BMJ Health & Care Informatics

Varying association of laboratory values with reference ranges and outcomes in critically ill patients: an analysis of data from five databases in four countries across Asia, Europe and North America

Haoran Xu,[1] Louis Agha-Mir-Salim ,[2,3] Zachary O'Brien,[4] Dora C Huang,[5] Peiyao Li,[6,7] Josep Gómez,[8,9] Xiaoli Liu,[2,10] Tongbo Liu,[11] Wesley Yeung,[2,12] Patrick Thoral,[13] Paul Elbers,[13] Zhengbo Zhang,[14] María Bodí Saera,[8,9] Leo Anthony Celi [2,15]

To cite: Xu H, Agha-Mir-Salim L, O'Brien Z, *et al*. Varying association of laboratory values with reference ranges and outcomes in critically ill patients: an analysis of data from five databases in four countries across Asia, Europe and North America. *BMJ Health Care Inform* 2021;**28**:e100419. doi:10.1136/bmjhci-2021-100419

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjhci-2021-100419).

HX and LA-M-S are joint first authors.

Received 27 May 2021
Accepted 17 September 2021

For numbered affiliations see end of article.

Correspondence to
Dr Louis Agha-Mir-Salim; mirsalim@mit.edu

ABSTRACT

Background Despite wide usage across all areas of medicine, it is uncertain how useful standard reference ranges of laboratory values are for critically ill patients.

Objectives The aim of this study is to assess the distributions of standard laboratory measurements in more than 330 selected intensive care units (ICUs) across the USA, Amsterdam, Beijing and Tarragona; compare differences and similarities across different geographical locations and evaluate how they may be associated with differences in length of stay (LOS) and mortality in the ICU.

Methods A multi-centre, retrospective, cross-sectional study of data from five databases for adult patients first admitted to an ICU between 2001 and 2019 was conducted. The included databases contained patient-level data regarding demographics, interventions, clinical outcomes and laboratory results. Kernel density estimation functions were applied to the distributions of laboratory tests, and the overlapping coefficient and Cohen standardised mean difference were used to quantify differences in these distributions.

Results The 259 382 patients studied across five databases in four countries showed a high degree of heterogeneity with regard to demographics, case mix, interventions and outcomes. A high level of divergence in the studied laboratory results (creatinine, haemoglobin, lactate, sodium) from the locally used reference ranges was observed, even when stratified by outcome.

Conclusion Standardised reference ranges have limited relevance to ICU patients across a range of geographies. The development of context-specific reference ranges, especially as it relates to clinical outcomes like LOS and mortality, may be more useful to clinicians.

Summary

What is already known?

- Laboratory results of critically ill patients are interpreted using reference ranges created on the basis of healthy outpatients.
- Correcting abnormal laboratory results to reference range standards can have beneficial or harmful effects.

What does this paper add?

- Laboratory results of critically ill patients often differ significantly from the reference range, even in those with the best clinical outcomes.
- Critically ill patients may require local, context-specific reference ranges for laboratory results to promote appropriate interpretation.

INTRODUCTION

The care of critically ill patients relies heavily on laboratory data—and, by extension, the laboratory reference ranges associated with them. However, these laboratory reference ranges are typically created by surveying healthy outpatients.[1] It remains unclear if these ranges are applicable to patients admitted to the intensive care unit (ICU).

Previous studies have shown that correcting abnormal values in critically ill patients, such as haemoglobin or glucose, to reference range standards may be harmful.[2–5] For example, clearly defined thresholds have been established for the initiation of packed red blood cell (PRBC) transfusions.[6–8] However, observational studies show that PRBCs are routinely administered at higher haemoglobin levels,[7 9 10] suggesting that clinicians may strive to correct laboratory values towards normality rather than adhere to evidence-based targets. As previously hypothesised, specific reference ranges tailored to scenarios and populations may be more meaningful, if these reference ranges are shown to relate to clinical outcome.[11–13]

Previous research considered whether the distributions of laboratory values for critically ill patients differed from reference ranges and if these differences were associated with outcomes. The single-centre, cross-sectional study found that laboratory values of ICU patients differed significantly from the reference range, even in those with the best clinical outcomes, suggesting that normal reference ranges may not apply to critically ill patients in reference or outcome.[14] This adds to ongoing discussions regarding the need to consider context in interpreting laboratory values, particularly in critical care settings, and advocates for further research into contextualising laboratory values.[11 15]

This study aims to expand on previous work by evaluating data from five ICU databases located across different continents to consider if similar patterns hold worldwide. In particular, this work aims to characterise how ICU laboratory values differ from typical reference ranges in ICUs across the USA, Netherlands, China and Spain, and determine if the relationship between laboratory values and patient outcomes varies across contexts.

METHODS

Design

We conducted an international, multi-centre, retrospective, cross-sectional study examining the most severely deranged (minimum or maximum as appropriate) laboratory results within the first 24 hours of a patient's first admission to the ICU.

Setting

We included all patients from five ICU databases: the Medical Information Mart for Intensive Care (MIMIC), the eICU Collaborative Research Database (eICU-CRD), the Amsterdam University Medical Center database (AUMCdb), the Chinese PLA General Hospital ICU database (PLAGH-ICUdb) and the Unitat de Cures Intensives de l'Hospital Joan XXIII database (UCIHJ23db). An overview of all databases is displayed in online supplemental table 1.

MIMIC contains data from the Beth Israel Deaconess Medical Center ICU, a tertiary hospital located in Boston, Massachusetts, USA, which comprises more than 70 beds with a broad case mix. We used data from the latest version available at the time of analysis, MIMIC-III, which contained granular data on more than 38 000 admissions between 2001 and 2012.[16] eICU-CRD contains similarly detailed, patient-level data on more than 200 000 admissions across 335 ICUs in the USA between 2014 and 2015.[17] AUMCdb is the first freely accessible European ICU database and contains data from the Amsterdam University Medical Center ICU, a mixed medical-surgical ICU with data on more than 20 000 patients including admissions between 2003 and 2016 (V.1.0.2).[18] The PLAGH-ICUdb integrates data from nine ICUs in the Chinese People's Liberation Army General Hospital in Beijing, China.[19] PLAGH-ICUdb includes data on more than 74 000 adult patients admitted between 2008 and 2019. Finally, the UCIHJ23db includes 4840 admissions between 2015 and 2019 from the Joan XXIII University Hospital in Tarragona, Spain.

Primary analysis

Our analysis was performed between February and May 2020. We included the first ICU admission of all adult patients from the five databases. Patients were stratified into those with the 'best' and 'worst' clinical outcomes. The best outcome group was defined as patients who survived the ICU admission and had an ICU length of stay (LOS) in the lowest quartile (shortest LOS). The worst outcome group was defined as those who died during the ICU admission.

For each patient, we extracted the most severely deranged laboratory values of commonly ordered investigations collected within the first 24 hours of ICU admission. The included investigations were maximum creatinine, minimum haemoglobin, maximum lactate and maximum sodium. No imputation was performed to replace missing data. We calculated the 95% CI for each investigation, stratified by the best and worst outcome patients within each database, and presented these as distribution plots. The locally used normal reference range for each investigation was added to these plots, to allow a visual assessment of the variance between these reference ranges and patient outcomes. We compared the difference in laboratory result distributions between the best and worst outcome groups by calculating the degree of overlap and divergence.

Statistical analysis

Data extraction was performed using SQL. Statistical analysis was then conducted using R and Python. The queries and code used for analyses were uploaded to a public GitHub repository.[20] Kernel density estimation plots were used to present the distribution of laboratory results. To then quantify the difference in distribution between best and worst outcome groups, we calculated the overlapping coefficient (OVL) and the Cohen standardised mean difference (SMD), as have been used for this purpose previously.[14 21 22] OVL quantifies the overlap of two distributions, with an OVL of 1 representing complete overlap and an OVL of 0 representing no overlap. SMD describes the difference in group means, relative to the variability observed within each group. The SMD value represents the divergence between groups in SD. An SMD of 0 indicates no difference in the means of the two groups; less than 0.2 is considered a small effect size, 0.2 to 0.8 a moderate effect size and greater than 0.8 a large effect size.[21]

Given the large sample size included in our analysis, tests of statistical significance were not performed, as it was anticipated that even very small and clinically irrelevant differences between groups would demonstrate statistical significance and may consequently have undue importance assigned to them.

RESULTS

Patients

Our study population included a total of 259 382 patients from five databases (MIMIC n=38 508, eICU-CRD n=132 994, PLAGH-ICUdb n=63 515, AUMCdb n=20 127 and UCIHJ23db n=4238). Substantial heterogeneity existed across the databases in patient demographics, the interventions they received and their clinical outcomes, as displayed in table 1. Notably, the proportion of patients who were admitted electively varied widely from 4.2% in UCIHJ23db to 71.7% in AUMCdb. By extension, the case mix also varied greatly, as reflected by the proportion of patients admitted following cardiac surgery (0.0% in UCIHJ23db vs 35.0% in AUMCdb). The interventions that patients received differed across databases, most appreciably in the delivery of mechanical ventilation and the administration of intravenous crystalloids, colloids and PRBCs.

The proportion of patients who died in the ICU ranged from 5.73% in eICU-CRD to 14.61% in UCIHJ23db. Similarly, the median ICU LOS ranged from 25 (20–73) hours in AUMCdb to 95 (46–173) hours in PLAGH-ICUdb.

Laboratory results

The IQR and median values of the most severely deranged measured laboratory investigations, for creatinine (maximum), haemoglobin (minimum), lactate (maximum) and sodium (maximum), stratified by database and patient outcomes, are displayed in online supplemental table 2. The locally used normal reference ranges for each database are also reported.

Regarding the distribution of investigation results and their corresponding reference ranges, the sodium measurements of best outcome patients consistently fell within the corresponding normal range (online supplemental figure 1), though other laboratory results did so variably. While the upper margin of the creatinine reference range includes the vast majority of best outcome patients from the PLAGH-ICUdb and UCIHJ23db, increasing proportions of best outcome patients had creatinine values beyond the upper margin in AUMCdb, MIMIC and eICU-CRD (online supplemental figure 2). The distribution of haemoglobin results shows that the majority of patients tended to record values below the lower margin of their local reference range across all databases, irrespective of whether they had best or worst outcomes (figure 1). Similarly, the distribution of lactate measurements indicates that a substantial proportion of patients with best outcomes had a measured lactate above the upper margin of local reference ranges, particularly in MIMIC and eICU-CRD (figure 2).

The 95% CIs for each laboratory value, stratified by best and worst clinical outcome group and by database, are reported in table 2. Overlapping and divergence coefficients are reported in table 3 and summarise the degree to which the distribution of laboratory results differed between best and worst outcome patients.

The best and worst outcome patients in the UCIHJ23db demonstrated the greatest overlap in the distribution of both creatinine (OVL=0.67, SMD=−0.46) and haemoglobin (OVL=0.86, SMD=0.32), while those from the PLAGH-ICUdb demonstrated the least overlap in the distribution of these laboratory results (creatinine OVL=0.48, SMD=−0.92 and haemoglobin OVL=0.67, SMD=0.8) (online supplemental figure 2 and figure 1).

Best and worst outcome patients from MIMIC demonstrated the greatest overlap in the distribution of highest measured lactate (OVL=0.65, SMD=−0.65), while those from AUMCdb demonstrated the least overlap (OVL=0.47, SMD=−1.01) (figure 2). AUMCdb also demonstrated the least overlap in highest measured sodium between best and worst outcome patients (OVL=0.67, SMD=−0.74), while the remaining databases consistently demonstrated OVL of approximately 0.75 for sodium measurements (online supplemental figure 1).

Overall, the mean overlap between best and worst patients across databases was greatest for measurements of haemoglobin (OVL=0.79) and lowest for measurements of lactate (OVL=0.45).

DISCUSSION

Differences between the most severely deranged laboratory results of patients admitted to the ICU and locally used normal reference ranges were observed in every database studied. These differences persisted even when comparing those patients with the best outcomes against the normal reference range.

In addition, among the databases, differences in the degree of overlap between best and worst group laboratory distributions were observed, which may represent variability in case mix and therapies applied, and/or imply variable discriminatory function among laboratory values based on region. Our findings build on the single-centre work by Tyler *et al*[14] by replicating similar observations across different contexts and geographies. They further support the need to consider context in reacting to abnormal laboratory values, as correcting abnormal values may not always be beneficial or benign.[5 11]

The differences observed between the reference range and selected ICU values across all five databases suggest that normal reference ranges are not useful in managing critically unwell patients. For instance, the haemoglobin results of most ICU patients fell outside normal reference ranges, irrespective of whether they had the best or worst outcome (figure 1). Patients with the best outcomes (here, ICU survival and shortest LOS) would be expected to have laboratory results that more closely align with the reference range, while those who die in the ICU should have results which are significantly worse. This is based on the assumption that the further patients' results deviate from the reference range, the more severely deranged their physiology and the more likely they are to have a poor clinical outcome. However, given the difference observed between the reference range and best outcome

Table 1 Baseline characteristics, interventions and outcomes of study population

	MIMIC	eICU-CRD	PLAGH-ICUdb	AUMCdb	UCIHJ23db
Country	USA	USA	China	Netherlands	Spain
ICU admissions, No	38 508	132 994	63 515	20 127	4238
Male sex, No (%)	21 793 (56.59)	71 983 (54.15)	38 155 (60.07)	12 807 (65.16)	2730 (64.42)
Age, years, median (IQR)	65.70 (52.40–77.90)	65.00 (53.00–76.00)	56.89 (45.89–66.77)	Not available*	63.00 (50.00–73.00)
Admission type, No (%):					
Elective	6039 (15.83)	24 631 (18.52)	29 398 (46.29)	14 429 (71.69)	179 (4.22)
Emergency:	32 122 (84.17)	108 363 (81.48)	34 117 (53.71)	5698 (28.31)	4059 (95.78)
Via emergency department	19 152 (49.77)	69 077 (51.94)	29 790 (46.90)	2515 (12.50)	2220 (52.38)
Via other (eg, ward)	13 248 (34.40)	39 312 (29.56)	4327 (6.81)	3183 (15.81)	1839 (43.39)
Cardiac surgery, No (%)	7606 (19.75)	12 103 (9.10)	7265 (11.44)	7036 (34.96)	0 (0)
Mechanical ventilation,					
No (%)	19 156 (49.78)	49 159 (36.96)	Not available†	14 055 (69.83)	1844 (43.51)
Duration—hours, median (IQR)	19.00 (6.53–82.25)	27.82 (11.28–74.35)	Not available†	Not available	94.65 (28.40–256.82)
RRT, No (%)	1790 (4.65)	2154 (4.38)‡	Not available†	918 (4.56)	343 (8.09)
Crystalloid administration,					
No (%)	27 218 (70.74)	Not available§	52 165 (82.13)	19 079 (94.79)	1894 (44.69)
mL, median (IQR)	2000 (900–3650)	Not available§	1500 (750–2500)	1500 (1000–2469)	1000 (500–1500)
Colloid administration, No (%)	2587 (6.72)	Not available§	19 527 (30.74)	8642 (42.94)	217 (5.12)
Packed red blood cell transfusion, No (%)	12 361 (32.12)	Not available§	8584 (13.51)	5435 (27.00)	1007 (23.76)
ICU length of stay, hours, median (IQR)	50.30 (28.50–98.20)	39.00 (21.00–71.00)	95.01 (45.67–173.16)	25.00 (20.00–73.00)	93.70 (48.20–191.52)
ICU mortality, No (%)	2947 (7.66)	7618 (5.73)	3743 (5.89)	1929 (9.58)	619 (14.61)

*AUMCdb only contains age categories.
†Data not recorded in PLAGH-ICUdb.
‡Data from subset of eICU-CRD sites with reliable data regarding RRT therapy.
§Excluded due to poor data collection regarding IV fluid and PRBC administration.
ICU, intensive care unit; PRBC, packed red blood cell; RRT, renal replacement therapy.

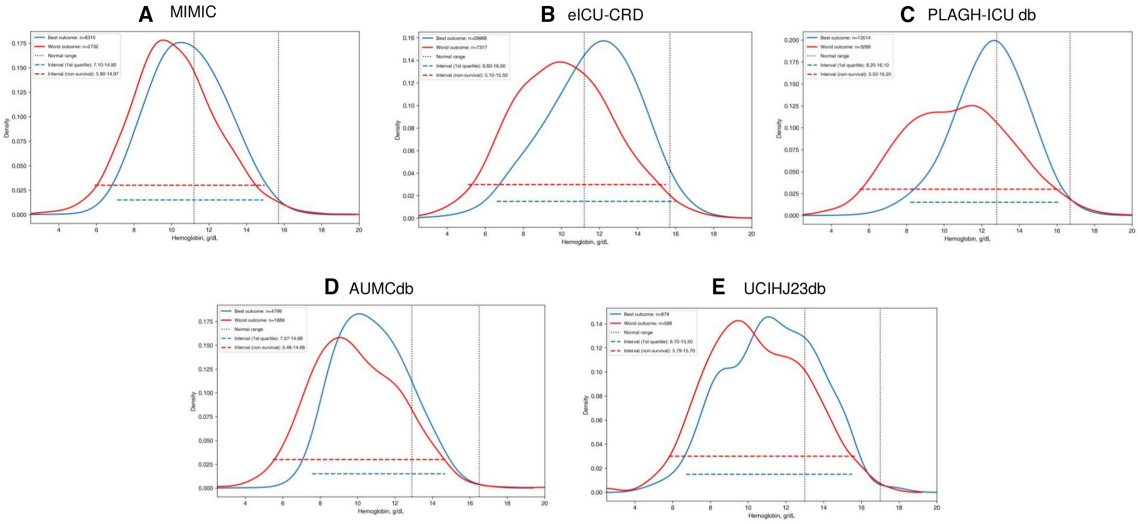

Figure 1 Minimum haemoglobin measurement on first intensive care unit admission—best versus worst outcome per database (A-E).

group, it is clear that normal reference ranges are not meaningful in ICU contexts. This discrepancy likely represents how reference ranges are formulated: reference ranges, though used to define normal and abnormal for both healthy individuals and critically unwell patients, are typically derived from samples of healthy outpatients.[1]

As expected, patients with the best and worst clinical outcomes had differing laboratory results across databases. However, between databases, we found that the extent to which these groups differed was variable, as the degree of overlap in distributions changed across investigations and the context in which they were utilised. For example, in the UCIHJ23db from Spain, the creatinine of patients with the best and worst clinical outcomes had substantial overlap (OVL=0.67), suggesting a decreased ability for creatinine to differentiate between patients with good and bad outcomes in this context. By comparison, creatinine results in the PLAGH-ICUdb from China demonstrated a lower overlap between groups (OVL=0.48). Consequently, creatinine may serve as a better prognosticator in this database, as it better discriminates between those with good and bad outcomes.

Variation in the overlap of laboratory results between patients with the best and worst outcomes was also seen

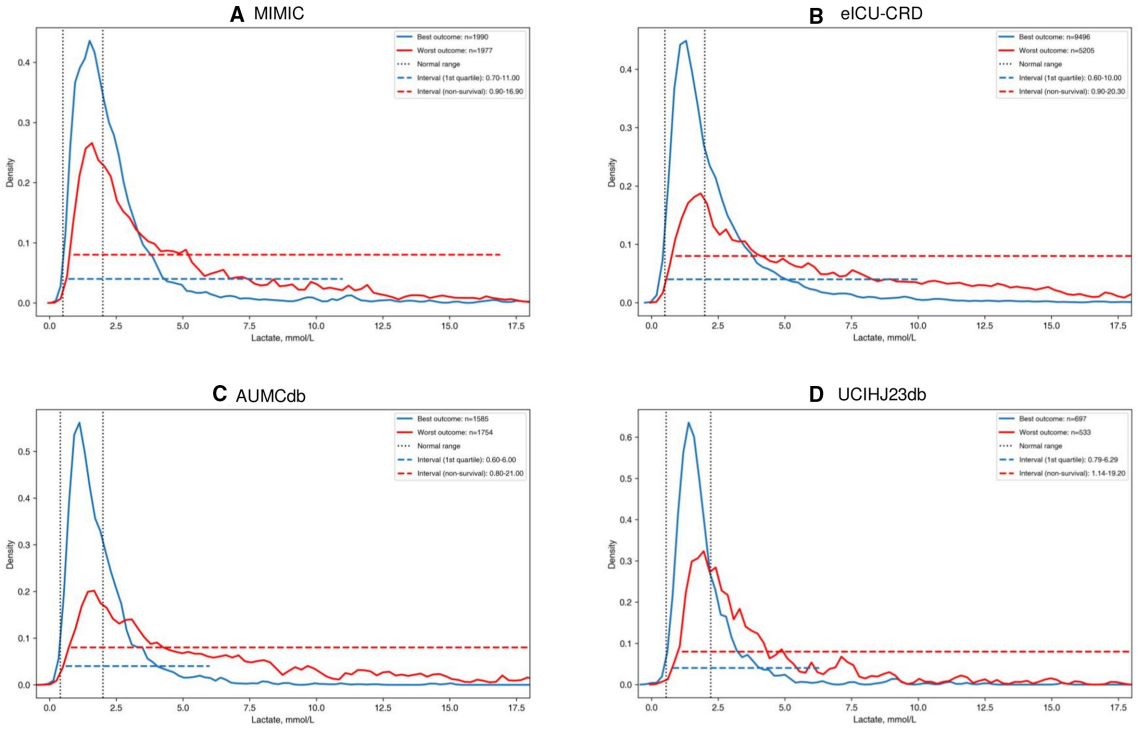

Figure 2 Maximum lactate measurement on first intensive care unit admission—best versus worst outcome per database (A-D). Data not recorded in PLAGH-ICUdb.

Table 2 Ninety-five per cent CIs of laboratory results stratified by best and worst outcome patient groups

	Creatinine, mg/dL		Haemoglobin, g/dL		Lactate, mmol/L		Sodium, mEq/L	
	Best	Worst	Best	Worst	Best	Worst	Best	Worst
MIMIC	0.50–4.50	0.50–6.40	7.10–14.90	5.90–14.97	0.70–11.00	0.90–16.90	131.00–146.00	129.00–155.78
eICU-CRD	0.50–6.45	0.54–7.50	6.60–16.00	5.10–15.50	0.60–10.00	0.90–20.30	130.00–147.00	128.00–157.00
PLAGH-ICUdb	0.44–1.51	0.42–7.49	8.20–16.10	5.50–16.20	Not available*	Not available*	133.10–147.90	128.80–156.60
AUMCdb	0.45–5.12	0.54–5.72	7.57–14.68	5.48–14.68	0.60–6.00	0.80–21.15	132.00–147.00	132.00–158.00
UCIHJ23db	0.40–5.04	0.44–6.33	6.70–15.50	5.79–15.70	0.79–6.29	1.14–19.20	132.00–148.93	129.00–155.28

*Data not recorded in PLAGH-ICUdb.

to different extents in the measurement of haemoglobin, lactate and sodium. These results imply that different investigations may represent good prognosticators in one context but not another and question the value of attempting to return these results to a healthy patient's reference range. As with currently utilised reference ranges, context-specific reference ranges developed from a heterogeneous cohort of patients would need to be interpreted with an understanding of individual patient factors and how their acute pathology may alter the significance of specific results.

While outside the scope of our project, we included data regarding patient demographics and common critical care interventions (renal replacement therapy (RRT), PRBC transfusion and crystalloid vs colloid resuscitation), which may have indicated mechanisms contributing to these variations in overlap. For instance, PLAGH-ICUdb was seen to have a narrower distribution of creatinine results. While RRT data was not available from this database, it can be seen that intravenous fluid administration was similar to that in other databases, so is unlikely to explain variations in renal function. However, the overall younger age of patients in PLAGH-ICUdb may have contributed to their lower creatinine. Furthermore, PLAGH-ICUdb also displayed the lowest overlap in haemoglobin results between best and worst outcome patients. While they also had the lowest PRBC administration rate, whether this represents a causative relationship is unknown. More broadly, we have demonstrated that substantial variability in the case mix and therapies provided existed across the databases, which may have contributed to differences in results across countries. Considering the differences between centres worldwide, the concept of context-specific reference ranges may prove even more useful by guiding practice with the goal of improving patient outcomes rather than unnecessarily normalising pathology results. Future research must consider the impact of case mix and clinical practices when developing new reference ranges, which would then require prospective validation to confirm them as appropriate treatment targets.

As such, our current study forms the foundation for several avenues of future enquiry. First, we intend to analyse and compare more homogeneous subgroups of patients (eg, cardiac surgery) and define context-specific laboratory result ranges, which are associated with the best clinical outcomes, and may therefore represent 'normality' for these groups of critically ill patients. Such reference ranges may then be prospectively validated to determine if they represent appropriate treatment targets and whether deviation from these ranges are associated with poorer outcomes. Furthermore, prospective studies will allow for the collection of data regarding the therapies provided to patients and thereby an investigation of the mechanism through which context-specific variations may arise. In addition, among databases that have collected data over a greater length of time (eg, PLAGH-ICUdb from 2008 until 2019), we intend to investigate

Table 3 Overlap between laboratory distributions of best and worst outcome patients with the local reference range

	Creatinine		Haemoglobin		Lactate		Sodium	
	OVL	SMD	OVL	SMD	OVL	SMD	OVL	SMD
MIMIC	0.65	−0.53	0.85	0.35	0.65	−0.65	0.76	−0.31
eICU-CRD	0.62	−0.48	0.74	0.61	0.57	−1.03	0.77	−0.35
PLAGH-ICUdb	0.48	−0.92	0.67	0.8	Not available*	Not available*	0.75	−0.09
AUMCdb	0.61	−0.48	0.81	0.42	0.47	−1.01	0.67	−0.74
UCIHJ23db	0.67	−0.46	0.86	0.32	0.57	−0.73	0.75	−0.31
Mean	0.61	−0.57	0.79	0.50	0.45	−0.68	0.74	−0.36

*Data not recorded in PLAGH-ICUdb.
OVL, overlapping coefficient; SMD, standardised mean difference.

whether the association between laboratory results and outcomes varies over time and therefore suggests that the prognostic value of results and their corresponding reference ranges require periodic review.

Strengths and limitations

Our study has several strengths. It is an analysis of an extremely large dataset, including more than 250 000 patients from three continents. Moreover, the ICUs included are varied in their case mixes and the corresponding severity of illness of their patients.

However, several limitations exist within this study. As with all retrospective research involving multiple large databases, variation in the design, collection and coding of variables may vary across datasets, creating inaccuracy in results. In our study, this is mitigated through the use of objective variables including laboratory results, ICU LOS and ICU mortality. Retrospective research of this nature is also inherently limited by missing data. Notably in our study this included missing data regarding lactate and interventions from PLAGH-ICUdb and intravenous fluid therapy and transfusions in eICU-CRD, respectively. However, other than lactate, these variables were used purely for hypothesis-generating purposes and do not alter our primary findings. The included databases collected information from ICU admissions across varying years, so differences in results may reflect changes in global practices over time rather than differences between centres or countries. Dichotomising patients into those with best and worse outcomes using ICU LOS and mortality does not reflect patient outcomes beyond ICU discharge. This includes the possibility that patients classified as having the 'best' outcome may have been discharged quickly from the ICU to receive end-of-life care. However, these definitions improved interpretability of our results and are consistent with those used previously.[14] Further, our study includes descriptive analyses without adjustment for potential confounders. Therefore, the associations between individual laboratory results and patient outcomes do not indicate independent causative relationships and should not be interpreted as such. Finally, comparing heterogeneous patient populations comprising varied case mixes is problematic. The possibility that context-specific reference ranges would also need to vary based on patient factors or specific conditions exists, though could not be concisely investigated in our present work.

CONCLUSION

In a cohort of more than 250 000 patients admitted to ICUs across four countries and three continents, there was substantial deviation in laboratory results when compared with normal reference ranges, even for those with the best clinical outcomes. Furthermore, when stratified by patients with the best and worst clinical outcomes, the degree of overlap between these patient groups varied widely across investigations and databases. These results suggest not only that specific reference ranges may be required for critically ill patients in different contexts but also that investigations may have a varying ability to discriminate between patients' outcomes depending on the setting.

Author affiliations

[1]School of Medicine, Chinese PLA General Hospital, Beijing, China
[2]Laboratory for Computational Physiology, Harvard-MIT Division of Health Sciences and Technology, Cambridge, Massachusetts, USA
[3]Institute of Medical Informatics, Charité - Universitätsmedizin Berlin (corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health), Berlin, Germany
[4]School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia
[5]Department of Internal Medicine, Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA
[6]Global Health Drug Discovery Institute, Beijing, China
[7]Department of Computer Science and Technology, Tsinghua University, Beijing, China
[8]Department of Intensive Care Medicine, Joan XXIII University Hospital in Tarragona, Tarragona, Catalunya, Spain
[9]Pere Virgili Health Research Institute, Reus, Catalunya, Spain
[10]School of Biological Science and Medical Engineering, Beihang University, Beijing, China
[11]Information Department, Chinese PLA General Hospital, Beijing, China
[12]Department of Cardiology, National University Health System, Singapore
[13]Department of Intensive Care Medicine, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
[14]Medical Innovation Research Department, Chinese PLA General Hospital, Beijing, China
[15]Division of Pulmonary Critical Care and Sleep Medicine, Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA

Twitter Leo Anthony Celi @MITCriticalData

Acknowledgements Article funding was supplied by MIT Libraries, the Beijing Municipal Science and Technology Project (Z181100001918023) and the Big Data R&D Project of Chinese PLA general hospital (2018MBD-009).

Funding LAC is funded by the National Institute of Health through NIBIB R01 EB017205.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval The study data and protocol were approved, where applicable, by the respective institutional review board (IRB) as follows: The data in MIMIC-III has been previously de-identified, and the IRBs of the Massachusetts Institute of Technology (0403000206) and Beth Israel Deaconess Medical Center (2001-P-001699/14) approved the use of the database for research. The use of eICU-CRD is exempt from IRB approval due to the retrospective design, lack of direct patient intervention, and the security schema, for which the re-identification risk was certified as meeting safe harborharbour standards by an independent privacy expert (Privacert, Cambridge, MAMassachusetts, USA). The use of AmsterdamUMCdb is likewise exempt from IRB approval due to a combination of de-identification, contractual and governance strategies where re-identification is not reasonably likely and can therefore be considered anonymous information in the context of the General Data Protection Regulation. Data from UCIHJ23db has been previously anonymizsed, removing any link with real identifiers. The IRB of the Hospital Universitari de Tarragona Joan XXIII approved the anonymizsation mechanism used for the present study. Finally, the data from People's Liberation Army (PLA) General Hospital was de-identified and approved for research use by the hospital ethics committee (S2021-050-01). Due to the de-identified nature of the data, the analysis is not considered human subject research. Hence, informed consent was waived for this study as approved by Beth Israel Deaconess Medical Center, Boston, MAMassachusetts, USA, and all other partaking hospitals and universities.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. Data are available upon reasonable request.

ORCID iDs
Louis Agha-Mir-Salim http://orcid.org/0000-0002-2733-5084
Leo Anthony Celi http://orcid.org/0000-0001-6712-6626

REFERENCES

1 Horowitz G, Jones G. Establishment and use of reference intervals. In: Burtis C, Ashwood E, Bruns E, eds. *Tietz textbook of clinical chemistry and molecular diagnostics*, 2017: 170–94.
2 Takala J, Ruokonen E, Webster NR, *et al*. Increased mortality associated with growth hormone treatment in critically ill adults. *N Engl J Med* 1999;341:785–92.
3 NICE-SUGAR Study Investigators for the Australian and New Zealand Intensive Care Society Clinical Trials Group and the Canadian Critical Care Trials Group, Finfer S, Chittock D, *et al*. Intensive versus conventional glucose control in critically ill patients with traumatic brain injury: long-term follow-up of a subgroup of patients from the NICE-SUGAR study. *Intensive Care Med* 2015;41:1037–47.
4 Holst LB, Haase N, Wetterslev J, *et al*. Lower versus higher hemoglobin threshold for transfusion in septic shock. *N Engl J Med* 2014;371:1381–91.
5 Aberegg SK, O'Brien JM. The normalization heuristic: an untested hypothesis that may misguide medical decisions. *Med Hypotheses* 2009;72:745–8.
6 Carson JL, Triulzi DJ, Ness PM. Indications for and adverse effects of red-cell transfusion. *N Engl J Med* 2017;377:1261–72.
7 Carson JL, Stanworth SJ, Roubinian N, *et al*. Transfusion thresholds and other strategies for guiding allogeneic red blood cell transfusion. *Cochrane Database Syst Rev* 2016;10:CD002042.
8 Carson JL, Guyatt G, Heddle NM, *et al*. Clinical practice guidelines from the AABB: red blood cell transfusion thresholds and storage. *JAMA* 2016;316:2025–35.
9 Soril LJJ, Noseworthy TW, Stelfox HT, *et al*. A retrospective observational analysis of red blood cell transfusion practices in stable, non-bleeding adult patients admitted to nine medical-surgical intensive care units. *J Intensive Care* 2019;7:19.
10 Sadana D, Kummangal B, Moghekar A, *et al*. Adherence to blood product transfusion guidelines-An observational study of the current transfusion practice in a medical intensive care unit. *Transfus Med* 2021;31:227–35.
11 Aberegg S. The normalization fallacy: why much of "critical care" may be neither. Pulm CCM, 2017. Available: http://pulmccm.org/main/2017/critical-care-review/normalization-fallacy-much-critical-care-may-neither/ [Accessed 7 Jan 2021].
12 Manrai AK, Patel CJ, Ioannidis JPA. In the era of precision medicine and big data, who is normal? *JAMA* 2018;319:1981–2.
13 Gräsbeck R. The evolution of the reference value concept. *Clin Chem Lab Med* 2004;42:692–7.
14 Tyler PD, Du H, Feng M, *et al*. Assessment of intensive care unit laboratory values that differ from reference ranges and association with patient mortality and length of stay. *JAMA Netw Open* 2018;1:e184521.
15 Ezzie ME, Aberegg SK, O'Brien JM. Laboratory testing in the intensive care unit. *Crit Care Clin* 2007;23:435–65.
16 Johnson AEW, Pollard TJ, Shen L, *et al*. MIMIC-III, a freely accessible critical care database. *Sci Data* 2016;3:1–9.
17 Pollard TJ, Johnson AEW, Raffa JD, *et al*. The eICU Collaborative research database, a freely available multi-center database for critical care research. *Sci Data* 2018;5:1–13.
18 Thoral PJ, Peppink JM, Driessen RH. Sharing ICU patient data Responsibly under the SCCM/ESICM joint data science collaboration: the AmsterdamUMCdb example. *Crit Care Med* 2021;49.
19 Liu T, Liu X, Fan Y. Constructing a Comprehensive Clinical Database Integrating Patients' Data from Intensive Care Units and General Wards. *Proc - 2019 12th Int Congr Image Signal Process Biomed Eng Informatics, CISP-BMEI 2019*, 2019.
20 Xu H. Online appendix—ICU reference ranges code. GitHub, 2021. Available: https://github.com/xhr0506/ICU-labtest
21 Inman HF, Bradley EL. The overlapping coefficient as a measure of agreement between probability distributions and point estimation of the overlap of two normal densities. *Commun Stat Theory Methods* 1989;18:3851–74.
22 Cohen J. *Statistical power analysis for the behavioral sciences. L. Erlbaum associates*, 1988.

Author/s:
Xu, H; Agha-Mir-Salim, L; O'Brien, Z; Huang, DC; Li, P; Gómez, J; Liu, X; Liu, T; Yeung, W; Thoral, P; Elbers, P; Zhang, Z; Saera, MB; Celi, LA

Title:
Varying association of laboratory values with reference ranges and outcomes in critically ill patients: an analysis of data from five databases in four countries across Asia, Europe and North America.

Date:
2021-10

Citation:
Xu, H., Agha-Mir-Salim, L., O'Brien, Z., Huang, D. C., Li, P., Gómez, J., Liu, X., Liu, T., Yeung, W., Thoral, P., Elbers, P., Zhang, Z., Saera, M. B. & Celi, L. A. (2021). Varying association of laboratory values with reference ranges and outcomes in critically ill patients: an analysis of data from five databases in four countries across Asia, Europe and North America.. BMJ Health Care Inform, 28 (1), pp.e100419-e100419. https://doi.org/10.1136/bmjhci-2021-100419.

Persistent Link:
http://hdl.handle.net/11343/296433

File Description:
Published version