



Published in final edited form as:

Emerg Radiol. 2022 December ; 29(6): 995–1002. doi:10.1007/s10140-022-02087-5.

A pilot study of deep learning-based CT volumetry for traumatic hemothorax

David Dreizin¹, Bryan Nixon², Jiazhen Hu³, Benjamin Albert³, Chang Yan³, Gary Yang³, Haomin Chen³, Yuanyuan Liang⁴, Nahye Kim², Jean Jeudy², Guang Li², Elana B. Smith¹, Mathias Unberath³

¹Department of Diagnostic Radiology and Nuclear Medicine, R Adams Cowley Shock Trauma Center, University of Maryland School of Medicine, 22 S Greene St, Baltimore, MD 21201, USA

²Department of Diagnostic Radiology and Nuclear Medicine, University of Maryland School of Medicine, Baltimore, MD, USA

³Department of Computer Science, Johns Hopkins University, Baltimore, MD, USA

⁴Epidemiology & Public Health, University of Maryland School of Medicine, Baltimore, MD, USA

Abstract

Purpose—We employ nnU-Net, a state-of-the-art self-configuring deep learning-based semantic segmentation method for quantitative visualization of hemothorax (HTX) in trauma patients, and assess performance using a combination of overlap and volume-based metrics. The accuracy of hemothorax volumes for predicting a composite of hemorrhage-related outcomes — massive transfusion (MT) and in-hospital mortality (IHM) not related to traumatic brain injury — is assessed and compared to subjective expert consensus grading by an experienced chest and emergency radiologist.

Materials and methods—The study included manually labeled admission chest CTs from 77 consecutive adult patients with non-negligible (> 50 mL) traumatic HTX between 2016 and 2018 from one trauma center. DL results of ensembled nnU-Net were determined from fivefold cross-validation and compared to individual 2D, 3D, and cascaded 3D nnU-Net results using the Dice similarity coefficient (DSC) and volume similarity index. Pearson's r , intraclass correlation coefficient (ICC), and mean bias were also determined for the best performing model. Manual and automated hemothorax volumes and subjective hemothorax volume grades were analyzed

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

✉ David Dreizin, daviddreizin@gmail.com.

Conflict of interest David Dreizin, MD, is founder of TraumaVisual, LLC.

Code availability Code can be made available on github.

Declarations

Ethics approval IRB approval was obtained for this study at the University of Maryland School of Medicine.

Consent for study The IRB waived informed consent due to retrospective observational nature and deidentified data.

Consent for publication IRB approved waiver of informed consent for publication of observational study of deidentified data as above.

as predictors of MT and IHM using AUC comparison. Volume cut-offs yielding sensitivity or specificity 90% were determined from ROC analysis.

Results—Ensembled nnU-Net achieved a mean DSC of 0.75 (SD: \pm 0.12), and mean volume similarity of 0.91 (SD: \pm 0.10), Pearson r of 0.93, and ICC of 0.92. Mean overmeasurement bias was only 1.7 mL despite a range of manual HTX volumes from 35 to 1503 mL (median: 178 mL). AUC of automated volumes for the composite outcome was 0.74 (95%CI: 0.58–0.91), compared to 0.76 (95%CI: 0.58–0.93) for manual volumes, and 0.76 (95%CI: 0.62–0.90) for consensus expert grading ($p = 0.93$). Automated volume cut-offs of 77 mL and 334 mL predicted the outcome with 93% sensitivity and 90% specificity respectively.

Conclusion—Automated HTX volumetry had high method validity, yielded interpretable visual results, and had similar performance for the hemorrhage-related outcomes assessed compared to manual volumes and expert consensus grading. The results suggest promising avenues for automated HTX volumetry in research and clinical care.

Keywords

Deep learning; nnU-Net; Computed tomography; Trauma; Hemothorax; Segmentation; Quantitative visualization; Artificial intelligence machine learning

Introduction

Chest trauma collectively results in a quarter of all fatal traumatic injuries annually, and hemorrhage into the thoracic cavity is an important proximate cause of early morbidity and mortality [1–7]. Since admission computed tomography (CT) is rapid and allows concurrent monitoring by the trauma team, the modality is routinely employed in patients sufficiently hemodynamically stable for transfer to a trauma bay-adjacent scanner. Hemothorax reflects the total amount of bleeding into the chest cavity from the time of injury and manifests as dependently layering accumulations of blood within the pleural space [3]. Hemothorax is difficult to quantify with chest x-ray, and CT is generally superior for characterization of thoracic injury [2, 8, 9]. The unobstructed anatomic detail afforded by CT opens the possibility for voxelwise quantification of hemothorax, and could facilitate objective personalized prognostication and decision support with respect to hemorrhage control; however, manual or semi-automated voxelwise measurement is labor-intensive and not feasible at the point of care [10, 11].

There is currently no universal and clinically validated system for grading hemothorax size on imaging, with coarse subjective descriptors such as “small” or “moderate” routinely used in the clinical setting [5, 12, 13]. Because the thoracic cavity can accommodate accumulation of a large volume of blood, traumatic hemothorax may result in exsanguination and mortality. Massive transfusion is an important life-saving intervention and there is an unmet need for objective criteria to guide the transfusion strategy in patients with chest trauma and hemothorax [14, 15].

Automated methods for voxelwise segmentation and quantification have been described for features of hemorrhage in other body cavities, including extraperitoneal pelvic hematomas

[16, 17] and hemoperitoneum [18]. Segmentation methods must be robust at multiple scales given the wide range of hemothorax volumes encountered.

The purpose of this pilot study is to assess method validity of deep learning-based automated hemothorax segmentation and quantification and explore the potential clinical role of automated hemothorax volumetry for predicting need for massive transfusion and hemorrhage-related mortality.

Materials and methods

Study dataset

De-identified admission contrast-enhanced chest CT studies that were performed on arrival at a level I trauma center were retrospectively compiled into the study dataset as part of an IRB-approved HIPAA compliant study. Consecutive adult (age ≥ 18 years) patients between August 2016 and December 2018 who experienced blunt or penetrating thoracic trauma were queried for the keyword “hemothorax” within radiologic reports, returning 97 patients. Four patients were excluded because of lack of archived CT images or clinical data. Ten cases with incomplete imaging data or corrupted meta-data could not be successfully loaded into the segmentation software and were also excluded.

Labeling in all patients was performed using 3D slicer [19] (details provided below), and trace hemothoraces (< 50 mL, reflecting a small fraction of one unit of transfused blood) expected to be too small to impact haemorrhage-related outcomes were excluded a priori ($n = 6$). The final dataset consisted of 77 labeled CT studies. These patients’ charts were reviewed by a research assistant for hemorrhage-related outcomes that were used to generate a composite endpoint including massive transfusion (MT-, defined as at least 4 units of packed red blood cells (PRBC) in a 4-h period or at least 10 units of PRBC over 24 h [20, 21]); or in-hospital mortality (IHM), excluding fatalities from traumatic brain injury (head abbreviated injury severity score of ≥ 6).

Image acquisition

Thoracic CT studies were completed using either a dual-source 128-section (Siemens Force; Siemens, Erlangen, Germany) or 64-section (Brilliance; Philips Healthcare, Andover, MA) scanners located within the trauma center. The scanning protocol included injection of 100 mL of intravenous contrast (Omnipaque [iohexol; 350 mg of iodine per milliliter]; GE Healthcare, Chicago, IL) and image acquisition in the arterial phase at 120 kVp (Philips) or 120 kVp equivalent blend (Siemens) at 150–159 reference mAs. Studies were archived with 1.5–3-mm section thickness.

Qualitative grading

Blinded qualitative evaluation of hemothoraces was performed by consensus of two board-certified radiologists — a chest radiologist with over 15 years of experience and an emergency radiologist with 10 years of experience. In keeping with clinical practice, hemothorax size was graded in a gestalt fashion as small, moderate, or large. The following guideline was used to inform qualitative grading [5]: small hemothoraces have an estimated

volume less than 250 mL, with pleural thickness less than 1.5 cm; moderate hemothoraces have estimated volumes between 250 and 1000 mL, with pleural thickness between 1.5 and 4 cm; and large hemothoraces have a volume > 1000 mL with pleural thickness >4 cm.

Manual labeling

Each study was converted to Neuroimaging Informatics Technology Initiative (NifTI) file format and manually labeled by trained staff at the voxel level using the segment editor module in 3D slicer (version 4.10.2, www.slicer.org). The 3D ROI paint tool was employed with thresholding between -20 and 100 HU to account for image noise and beam hardening from adjacent ribs. Labels were created in the axial planes with further editing in the coronal and sagittal planes. A morphologic smoothing filter was uniformly applied to eliminate label artifact. All studies underwent further editing by a radiologist. Manual hemothorax volumes were calculated using voxel counting with the slicer quantification module.

High-level overview of the deep learning method

Convolutional neural networks with encoder and decoder architectures such as U-Net form the backbone of contemporary automated semantic segmentation methods [22]. At baseline, U-Net and similar fully convolutional neural networks (FCNs) involve gradual hierarchical increases in size of receptive fields in progressively deeper layers and are therefore suboptimal for segmentation of targets that vary greatly in scale [23]. Traumatic bleeding in the large body cavities, including pelvic hematoma [17], hemoperitoneum [18], and hemothorax, ranges widely in volumes and best results have been achieved using tailored multiscale solutions.

In recent years, numerous approaches have proliferated to address multiscale problems in cross-sectional imaging. Commonly, low-resolution networks that maximize global contextual awareness are combined with high resolution networks that produce fine detail segmentations, operating in a cascaded [17] or tandem [18, 24–26] fashion. nnU-Net, introduced in *Nature* in 2021 by Isensee et al. [27], is based on the principle that design choices in pre-processing and hyper-parameter tuning, including intensity thresholding, cropping, slice thickness resampling, data augmentation, batch size, and learning rate [28], have more impact on performance than do architectural variations. nnU-net is *self-configuring*, automatically determining the aforementioned design choices in a data-driven manner derived from best results on over 20 public cross-sectional imaging datasets. The selected design choices produce a unique *dataset and pipeline fingerprint* for the new data at hand. The rules-based method for determining the fingerprint is based on features which include median number of positive voxels, spacing and size anisotropy, number of training cases, and imaging modality. It is shorthand for “no new U-Net” as it employs generic 2D, 3D, and cascaded 3D U-Net architectures in fivefold cross-validation using a Dice similarity coefficient loss function. Ensembled results of the three methods are also determined, and the best-performing method (2D, 3D, cascaded 3D, or ensembled) is selected.

In spite of the simplicity of the backbone architectures, nnU-Net surpasses most existing highly-specialized solutions for 53 different segmentation tasks from international

biomedical image segmentation challenges, and serves as a state-of-the-art out-of-the-box automated segmentation tool that is robust for multiscale problems [27].

Statistical analyses

nnU-Net training and validation were performed with an RTX 3090 NVIDIA GPU with 24 GB of RAM using Python version 3.7.11. The best performing model was selected based on highest mean Dice similarity coefficient (DSC) and volume similarity [29]. Correlation and agreement between automated and manual volumes were further interrogated using intraclass correlation coefficient (ICC), Pearson's r , and mean bias derived from Bland–Altman analysis.

Median automated volumes were compared by composite outcome using the Mann–Whitney U test. Automated volumes were compared to manual volumes and qualitative consensus expert grading for the composite outcome of massive transfusion and in-hospital mortality using the area under the ROC curve (AUC). Statistical comparisons were performed using the DeLong method. Receiver operating characteristic curve (ROC) analysis was also used to determine volume thresholds resulting in 90% sensitivity and 90% specificity.

Results

Ensembled nnU-Net was selected for further analysis as it demonstrated the best overall combined performance using overlap and volume-based segmentation metrics, achieving high mean volume similarity of 0.91 — equivalent to high-resolution 3D and cascaded 3D methods — and highest mean DSC of 0.75 (Table 1). The distribution of manual and automated (ensembled nnU-Net) hemothorax volumes ranged between 35 and 1503 mL (manual) and 34 and 1451 mL (auto), with median volumes of 178 mL (manual) and 185 mL (auto). Standard quantitative imaging biomarker evaluation metrics [30] comparing manual to automated volumes included an ICC of 0.92 (95%CI: 0.88 to 0.95), and Pearson's r of 0.93 (95%CI: 0.90 to 0.96). Pearson's r values ≥ 0.80 are indicative of strong correlation and ICC values ≥ 0.75 are considered to indicate excellent agreement [31]. The automated method had a mean overmeasurement bias of 1.7 mL compared to manual volumes.

Patient baseline characteristics are provided in Table 2. Among the 77 patients included in the study, the median age was 50 years (interquartile range [IQR] 36–62), and 74% were male. Mean Abbreviated Injury Score (AIS) score was 3.2 (SD: 0.6), and mean injury severity score (ISS) was 24.2 (SD: 10.3). Four patients (5%) suffered penetrating injuries, and the other 73 patients (95%) experienced blunt injuries. Twelve patients (16%) underwent massive transfusion, and 8 patients (10%) died during their hospital admission. A total of 15 patients (18%) had either or both outcomes.

Automated volumes were significantly higher in patients with the composite outcome (massive transfusion or in-hospital mortality) than those without (341.1 mL versus 158.4 mL, $p = 0.004$). Examples of automated segmentation results for a range of HTX volumes are illustrated in Fig. 1.

The AUCs of traumatic automated HTX volumes, manual volumes, and qualitative expert grading for the composite outcome were 0.74 (95%CI: 0.58–0.91), 0.76 (95%CI: 0.58–0.93), and 0.76 (95%CI: 0.62–0.90) ($p = 0.93$ for testing the null hypothesis that the three AUCs are equal). A traumatic automated HTX volume cut-off of 81 mL predicted the outcome with 93% sensitivity and 11% specificity and a volume cut-off 341 mL with 90% specificity and 53% sensitivity.

Discussion

Persistent hemorrhage into the pleural space may be fatal, and clinical decisions in managing traumatic hemothorax early in the course of treatment relate to addressing the origin of hemorrhage and transfusing to compensate for blood loss until the source can be controlled [1, 3, 13–15]. CT is integral to the evaluation of patients after severe thoracic trauma and is an important component of trauma care algorithms where intracavitary hemorrhage is suspected [4, 6]. Despite advancements in image technology that enables rapid acquisition of anatomic data, point of care hemothorax volume estimation is typically achieved with coarse subjective assessment. In comparison to broad, qualitative categorizations that are traditionally given in radiology reports, automated quantitative visualization of hemothorax could potentially provide objective information for prognosticating hemorrhage-related outcomes. Our automated hemothorax results demonstrate high method validity and outcome prediction on par with expert level qualitative grading undertaken in a controlled research setting. Automated hemothorax volumes that can be visually verified by any radiologist or trauma surgeon end-user are advantageous because they are rapid, consistent, and objective, not depending on experience level, subspecialization, reader fatigue, reading room distractions, or the lengthy study turnaround times of trauma CT exams, which can exceed 20–30 min [32–34]. By comparison, nnU-Net has inference times of under 2 min.

To engender trust among radiologists or trauma surgeons, hemothorax segmentation masks must conform to the boundaries of hemothorax while avoiding neighboring structures of similar attenuation values such as compressed lung. We achieved high-saliency interpretable visual results with corresponding high mean DSC values, volume similarity scores, correlation, and agreement using nnU-Net.

In this dataset, injury severity scores, and lactate levels were significantly higher in patients requiring massive transfusion or that expired from hemorrhage-related complications, and admission systolic blood pressure was significantly lower. This is likely accounted for by multicavitary hemorrhage, which is not addressed in this pilot study. Bleeding features on routine admission trauma CT are often leading indicators of hemodynamic collapse [4, 6, 35–37] and can soon be accounted for more holistically using similar deep learning-based quantitative visualization approaches. Proof-of-concept automated methods for quantitative visualization of pelvic hematoma volumes and hemoperitoneum are described [16–18], and comprehensive assessment of hemorrhage burden using automated hemothorax, hemoperitoneum, and pelvic hematoma segmentation is planned in future work scaled to larger datasets. Once containerized as software and deployed, high-throughput automated hemorrhage burden quantification tools will likely accelerate cross-institutional research

and may play an important role in objective decision-making in the early management of patients with torso hemorrhage.

Additional outcomes specific to hemothorax may also be of interest, including the risk of retained hemothorax and empyema [13]. Other manifestations of injury including lung contusion can impact the risk of developing traumatic acute respiratory distress syndrome (ARDS) [38], and in the future, automated multi-class voxelwise labeling and quantification of traumatic chest injuries on CT could facilitate prediction of a variety of outcomes based on multiple quantitative imaging parameters.

The computer vision aspect of our work is currently limited to the canonical task of segmentation. The method is currently not intended as a screening or workflow prioritization tool. A pipeline that screens for, segments, and measures hemothorax and hemorrhage in other body cavities autonomously in the clinical background and automatically notifies practitioners of critical quantitative results, as RAPID software has achieved for stroke imaging [39], is likely feasible but will require initial deep learning object detection steps with algorithms such as mask R-CNN preceding segmentation tasks [40, 41].

Finally, our pilot study used a CT dataset from a single institution. A large multi-institutional dataset will eventually be needed to ensure robust performance in the face of domain shift from scanner and scanning protocol heterogeneity and variable patient populations [42, 43] to achieve a technology readiness level suitable for deployment across institutions.

In conclusion, in this pilot study, we found that a state-of-the-art deep learning method resulted in objective and interpretable automated quantitative visualization of traumatic hemothorax, with hemorrhage-related outcome prediction comparable to expert qualitative grading. Future work will entail combining the method with similar pelvic hematoma and hemoperitoneum pipelines for comprehensive torso hemorrhage burden quantification, using initial detection algorithms to facilitate full-throughput autonomous function in the clinical background, and scaling to large multicenter datasets to ensure robustness.

Funding

1. NIH K08 EB027141-01A1 (PI: David Dreizin, MD).
2. Accelerated Translational Incubator Pilot (ATIP) award, University of Maryland (PI: David Dreizin, MD).

Data availability

Labeled dataset has not been made available at this time.

References

1. Mowery NT, Gunter OL, Collier BR, Jose' J D Jr, Haut E, Hildreth A, Holevar M, Mayberry J, Streib E (2011) Practice management guidelines for management of hemothorax and occult pneumothorax. *J Trauma Acute Care Surg* 70(2):510–518
2. Sangster GP, González-Beicos A, Carbo AI, Heldmann MG, Ibrahim H, Carrascosa P, Nazar M, D'Agostino HB (2007) Blunt traumatic injuries of the lung parenchyma, pleura, thoracic wall,

and intrathoracic airways: multidetector computer tomography imaging findings. *Emerg Radiol* 14(5):297–310 [PubMed: 17623115]

3. Mahmood I, Abdelrahman H, Al-Hassani A, Nabir S, Sebastian M, Maull K (2011) Clinical management of occult hemothorax: a prospective study of 81 patients. *Am J Surg* 201(6):766–769 [PubMed: 21741510]
4. Dreizin D, Munera F (2012) Blunt polytrauma: evaluation with 64-section whole-body CT angiography. *Radiographics* 32(3):609–631 [PubMed: 22582350]
5. Bilello JF, Davis JW, Lemaster DM (2005) Occult traumatic hemothorax: when can sleeping dogs lie? *Am J Surg* 190(6):844–848
6. Dreizin D, Munera F (2015) Multidetector CT for penetrating torso trauma: state of the art. *Radiology* 277(2):338–355 [PubMed: 26492022]
7. Demetri L, Aguilar MMM, Bohnen JD, Whitesell R, Yeh DD, King D, de Moya M (2018) Is observation for traumatic hemothorax safe? *J Trauma Acute Care Surg* 84(3):454–458 [PubMed: 29298241]
8. Stafford RE, Linn J, Washington L (2006) Incidence and management of occult hemothoraces. *Am J Surg* 192(6):722–726 [PubMed: 17161082]
9. Marts B, Durham R, Shapiro M, Mazuski JE, Zuckerman D, Sundaram M, Luchtefeld WB (1994) Computed tomography in the diagnosis of blunt thoracic injury. *Am J Surg* 168(6):688–692 [PubMed: 7978020]
10. Dreizin D, Bodanapally UK, Neerchal N, Tirada N, Patlas M, Herskovits E (2016) Volumetric analysis of pelvic hematomas after blunt trauma using semi-automated seeded region growing segmentation: a method validation study. *Abdom Radiol* 41(11):2203–2208
11. Battey TW, Dreizin D, Bodanapally UK, Wnorowski A, Issa G, Iacco A, Chiu W (2019) A comparison of segmented abdomin-opelvic fluid volumes with conventional CT signs of abdominal compartment syndrome in a trauma population. *Abdom Radiol* 44(7):2648–2655
12. Moy MP, Levsky JM, Berko NS, Godelman A, Jain VR, Haramati LB (2013) A new, simple method for estimating pleural effusion size on CT scans. *Chest* 143(4):1054–1059 [PubMed: 23632863]
13. Meyer DM (2007) Hemothorax related to trauma. *Thorac Cardiovasc Surg* 17(1):47–55
14. Huang J-F, Hsu C-P, Fu C-Y, Yang C-HO, Cheng C-T, Liao C-H, Kuo I-M, Hsieh C-H (2021) Is massive hemothorax still an absolute indication for operation in blunt trauma? *Injury* 52(2):225–30 [PubMed: 33386159]
15. Ma H-S, Ma J-H, Xue F-L, Fu X-N, Zhang N (2016) Clinical analysis of thoracoscopic surgery combined with intraoperative autologous blood transfusion in the treatment of traumatic hemothorax. *Chin J Traumatol* 19(6):371–372 [PubMed: 28088945]
16. Dreizin D, Zhou Y, Chen T, Li G, Yuille AL, McLenithan A, Morrison JJ (2020) Deep learning-based quantitative visualization and measurement of extraperitoneal hematoma volumes in patients with pelvic fractures: potential role in personalized forecasting and decision support. *J Trauma Acute Care Surg* 88(3):425–433 [PubMed: 32107356]
17. Dreizin D, Zhou Y, Zhang Y, Tirada N, Yuille AL (2020) Performance of a deep learning algorithm for automated segmentation and quantification of traumatic pelvic hematomas on CT. *J Digit Imaging* 33(1):243–251 [PubMed: 31172331]
18. Dreizin D, Zhou Y, Fu S, Wang Y, Li G, Champ K, Siegel E, Wang Z, Chen T, Yuille AL (2020) A multiscale deep learning method for quantitative visualization of traumatic hemoperitoneum at CT: assessment of feasibility and comparison with subjective categorical estimation. *Radiol: Artif Intell* 2(6):e190220
19. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J-C, Pujol S, Bauer C, Jennings D, Fennessy F, Sonka M (2012) 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn Reson Imaging* 30(9):1323–1341 [PubMed: 22770690]
20. Mackenzie CF, Wang Y, Hu PF, Chen S-Y, Chen HH, Hagegeorge G, Stansbury LG, Shackelford S, Group OS (2014) Automated prediction of early blood transfusion and mortality in trauma patients. *J Trauma Acute Care Surg* 76(6):1379–85 [PubMed: 24854304]
21. Parimi N, Hu PF, Mackenzie CF, Yang S, Bartlett ST, Scalea TM, Stein DM (2016) Automated continuous vital signs predict use of uncrossed matched blood and massive transfusion following

- trauma. *J Trauma Acute Care Surg* 80(6):897–906. 10.1097/TA.0000000000001047 [PubMed: 27027555]
22. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, Van Der Laak JA, Van Ginneken B, Sánchez CI (2017) A survey on deep learning in medical image analysis. *Med Image Anal* 42:60–88 [PubMed: 28778026]
 23. Fang X, Yan P (2020) Multi-organ segmentation over partially labeled datasets with multi-scale feature abstraction. *IEEE Trans Med Imaging* 39(11):3619–3629 [PubMed: 32746108]
 24. Dreizin D, Chen T, Liang Y, Zhou Y, Paes F, Wang Y, Yuille AL, Roth P, Champ K, Li G, McLenithan A, Morrison JJ (2021) Added value of deep learning-based liver parenchymal CT volumetry for predicting major arterial injury after blunt hepatic trauma: a decision tree analysis. *Abdom Radiol (NY)* 46(6):2556–2566. 10.1007/s00261-020-02892-x [PubMed: 33469691]
 25. Zhou Y, Dreizin D, Li Y, Zhang Z, Wang Y, Yuille A (2019) Multi-scale attentional network for multi-focal segmentation of active bleed after pelvic fractures. In *International Workshop on Machine Learning in Medical Imaging*. Springer, Cham, pp 461–469
 26. Zhou Y, Dreizin D, Wang Y, Liu F, Shen W, Yuille AL (2022) External attention assisted multi-phase splenic vascular injury segmentation with limited data. *IEEE Trans Med Imaging* 41(6):1346–1357. 10.1109/TMI.2021.3139637 [PubMed: 34968179]
 27. Isensee F, Jaeger PF, Kohl SA, Petersen J, Maier-Hein KH (2021) nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat Methods* 18(2):203–211 [PubMed: 33288961]
 28. Chartrand G, Cheng PM, Vorontsov E, Drozdal M, Turcotte S, Pal CJ, Kadoury S, Tang A (2017) Deep learning: a primer for radiologists. *Radiographics* 37(7):2113–2131 [PubMed: 29131760]
 29. Maier-Hein L, Reinke A, Christodoulou E, Glocker B, Godau P, Isensee F, Kleesiek J, Kozubek M, Reyes M, Riegler MA (2022) Metrics reloaded: pitfalls and recommendations for image analysis validation. *arXiv preprint arXiv:2206.01653*
 30. Kessler LG, Barnhart HX, Buckler AJ, Choudhury KR, Kondratovich MV, Toledano A, Guimaraes AR, Filice R, Zhang Z, Sullivan DC (2015) The emerging science of quantitative imaging biomarkers terminology and definitions for scientific studies and regulatory submissions. *Stat Methods Med Res* 24(1):9–26 [PubMed: 24919826]
 31. Zou KH, Tuncali K, Silverman SG (2003) Correlation and simple linear regression. *Radiology* 227(3):617–628 [PubMed: 12773666]
 32. Hunter TB, Taljanovic MS, Krupinski E, Ovitt T, Stubbs AY (2007) Academic radiologists' on-call and late-evening duties. *J Am Coll Radiol* 4(10):716–719 [PubMed: 17903757]
 33. Brant-Zawadzki MN (2007) Special focus—outsourcing after hours radiology: one point of view—outsourcing night call. *J Am Coll Radiol* 4(10):672–674 [PubMed: 17903747]
 34. Banaste N, Caurier B, Bratan F, Bergerot J-F, Thomson V, Millet I (2018) Whole-body CT in patients with multiple traumas: factors leading to missed injury. *Radiology* 289(2):374–383 [PubMed: 30084754]
 35. Dreizin D, Bodanapally U, Boscak A, Tirada N, Issa G, Nascone JW, Bivona L, Mascarenhas D, O'Toole RV, Nixon E (2018) CT prediction model for major arterial injury after blunt pelvic ring disruption. *Radiology* 287(3):1061–1069 [PubMed: 29558295]
 36. Borrer W, Gaski GE, Steenburg S (2019) Abdominopelvic bleed rate on admission CT correlates with mortality and transfusion needs in the setting of blunt pelvic fractures: a single institution pilot study. *Emerg Radiol* 26(1):37–44 [PubMed: 30259226]
 37. Dreizin D, Champ K, Dattwyler M, Bodanapally U, Smith EB, Li G, Singh R, Wang Z, Liang Y (2022) Blunt splenic injury in adults: association between volumetric quantitative CT parameters and intervention. *J Trauma Acute Care Surg*. 10.1097/TA.0000000000003684
 38. Miller PR, Croce MA, Bee TK, Qaisi WG, Smith CP, Collins GL, Fabian TC (2001) ARDS after pulmonary contusion: accurate measurement of contusion volume identifies high-risk patients. *J Trauma Acute Care Surg* 51(2):223–230
 39. Laughlin B, Chan A, Tai WA, Moftakhar P (2019) RAPID automated CT perfusion in clinical practice. *Pract Neurol* 2019:41–55
 40. Zhang T, Song Z, Yang J, Zhang X, Wei J (2021) Cerebral hemorrhage recognition based on Mask R-CNN network. *Sens Imaging* 22(1):1–16

41. Zapaishchykova A, Dreizin D, Li Z, Wu JY, Faghihroohi S, Unberath M (2021) An interpretable approach to automated severity scoring in pelvic trauma. In International Conference on Medical Image Computing and Computer-Assisted Interventio. Springer, Cham, pp 424–433
42. Chen H, Gomez C, Huang C-M, Unberath M (2021) INTRPRT: a systematic review of and guidelines for designing and validating transparent AI in medical image analysis. arXiv preprint arXiv:2112.12596
43. Vlontzos A, Rueckert D, Kainz B (2022) A review of causality for learning algorithms in medical image analysis. arXiv preprint arXiv:2206.05498

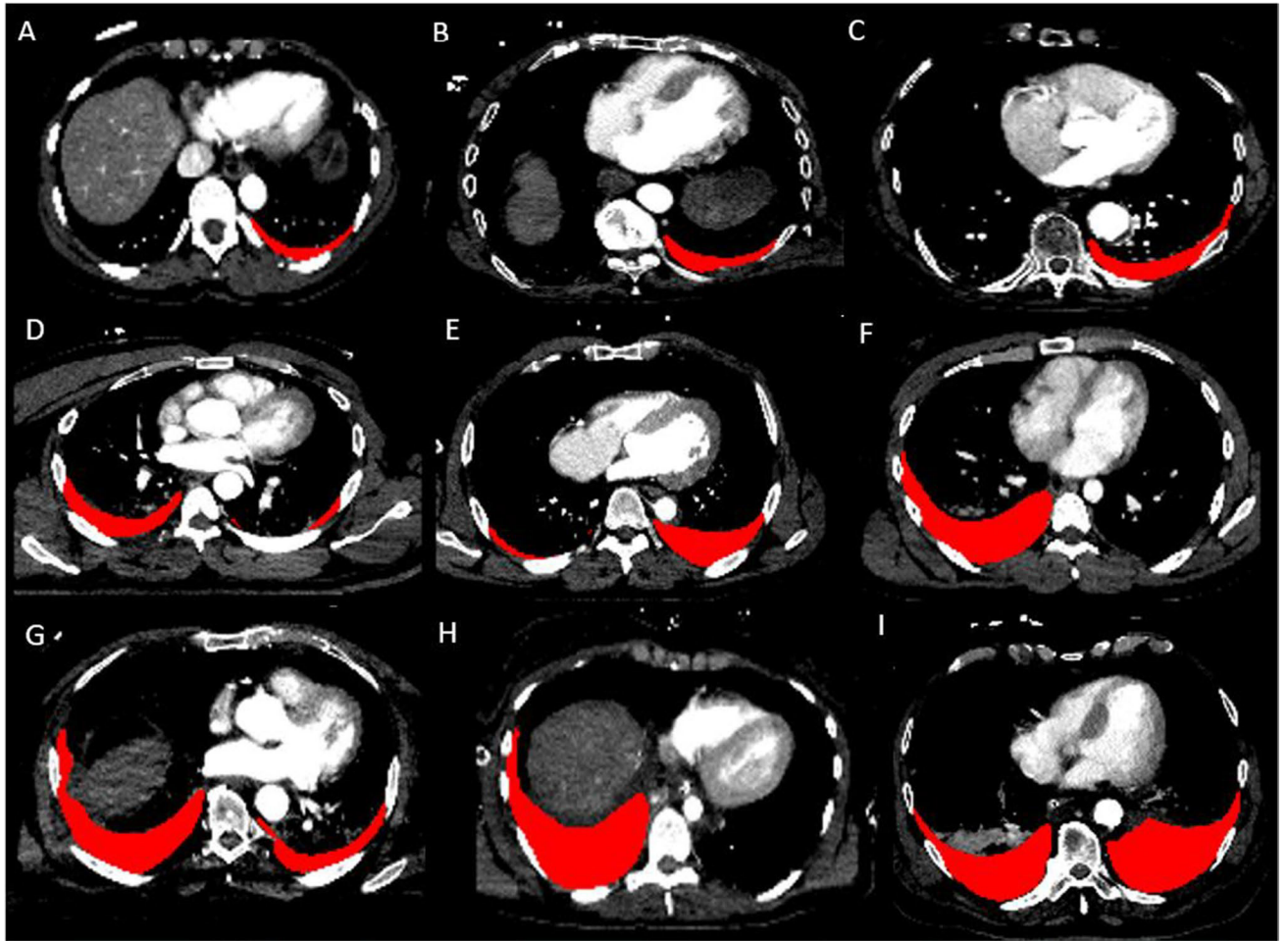


Fig. 1. Automated results using nnU-Net are shown for a wide range of hemothorax volumes (A 96.6 mL, B 147.8 mL, C 185.4 mL, D 287.3 mL, E 419.6 mL, F 597.4 mL, G 714.0 mL, H 898.6 mL, I 1450.6 mL)

Table 1

Performance metrics for nnU-Net

nnU-Net component network	Mean DSC (\pm SD)	Volume similarity score (\pm SD)
2D U-Net	0.65 (\pm 0.14)	0.87 (\pm 0.13)
3D U-Net (high resolution)	0.74 (\pm 0.12)	0.91 (\pm 0.10)
3D U-Net (cascaded)	0.74 (\pm 0.12)	0.91 (\pm 0.10)
Ensembled nnU-Net	0.75 (\pm 0.12)	0.91 (\pm 0.10)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Demographic and clinical characteristics

	Total cohort <i>n</i> = 77		<i>p</i> value
	Yes <i>n</i> = 15	No <i>n</i> = 62	
autoHTX vol (median [IQR])	185.5 [118.0, 277.8]	341.1 [211.2, 485.6]	158.4 [114.4, 263.1] 0.004
Age (median [IQR])	50 [36–62]	62 [36–76]	49 [36, 60] 0.35
Gender (<i>n</i> (%))			
Male	57 (74)	11 (73)	46 (74) 1
Female	20 (26)	4 (27)	16 (26) 0.12
HR (mean [SD])	94 [19]	100 [19]	94 [19] <0.001
SBP (mean [SD])	136 [33]	103 [23]	144 [30] 0.02
Shock (SBP < 90) (<i>n</i> (%))	6 (8)	4 (27)	2 (3) <0.001
Lactate (mean [SD])	3.8 [3.0]	7.6 [4.5]	2.9 [1.4] 0.58
Mechanism (<i>n</i> (%))			
Blunt	73 (95)	15 (100)	58 (93) <0.001
Penetrating	4 (5)	0 (0)	4 (7) 0.01
Chest AIS (mean [SD])	3.2 [0.6]	3.8 [0.7]	3.1 [0.6] <0.001
ISS (mean [SD])	24.2 [10.3]	31.4 [10.7]	22.5 [9.5] <0.001
Qualitative grade (<i>n</i> (%))			
Mild	28 (36)	2 (13)	26 (42) 0.01
Moderate	39 (51)	6 (40)	33 (53) <0.001
Severe	10 (13)	7 (47)	3 (5) <0.001

Two-tailed *p*-values less than 0.05 indicate statistical significance

Abbreviations: *autoHTX vol*, automated hemothorax volumes; *IQR*, interquartile range; *HR*, heart rate; *SBP*, admission systolic blood pressure; *AIS*, abbreviated injury scale; *ISS*, Injury Severity Score