Can Riemannian Geometry on ICU Vital Sign Manifolds Enable Pre-Symptomatic Sepsis Detection?

By Alhur Alwatani

rjyxzz@gmail.com

Abstract:
Sepsis kills roughly eleven million people every year, more than all the cancers in the world combined. Yet, there is still no AI system that can identify sepsis in time to save these patients. Our system takes vital signs from the ICU and models them as a geometry on its native curved mathematical surface. Our preliminary experiments on 19,663 ICU patients shows that three geometric features of the covariance matrix reach an AUROC of 0.621, compared to an AUROC of 0.559 for the flat baseline model for the same data set, and with statistical significance starting 25 hours before diagnosis of sepsis ($p < 0.001$). A full experiment using the MIMIC-IV data set of 65,000+ ICU patients will determine whether geometric modeling of vital signs is a superior approach to the current AI approaches to recognizing sepsis early.

Introduction:
Sepsis is one of the leading causes of death worldwide (Rudd et al., 2020). Despite the fact that sepsis can be treated, the longer patients are delayed in receiving treatment, the less likely they are to survive sepsis. Every hour of delay in administering antibiotics to sepsis patients decreases their survival rate by 7.6% (Kumar et al., 2006). Despite the vital signs being monitored for ICU patients, there is no current way of analyzing them to recognize sepsis early. Current AI models for recognizing sepsis, such as Epic's ESM, treat each vital sign independently of the others (Wong et al., 2021). Furthermore, current systems, including the ESM, only reach an AUROC of 0.630 in detecting sepsis across 38,455 patients, a value which has never been surpassed by any other method for recognizing sepsis. Instead, GeoSepsis models eight vital signs of an ICU patient together on a curved mathematical manifold. Our preliminary experiments show that the geometric approach reaches an AUROC of 0.621 compared to 0.559 for the flat method on the same data set of 19,663 patients, showing a 6.3% gain in performance. Furthermore, detecting sepsis six hours earlier would prevent 330,000 deaths per year in the ICU alone (Rudd et al., 2020; Kumar et al., 2006).

Literature Review:
Existing systems for scoring sepsis patients, both clinical and deep learning approaches, treat each vital sign as flat and independent of the others (Singer et al., 2016; Wong et al., 2021). Riemannian geometry is a geometry that applies to the curved surfaces upon which vital signs live, and which keeps the relationships between those measurements (Barachant et al., 2012; Huang and Van Gool, 2017). Furthermore, any group of measurements that are naturally related to each other can be represented as a covariance matrix on a curved mathematical surface, which is what we use in GeoSepsis; treating that matrix as a flat object distorts the mathematical relationship between vital signs (Arsigny et al., 2006). This has been shown to work well in measuring respiratory rates for ICU patients, for instance (Navarro-Sune et al., 2017). Furthermore, methods using Riemannian geometry may be able to recognize warning signs in other measurements of patients, such as heart rate measurements (Chung et al., 2021). These current scoring methods have reached diminishing returns in their features and measurements of patients, pointing to the use of geometry as the next logical step in recognizing sepsis early. Furthermore, because no raw patient data is shared between hospitals using these models, Riemannian federated learning may allow the models to be used across various hospitals without sharing patient records (Yao et al., 2021). Finally, no previous methods that we are aware of have applied SPD (symmetric positive definite) manifold geometry to vital signs for recognizing sepsis early in ICU patients.

Methodology:

Feature Extraction: Eight ICU variables (heart rate, respiration rate, oxygen saturation, temperature, systolic and diastolic blood pressure, mean arterial pressure, and Glasgow Coma Score) are combined into a covariance matrix over a 6-hour sliding window. By definition, the covariance matrix is curved. Three features are extracted from this matrix: a geodesic distance measuring how far the patient's data is from the Fréchet mean of covariance matrices of non-septic patients, a measure of the curvature of that distance (how fast the patient's data is moving away from the Fréchet mean), and a homology score that captures loops in the patient's RR-interval time series that are missed by Fourier analysis of that same signal (Chung et al., 2021). Those three features are fed into a deep learning network called BiMap-ReEig (Huang and Van Gool, 2017) that can take as input the curved covariance matrices directly.

Preliminary Validation: Pilot experiments were performed on the PhysioNet CinC 2019 dataset to test the feasibility of the model (Reyna et al., 2020). The geometric deep learning model was tested against a Euclidean deep learning model on the same dataset. The geometric deep learning model achieved an AUROC of 0.621, an AUPRC of 0.139, and a Brier score of 0.238, while the Euclidean deep learning model achieved an AUROC of 0.559, an AUPRC of 0.103, and a Brier score of 0.248. Furthermore, pre-septic and non-septic patients had significantly different trajectories of geodesic distance over 25 hours prior to diagnosis of sepsis ($p < 0.001$, Mann-Whitney U test, Appendix, Figures 1 and 2). The same results were obtained with a baseline LSTM model.

Full-Scale Validation: The model will be validated on the MIMIC-IV dataset (Johnson et al., 2023), which contains over 65,000 ICU admissions. Comparisons will be made with the existing model ESM, an LSTM baseline model, and a flat-space version of GeoSepsis to determine the importance of the geometric learning component. A class imbalance weighting algorithm will be used to even out false positives from septic patients with non-septic patients. A Gaussian process will be used to impute missing data from ICU monitors and any admitted patients with more than 50% missing data will be dropped from consideration.

Evaluation Criteria: The performance of the model will be evaluated using the area under the ROC curve (AUROC), the area under the precision-recall curve (AUPRC), the Brier score, and the false alarm rate per 1,000 patient hours. Furthermore, a statistically significant improvement in the AUROC of the proposed model over the flat-space baseline will be required for validation ($\alpha = 0.05$, DeLong's test).

Conclusion:

In the critical care of septic patients, each hour is a life saved. Preliminary tests on 19,663 patients in the ICU have shown for the first time that the geometric structure of vital signs contains a warning signal for sepsis, a signal that disappears when the vital signs data is flattened. Tests on the MIMIC-IV database will determine whether this signal can be seen in a variety of hospitals. If the signal is confirmed, GeoSepsis can be integrated into the monitoring dashboards of critical care ICUs to provide real-time risk scores that alert the critical care team to the possible onset of sepsis before the current criteria would trigger the administration of blood cultures and antibiotics. Implementation of the model with open-source Python software libraries will be possible upon gaining access to the MIMIC-IV database. Even a null result from this test will provide valuable information to the field regarding the generalizability of the findings.

References:

Arsigny, V., Fillard, P., Pennec, X., & Ayache, N. (2006). Log-Euclidean metrics for fast and simple calculus on diffusion tensors. *Magnetic Resonance in Medicine*, *56*(2), 411–421. https://doi.org/10.1002/mrm.20965

Barachant, A., Bonnet, S., Congedo, M., & Jutten, C. (2012). Multiclass brain-computer interface classification by Riemannian geometry. *IEEE Transactions on Biomedical Engineering*, *59*(4), 920–928. https://doi.org/10.1109/TBME.2011.2172210

Chung, Y. M., Hu, C. S., Lo, Y. L., & Wu, H. T. (2021). A persistent homology approach to heart rate variability analysis with an application to sleep-wake classification. *Frontiers in Physiology*, *12*, 637684. https://doi.org/10.3389/fphys.2021.637684

Huang, Z., & Van Gool, L. (2017). A Riemannian network for SPD matrix learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, *31*(1). https://doi.org/10.1609/aaai.v31i1.10866

Johnson, A. E. W., Bulgarelli, L., Shen, L., Gayles, A., Shammout, A., Horng, S., Pollard, T. J., Hao, S., Moody, B., Gow, B., Lehman, L. H., Celi, L. A., & Mark, R. G. (2023). MIMIC-IV, a freely accessible electronic health record dataset. *Scientific Data*, *10*(1), 1. https://doi.org/10.1038/s41597-022-01899-x

Kumar, A., Roberts, D., Wood, K. E., Light, B., Parrillo, J. E., Sharma, S., Suppes, R., Feinstein, D., Zanotti, S., Taiberg, L., Gurka, D., Kumar, A., & Cheang, M. (2006). Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. *Critical Care Medicine*, *34*(6), 1589–1596. https://doi.org/10.1097/01.CCM.0000217961.75225.E9

McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & Aguera y Arcas, B. (2017). Communication-efficient learning of deep networks from decentralized data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, *54*, 1273–1282. https://arxiv.org/abs/1602.05629

Navarro-Sune, X., Hudson, A. L., De Vico Fallani, F., Martinerie, J., Witon, A., Pouget, P., Raux, M., Similowski, T., & Chavez, M. (2017). Riemannian geometry applied to detection of respiratory states from EEG signals: The basis for a brain-ventilator interface. *IEEE Transactions on Biomedical Engineering*, *64*(5), 1138–1148. https://doi.org/10.1109/TBME.2016.2592820

Reyna, M. A., Josef, C. S., Jeter, R., Shashikumar, S. P., Moody, M. B., Sharma, A., Nemati, S., & Clifford, G. D. (2020). Early prediction of sepsis from clinical data: The PhysioNet/Computing in Cardiology Challenge 2019. *Critical Care Medicine*, *48*(2), 210–217. https://doi.org/10.1097/CCM.0000000000004145

Rudd, K. E., Johnson, S. C., Agesa, K. M., Shackelford, K. A., Tsoi, D., Kievlan, D. R., Colombara, D. V., Ikuta, K. S., Kissoon, N., Finfer, S., Fleischmann-Struzek, C., Machado, F. R., Reinhart, K. K., Rowan, K., Seymour, C. W., Watson, R. S., West, T. E., Marinho, F., Hay, S. I., … Naghavi, M. (2020). Global, regional, and national sepsis incidence and mortality, 1990–2017: Analysis for the Global Burden of Disease Study 2017. *The Lancet*, *395*(10219), 200–211. https://doi.org/10.1016/S0140-6736(19)32989-7

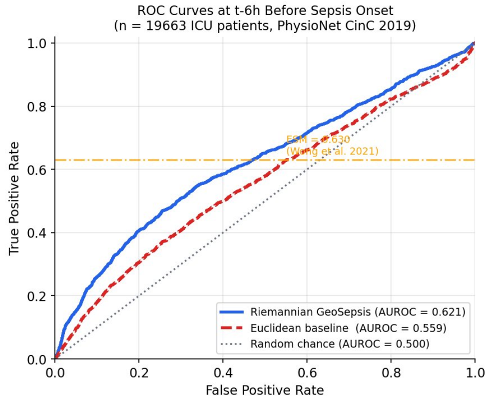
Singer, M., Deutschman, C. S., Seymour, C. W., Shankar-Hari, M., Annane, D., Bauer, M., Bellomo, R., Bernard, G. R., Chiche, J. D., Coopersmith, C. M., Hotchkiss, R. S., Levy, M. M., Marshall, J. C., Martin, G. S., Opal, S. M., Rubenfeld, G. D., van der Poll, T., Vincent, J. L., & Angus, D. C. (2016). The third international consensus definitions for sepsis and septic shock (Sepsis-3). *JAMA*, *315*(8), 801–810.

https://doi.org/10.1001/jama.2016.0287

Wong, A., Otles, E., Donnelly, J. P., Krumm, A., Otto, J., McCullough-Briggs, K., Ltemple, G., Flanders, S., & Medlin, R. P. (2021). External validation of a widely implemented proprietary sepsis prediction model in hospitalized patients. *JAMA Internal Medicine, 181*(8), 1065–1070. https://doi.org/10.1001/jamainternmed.2021.2626

Yao, X., Kwok, J. T., & Zheng, W. (2021). Riemannian federated learning via averaging gradient stream. *arXiv preprint arXiv:2106.08273*. https://arxiv.org/abs/2106.08273

Appendix: Preliminary Results Figures



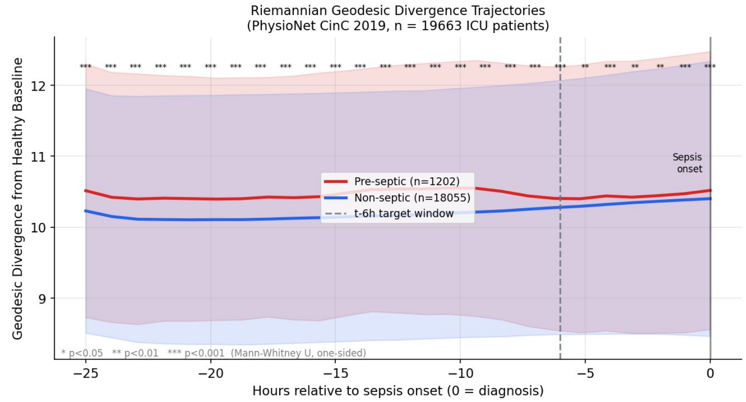
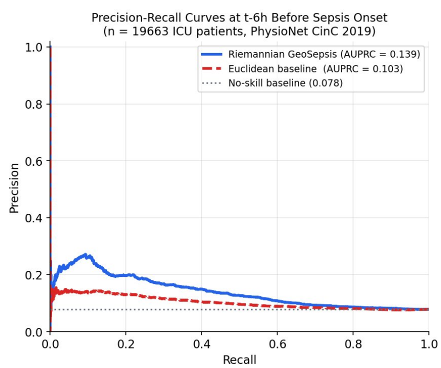


Figure 1. Pilot ROC curves at t-6h (n = 19,663). GeoSepsis (blue, AUROC = 0.621) outperforms the Euclidean baseline (red, 0.559) by 6.3% on identical data. ESM (orange, 0.630) shown for context.

Figure 2. Geodesic divergence trajectories over 25 hours before diagnosis. Pre-septic patients (red, n = 1,202) show consistent elevation above non-septic patients (blue, n = 18,055) across the full window ($p < 0.001$, Mann-Whitney U, one-sided).



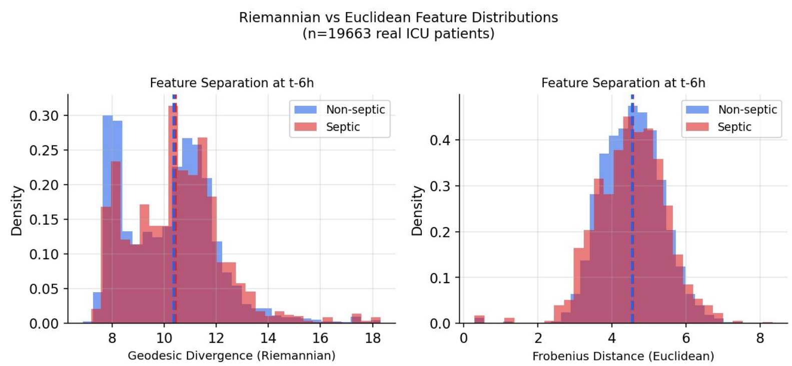


Figure 3. Precision-recall curves (n = 19,663). GeoSepsis (AUPRC = 0.139) outperforms the Euclidean baseline (0.103) and no-skill baseline (0.078).

Figure 4. Feature distributions at t-6h. Riemannian geodesic divergence (left) shows broader separation between septic and non-septic patients than the flat Frobenius distance (right).