans were unreadable or unusable to the pipeline?
2. How do I interpret the following variables: adc_rate_mean_csf, adc-rate_mean_tisse, adc_rate_mean_lesion. Also t2_rate_mean_(tissue, csf, lesion). Most importantly, adc_qa_snr and t2_qa_snr.
3. If the adc_rate variables are important, they need to converted to larger numbers because Marcio's output is limited to 00.00 format. He can do this in R for us, but I need to know if its worth his time.
4. What are the units for the volume measures, milliliters?
5. Tissue volume was the same Day 2 and Day 30 at 340. This is not possible if we had reasonable strokes. Or this indicates a significant 'graveyard effect' and the severe strokes have died before scanning. Any ideas?
6. Your reported volume of CSF would include only ventricles right? The sudural space is stripped off with the skull stripping step?
7. We would like to have estimates of ipsilateral and contralateral (right and left) hemispheres. Would the midline shift right and left variables approximate these?
8. Looking at the QA " rate" variables, Yale is consistently different than the other sites. Looking at the QA "SNR" variables, Augusta seems quite different. Are any of these differences meaningful?

Many thanks Ryan,

PL

From: **Patrick Lyden** | plyden@usc.edu Monday, Jul 12, 4:39 AM

To: **Ryan Cabeen** | Ryan.Cabeen@loni.usc.edu

This certainly answers—finally—our key concerns. Beautiful. It is clear that the hemisphere ipsi to the stroke (right) is smaller early and

recorvers. But comparing ot the contra (left) side, tissue has been lost. It is particularly interesting that UI has smaller contra hemisphere —somehow they ended up with smaller animals.

Thank you these data are most illuminating.

> From: **Ryan Cabeen** | Ryan.Cabeen@loni.usc.edu To: **Patrick Lyden** | plyden@usc.edu Thursday, Jul 8, 3:50 AM
>
> Here you go, the attached plots show the per-hemisphere tissue volumes by site and timepoint. Let me know if anything else would be helpful
>
> Ryan P. Cabeen, PhD
> Chan Zuckerberg Imaging Scientist
> Assistant Professor of Research Neurology
> Laboratory of Neuro Imaging
> USC Stevens Neuroimaging and Informatics Institute
> Keck School of Medicine of USC
> University of Southern California
> 2025 Zonal Ave.
> Los Angeles, CA 90033
> Tel: (323) 44-BRAIN
> Email: rcabeen@loni.usc.edu
> Web: cabeen.io
> www.ini.usc.edu
>
> > From: **Patrick Lyden** | plyden@usc.edu Wednesday, Jul 7, 10:58 AM
> >
> > Thank you.
> >
> > > From: **Ryan Cabeen** | Ryan.Cabeen@loni.usc.edu To: **Patrick Lyden** | plyden@usc.edu Tuesday, Jul 6, 9:48 PM
> > >
> > > For point #4, that's correct, the units are reported in mm^3 — my apologies for my lapse in recalling unit conversion!
> > >
> > > For point #5, yes, I'll create additional plots that show left and right volumes as well. Should be interesting!
> > >
> > > Ryan P. Cabeen, PhD
> > > Chan Zuckerberg Imaging Scientist
> > > Assistant Professor of Research Neurology
> > > Laboratory of Neuro Imaging
> > > USC Stevens Neuroimaging and Informatics Institute
> > > Keck School of Medicine of USC
> > > University of Southern California
> > > 2025 Zonal Ave.
> > > Los Angeles, CA 90033
> > > Tel: (323) 44-BRAIN
> > > Email: rcabeen@loni.usc.edu
> > > Web: http://cabeen.io
> > > www.ini.usc.edu
> > >
> > > > From: **Patrick Lyden** | plyden@usc.edu Tuesday, Jul 6, 6:39 AM
> > > >
> > > > Two more follow up points:
> > > >
> > > > Point #4 below, what are the units. You agreed with me that the volume units are in ml. Published data used units of mm3. I don't think these are the same thing: ml = cc3 = 1000mm3. Can you confirm that the linear units are mm, areal units mm2, and volume

Those first six variables are the average parameter values of ADC and R2 in CSF, tissue, and lesion. I now remember that you said this wasn't relevant, so perhaps I should have excluded them.

The "snr" variables are the ration of the average foreground signal to the average background signal, so a rough measure of image quality. I've attached a data dictionary that explains each variable in more detail, in case it's helpful.

3. If the adc_rate variables are important, they need to converted to larger numbers because Marcio’s output is limited to 00.00 format. He can do this in R for us, but I need to know if its worth his time.

That makes sense to me. That variable is in units of mm2/s, so if you multiple by 1000, then they could be easily reported in mm2/ms. If that loses some precision, multiplying by 10000 makes sense to me too, as long as we report the units accordingly.

4. What are the units for the volume measures, milliliters?

Yes, that's what I understand, going by the coordinates reported by the scanner.

5. Tissue volume was the same Day 2 and Day 30 at 340. This is not possible if we had reasonable strokes. Or this indicates a significant ‘graveyard effect’ and the severe strokes have died before scanning. Any ideas?

This seems odd to me as well, and I noticed it in looking at the data with Dr Ayata. I'm not sure why this is happening, but I can help look into it more. One more observation, this was reflected in the systematic differences between sites as well, e.g. Iowa was similarly lower in volume in both early and late timepoints (please find plot attached).

6. Your reported volume of CSF would include only ventricles right? The sudural space is stripped off with the skull stripping step?

I believe so, the skull stripping appears to exclude subdural space. I should look again carefully to be sure that it isn't being misclassified as tissue for some reasons, given #5 though.

7. We would like to have estimates of ipsilateral and contralateral (right and left) hemispheres. Would the midline shift right and left variables approximate these?

Yes, actually the "midline_tissue_volume_left" and "midline_tissue_volume_right" variables report the hemispheric normal-appearing tissue volumes. Does that sound right? Or do you need hemispheric CSF volume as well?

8. Looking at the QA “ rate” variables, Yale is consistently different than the other sites. Looking at the QA “SNR” variables, Augusta seems quite different. Are any of these differences meaningful?

I believe they are meaningful, for example, looking at the ADC scans from Augusta, it is visually evident that the background noise is higher than other sites. We have a denoising step that mitigates this to some extent, so I wouldn't expect it to have a big impact on the resulting lesion volumes. Maybe that's something to test statistically though...

Hope that helps, and please let me know if I can help with anything else. I'll plan to look more into the early vs late tissue

volumes to see if I can make any sense of that.

Cheers,
Ryan

Ryan P. Cabeen, PhD
Chan Zuckerberg Imaging Scientist
Assistant Professor of Research Neurology
Laboratory of Neuro Imaging
USC Stevens Neuroimaging and Informatics Institute
Keck School of Medicine of USC
University of Southern California
2025 Zonal Ave.
Los Angeles, CA 90033
Tel: (323) 44-BRAIN
Email: rcabeen@loni.usc.edu
Web: http://cabeen.io
www.ini.usc.edu

> From: **Patrick Lyden** | plyden@usc.edu Monday, May 3, 3:50 PM
>
> Ryan,
>
> I am working through Marcio's report of the MRI volumetry. I have a few questions:
>
> 1. Of 710 subjects who started surgery, there are 13 Day 2 scans and 142 Day 30 scans that have no data. Some of these are no doubt lost due to subject death, or for whatever reason were never scanned. But can you tell me how many scans were unreadable or unusable to the pipeline?
> 2. How do I interpret the following variables: adc_rate_mean_csf, adc-rate_mean_tisse, adc_rate_mean_lesion. Also t2_rate_mean_(tissue, csf, lesion). Most importantly, adc_qa_snr and t2_qa_snr.
> 3. If the adc_rate variables are important, they need to converted to larger numbers because Marcio's output is limited to 00.00 format. He can do this in R for us, but I need to know if its worth his time.
> 4. What are the units for the volume measures, milliliters?
> 5. Tissue volume was the same Day 2 and Day 30 at 340. This is not possible if we had reasonable strokes. Or this indicates a significant 'graveyard effect' and the severe strokes have died before scanning. Any ideas?
> 6. Your reported volume of CSF would include only ventricles right? The sudural space is stripped off with the skull stripping step?
> 7. We would like to have estimates of ipsilateral and contralateral (right and left) hemispheres. Would the midline shift right and left variables approximate these?
> 8. Looking at the QA " rate" variables, Yale is consistently different than the other sites. Looking at the QA "SNR" variables, Augusta seems quite different. Are any of these differences meaningful?
>
> Many thanks Ryan,
>
> PL