



Accuracy of Point-of-Care Blood Glucose Meters Compared With Central Laboratory Reference Methods in Critically Ill Adults

A Systematic Review

Cyrene Elijah Amorado, MSMLS¹; Leah Quinto, PhD²; Mark Kerwin Sayas MSMLS³; Ramon George Atento, PhD⁴

¹ University of Santo Tomas, Manila Philippines

²De La Salle Medical and Health Sciences Institute, Dasmariñas City, Cavite, Philippines

³Lyceum of the Philippines University-Batangas, ⁴ First Asia Institute of Technology and Humanities, Batangas

Abstract

Background: Point-of-care (POC) blood glucose meters shorten turnaround time for bedside glycemic management, but critical illness introduces physiologic and analytical conditions that can degrade agreement with central laboratory measurements. Objective: To synthesize evidence on the analytical and clinical accuracy of POC blood glucose meters versus central laboratory reference methods in critically ill adults and identify conditions associated with disagreement. Methods: Searches of PubMed, Google Scholar, and SciSpace identified 1,003 records published from 2010 through September 13, 2026. After removal of 280 duplicates, 723 unique records remained. An initial screening limit left 223 deduplicated records without title/abstract assessment; a subsequent completeness audit applied the same eligibility criteria to these records and re-evaluated full-text decisions. Thirty studies were included. Risk of bias was assessed with an adapted QUADAS-2 framework. Results: POC performance was strongly device-, specimen-, and context-dependent. Several modern systems showed small average bias and high standards compliance in selected ICU cohorts, whereas capillary measurements deteriorated substantially in shock, hypotension, catecholamine/vasopressor exposure, anemia or hematocrit extremes, acidemia, hypoglycemia, and selected drug/interference states. Reported ISO compliance ranged from very high in stable or validated device-specific conditions to markedly poor in shock subgroups; these percentages were not directly pooled because studies used different devices, specimens, standards, thresholds, and outcome metrics. Conclusion: POC glucose meters cannot be treated as analytically interchangeable with central laboratory testing across all critically ill adults. Risk-stratified confirmation with laboratory testing is warranted when POC results are clinically discordant or obtained under physiologic conditions known to impair meter performance.

Keywords: *point-of-care testing; blood glucose; glucometer; intensive care; analytical accuracy; bias; ISO 15197; QUADAS-2; systematic review*

1. Introduction

Rapid glucose measurement is operationally attractive in intensive care because insulin titration and hypoglycemia management are time-sensitive. Central laboratory methods provide higher-order analytical control but impose transport, processing, and reporting delays. POC meters move testing to the bedside, reducing turnaround time but exposing measurement to specimen-site effects, variable hematocrit, tissue hypoperfusion, oxygen tension, acid-base disturbances, drugs, and operator or preanalytical factors that are more common in critically ill populations than in ambulatory self-monitoring.

The clinical concern is not simply whether a meter correlates with a laboratory analyzer. High correlation can coexist with clinically important systematic bias or wide limits of agreement. In ICU insulin protocols, errors around hypoglycemic thresholds or rapid treatment cut-points may change insulin dosing, delay dextrose administration, or conceal true hypoglycemia. Accordingly, method-comparison evidence should be interpreted using bias, limits of agreement, error-grid

performance, compliance with relevant analytical standards, and clinically consequential misclassification rather than correlation alone.

This review therefore addresses four questions: (1) the level of analytical agreement and accuracy between POC meters and central laboratory reference methods; (2) the magnitude and direction of reported bias; (3) patient-, specimen-, and clinical factors associated with reduced accuracy; and (4) the extent to which devices meet accepted analytical or clinical accuracy criteria.

2. Methods

2.1 Reporting framework and review design

The review was conducted and reported using PRISMA 2020 principles. Because the target evidence consists primarily of analytical method-comparison studies rather than diagnostic classification studies, PRISMA-DTA concepts were used where applicable, while risk-of-bias judgments were adapted from QUADAS-2 to the method-comparison context.

2.2 Eligibility criteria

Eligible studies were original peer-reviewed empirical reports published from 2010 through 2026 involving critically ill adults aged 18 years or older in ICU, critical-care, or equivalent acute-care settings. Studies had to compare POC blood glucose meter measurements directly with a central laboratory or validated core-laboratory reference method and report quantitative accuracy or agreement outcomes. Studies evaluating hematocrit, shock, vasopressor use, specimen type, glucose range, perfusion, or other modifiers were eligible when the required meter-to-reference comparison was present.

Studies were excluded when they were exclusively pediatric/neonatal without separately extractable adult critical-care data; were conducted only in outpatient, home, routine diabetes, emergency-only, non-critical inpatient, or intraoperative settings; lacked an eligible central/core laboratory comparator; evaluated only continuous/flash monitoring; lacked extractable quantitative agreement outcomes; or were reviews, editorials, commentaries, letters, protocols, case reports, dissertations, conference abstracts without sufficient data, or other non-peer-reviewed reports. The saved SciSpace screening criteria did not operationalize an English-only restriction; therefore language alone was not used as a post hoc exclusion criterion.

2.3 Search strategy

Searches were conducted in PubMed, Google Scholar, and SciSpace using AI/deep-search and full-text search functions. The search covered January 2010 through September 13, 2026 and combined terms for point-of-care or bedside glucose meters, blood glucose, accuracy, agreement, bias, validation, critically ill or ICU populations, and laboratory or reference methods. PubMed used title/abstract Boolean queries; Google Scholar and SciSpace used parallel keyword and semantic searches. Databases that were not queried, including Scopus, Web of Science, and CINAHL, are not presented as search sources.

The search log reported 532 records from SciSpace deep/AI searches, 200 from SciSpace full-text searches, 200 from Google Scholar, and 71 from PubMed, for 1,003 identified records.

2.4 Screening and audit reconciliation

The search log reported removal of 280 duplicates, leaving 723 unique records. The initial screening workflow was limited to 500 records, leaving 223 unique records without title/abstract assessment. To preserve completeness, these records were carried forward to a subsequent audit using the same prespecified eligibility criteria.

Records omitted from the initial 500-record screening set were reconciled against the raw retrieval exports and screened using the same saved inclusion and exclusion criteria. Fourteen additional plausible reports underwent detailed eligibility verification; 11 met the criteria and three were excluded. In parallel, the 57 reports initially advanced to full-text screening were re-audited. Of 55 reports initially classified as included, 18 met the saved criteria; Ellis et al. (2013), one of two reports initially excluded, met the criteria and was reinstated, while the dissertation remained excluded. The corrected final set comprised 30 studies.

2.5 Data extraction and synthesis

Extracted variables included study year and setting, patient population, specimen source, POC device, laboratory reference method, number of paired measurements where available, bias and agreement measures, ISO or guideline compliance, error-grid classification, and prespecified clinical/analytical modifiers. Because studies reported non-commensurate outcomes (mean difference in mg/dL or mmol/L, percentage bias, MARD, error-grid zones, ISO 15197 versions, CLSI/FDA boundaries, correlation, and heterogeneous limits of agreement), a pooled meta-analysis was not performed. Pooling a single mean-bias estimate across different devices, specimen types, glucose ranges, reference methods, and standards would violate exchangeability assumptions and create a clinically misleading summary.

2.6 Risk-of-bias appraisal

An adapted QUADAS-2 appraisal evaluated four domains: patient selection, index-test conduct, reference-standard adequacy, and flow/timing. Ratings were conservative when full methodological details were unavailable in the extracted records or published report. Applicability was considered acceptable when the study directly represented adult critical-care meter testing, but highly selected interference studies were interpreted as context-specific rather than population-wide estimates.

3. Results

3.1 Study selection

The reconciled PRISMA flow retained all 723 unique retrieved records for eligibility screening. After title/abstract review, 71 reports were assessed for eligibility and 30 studies were included in the qualitative synthesis. Forty-one reports were excluded at eligibility assessment: 10 were outside the 2010-2026 period, 15 had an ineligible population or setting or no separable adult critical-care data, seven were ineligible publication types, four lacked an eligible index/reference/outcome comparison, and five were duplicate, overlapping, or unverifiable records.

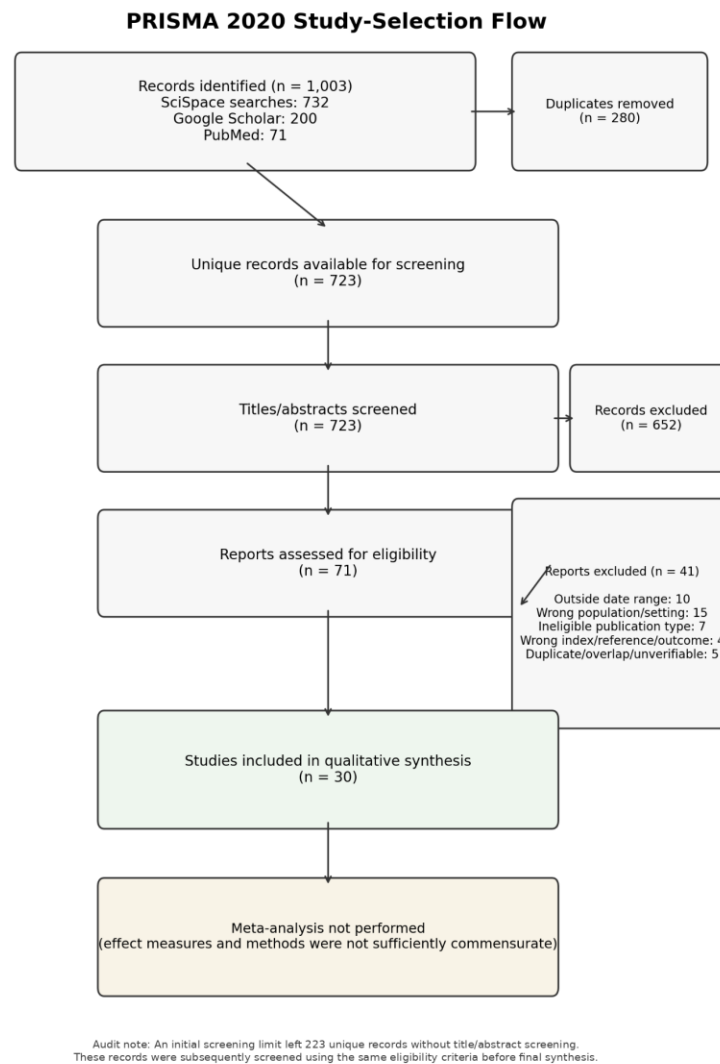


Figure 1. PRISMA 2020 study-selection flow after reconciliation of the initial screening limit.

3.2 Characteristics of included studies

The 30 included studies spanned a broad range of critical-care designs: prospective and cross-sectional method comparisons, retrospective paired-result analyses, postmarket electronic-record surveillance, multicenter analytical evaluations, and focused interference studies. Sample sizes ranged from small mechanistic or pilot cohorts to large real-world datasets. Devices included Accu-Chek Inform/Inform II, Nova StatStrip, CONTOUR systems, Precision/Abbott platforms, HemoCue, Menarini, and other hospital meters. Reference methods included central-laboratory hexokinase, glucose oxidase, dry chemistry, Vitros, Cobas, and YSI systems.

Study	Focus	Design / setting	Principal accuracy finding
Pidcoke et al., 2010	Anemia Causes Hypoglycemia in Intensive Care Unit Patients Due to Error in Single-Channel Glucometers:...	Study Type: Both retrospective and prospective phases. Setting: Burn and trauma ICUs at a single research institution. Sample Size: Healthy volunteers for in-vitro studies....	ICU anemia produced systematic positive bias in a single-channel meter; hematocrit correction substantially reduced error.
Fekih Hassen et al., 2010	Bedside capillary blood glucose measurements in	Study type: Not reported in available record. Setting:	43 adult ICU patients; catecholamine therapy increased discordant capillary-

Study	Focus	Design / setting	Principal accuracy finding
	critically ill patients: influence of catecholamine therapy	Critically ill patients (ICU patients). Sample size: Not reported in available record. Number of paired measurements: Not available in...	laboratory results (40% versus 29% without catecholamines); earlobe sampling did not solve the problem.
Argollo et al., 2010	Valores glicêmicos oferecidos pelo glicosímetro portátil, utilizando sangue de diferentes vias de coleta:...	Study Type: Quantitative validity study Setting: A private hospital's ICU in the city of Salvador, Bahia, Brazil, with 16 beds for adult clinical and surgical patients. Sample...	34 adult ICU patients/136 samples; arterial and central-venous POC results correlated better with laboratory values than capillary measurements, but agreement remained imperfect.
Hopf et al., 2011	Comparison of point-of-care testing glucose results from intensive care patients measured with network-ready...	Study type: Not reported in available record. Setting: Intensive care units. Sample size: Not reported in available record. Number of paired measurements: Not reported in available record. Blood...	Four network-ready meters tested against YSI reference using ICU blood; performance and susceptibility to oxygen/hematocrit effects were device-specific.
Creed et al., 2011	Point-of-Care Glucose Monitoring in a Critical Care Setting: Evaluation of a New Maltose Independent...	Critical-care evaluation of Accu-Chek Inform II, Nova StatStrip, and Abbott Medisense PXP against a reference method, including interference challenges.	Three POC systems were evaluated in critical care against reference methodology; accuracy and resistance to maltose, sodium, and hematocrit effects varied by device.
DuBose et al., 2012	Discrepancies between capillary glucose measurements and traditional laboratory assessments in both shock and...	Not reported in available record.	1,935 paired trauma-ICU samples; >15% disagreement occurred in 37.1% during shock and 40.8% during non-shock; agreement was poorest near hypoglycemia.
Feng et al., 2012	Evaluation of accuracy and influence factors of bedside blood glucose monitoring in critically ill patients	Study of 240 critically ill adults; synchronous venous, arterial and capillary specimens measured by glucose oxidase- and dehydrogenase-based POC methods versus laboratory...	240 critically ill adults; POC inaccuracy was greatest in hypoglycemia and varied by specimen and enzyme chemistry; hematocrit and illness severity contributed.
Gijzen et al., 2012	Is there a suitable point-of-care glucose meter for tight glycemic control? Evaluation of one home-use and...	Study type: Not reported in available record. Setting: Intensive care unit (ICU) Sample size: 80 patients Number of paired measurements: 390 glucose measurements Blood specimen types:...	80 ICU patients/390 measurements; ISO compliance ranged from 63.3% to 99.5% across five meters, demonstrating substantial device heterogeneity.
Lonjaret et al., 2012	Relative accuracy of arterial and capillary glucose meter measurements in critically ill patients.	Not reported in available record.	75 ICU patients/302 samples; ISO compliance 74.8% for capillary and 88.7% for arterial meter values, both tending to underestimate laboratory glucose.
López et al., 2012	Validation of a glucose meter at an intensive care unit	Study type: Not reported in available record. Setting: Intensive care units (ICUs) Sample size: Eighty-nine different samples Number of paired measurements: Not	89 ICU samples; StatStrip versus Cobas laboratory glucose showed $r=0.99$ and $ICC=0.99$, supporting good agreement in the tested setting.

Study	Focus	Design / setting	Principal accuracy finding
		reported in available record....	
Mitsios et al., 2013	Analytic Evaluation of a New Glucose Meter System in 15 Different Critical Care Settings	Study Type: Analytic evaluation to determine accuracy and precision. Setting: 15 different critical care settings across two academic medical centers (Washington University and...	600 patients/1,200 comparisons across 15 critical-care settings; combined bias 0.71 mg/dL and 98.8% within a stringent comparison boundary.
Corl et al., 2013	Evaluation of Point-of-Care Blood Glucose Measurements in Patients with Diabetic Ketoacidosis or...	Study Type: Prospective, cross-sectional study. Setting: Harborview Medical Center medical cardiac intensive care unit (MCICU). Sample Size: 28 patients. Number of Paired...	28 DKA/HHS critical-care patients; venous and capillary POC readings were higher than laboratory glucose, with capillary divergence ~30 mg/dL; acidemia amplified error.
Ellis et al., 2013	Suitability of capillary blood glucose analysis in patients receiving vasopressors	Prospective convenience sample of 50 adult postoperative cardiothoracic ICU patients receiving insulin and vasopressors; 162 simultaneous capillary, arterial POC, and arterial...	50 postoperative cardiothoracic ICU patients on vasopressors; capillary POC was less accurate than arterial POC against clinical chemistry reference.
Garingarao et al., 2014	Accuracy of Point-of-Care Blood Glucose Measurements in Critically Ill Patients in Shock	Study Type: Not reported in available record. Setting: ICU patients. Sample Size: Not reported in available record. Number of Paired Measurements: 186 paired glucose measurements. Blood...	186 paired samples; ISO 2003 compliance 95.7% in normotensive versus 79.8% in hypotensive patients; mean bias -12.4 versus -34.9 mg/dL.
Karon et al., 2014	Retrospective evaluation of the accuracy of Roche AccuChek Inform and Nova StatStrip glucose meters when used...	Study Type: Retrospective evaluation of two glucose systems, and a prospective evaluation of the Nova StatStrip. Setting: Intensive care unit (ICU) patients. Sample Size: Not...	Retrospective/prospective ICU evaluation; Nova StatStrip median bias ~1 mg/dL versus ~11 mg/dL for Roche AccuChek Inform in paired data.
Pereira et al., 2015	Inaccuracy of Venous Point-of-Care Glucose Measurements in Critically Ill Patients: A Cross-Sectional Study.	Study Type: Cross-sectional study with prospective data collection. Setting: 41-bed adult mixed-ICU in a tertiary care hospital. Sample Size: 145 patients enrolled. Number of...	145 adult mixed-ICU patients; mean bias varied by device/specimen (about 10-19 mg/dL) with wide limits of agreement; low hematocrit increased error.
Louie et al., 2015	Performance Evaluation of a Glucose Monitoring System for Point-of-Care Testing With the Critically Ill...	Multicenter evaluation using remnant deidentified critically ill whole-blood samples from three institutions.	Multicenter critically ill evaluation of Accu-Chek Inform II against hexokinase reference, including pO ₂ , hematocrit and sodium stress testing.
Bhurayanontachai, 2016	Agreement of blood glucose levels between venous blood test and arterial blood glucometer in the ICU	Prospective ICU method-comparison study; 15 hemodynamically stable critically ill adults; 45 paired arterial glucometer and venous central-laboratory samples.	15 stable critically ill adults/45 pairs; arterial glucometer and venous central laboratory showed $r=0.973$, 100% Clarke zone A, but wide Bland-Altman limits supported caution.

Study	Focus	Design / setting	Principal accuracy finding
Claerhout et al., 2016	Performance of strip-based glucose meters and cassette-based blood gas analyzer for monitoring glucose levels...	Comparative split-sample protocol Surgical intensive care unit (ICU) setting Blood gas samples used Samples below (n=9), in (n=51) or above (n=80) 4.2-5.9 mM range for...	Surgical ICU comparison: MARD 3.9% for BGA and 4.4-6.5% for tested strip meters; device-specific deviations from laboratory reference were evident.
Schroeder et al., 2016	Postmarket Surveillance of Point-of-Care Glucose Meters through Analysis of Electronic Medical Records	Study Type: Laboratory validation study, prospective bedside intensive care unit (ICU) study, and EMR data mining method (DETECT). Setting: Intensive care unit (ICU) for the...	Prospective ICU and EMR postmarket surveillance; average bias ~4.5-5.0 mg/dL before glycolysis correction; only a minority exceeded CLSI quality goals.
Schmolze et al., 2016	Using point of care glucose meters in the critically ill: Assessing meter performance in the clinical context	Real-world retrospective matching of six months of hospital POC results with central laboratory glucose within 10 minutes; 854 final matched pairs.	854 closely timed POC-laboratory pairs; clinically reasonable error-grid performance, but critical-care results showed substantial analytical scatter.
Mahmoodpoor et al., 2016	Validity of bedside blood glucose measurement in critically ill patients with intensive insulin therapy	Descriptive study of 300 critically ill patients; repeated arterial samples analyzed by bedside glucometer and central laboratory.	300 critically ill patients; overall correlation $r=0.81$, but unstable hemodynamics and multiple drugs reduced confidence in bedside readings.
DuBois et al., 2017	Bedside glucose monitoring—is it safe? A new, regulatory-compliant risk assessment evaluation protocol in...	International multicenter observational assessment in critically ill patients using multiple analytical and clinical risk tools.	International multicenter evaluation used laboratory comparison, error-grid and insulin-dosing risk methods; showed clinically acceptable performance for the tested hospital meter system.
Zhang et al., 2017	Performance of a Modern Glucose Meter in ICU and General Hospital Inpatients: 3 Years of Real-World Paired...	Not reported in available record.	Three years of paired ICU/general-inpatient data showed modern meter performance could be acceptable, but error patterns and specimen-timing effects remained clinically relevant.
Smith et al., 2018	Accuracy of point-of-care blood glucose level measurements in critically ill patients with sepsis receiving...	Prospective observational pilot in five critically ill septic adults receiving high-dose IV vitamin C; 16 paired POC and laboratory measurements.	Five septic ICU adults/16 pairs during high-dose IV vitamin C; 81.3% met ISO 15197:2013 and mean difference was 8.9 mg/dL.
Pilackas et al., 2020	Clinical reliability of point-of-care glucose testing in critically ill patients	Prospective cohort of 46 adult ICU patients with 85 paired capillary POC and venous laboratory samples.	46 adult ICU patients/85 pairs; mean difference 5.23 mg/dL and few treatment-changing discrepancies, suggesting acceptable performance in a relatively stable cohort.
Alshaer et al., 2022	Comparing the accuracy of point-of-care with laboratory (capillary, venous, and arterial) blood glucose...	Prospective case-control study of 268 critically ill patients with capillary, venous, arterial and central blood sampling.	268 critically ill patients; mean POC-laboratory differences varied by specimen, with the largest discrepancy for arterial comparisons; shock itself was not uniformly associated with larger error.
Rosengrave et al., 2022	Effect of intravenous vitamin C on arterial blood gas analyser and Accu-Chek point-of-care glucose	In vivo ICU comparison in 16 septic patients receiving IV vitamin C; 41 matched samples, plus an in-vitro	16 septic ICU patients/41 samples; Accu-Chek met ISO 15197:2013 in 83% versus 98% for ABG; high ascorbate with renal impairment could

Study	Focus	Design / setting	Principal accuracy finding
	monitoring...	ascorbate interference experiment.	cause clinically relevant interference.
Sheen et al., 2023	Accuracy of point-of-care blood glucometers in neonates and critically ill adults	Not reported in available record.	106 critically ill adults analyzed separately from neonates; 98.1% and 97.2% of adult arterial results met ISO 15197:2013 for two CONTOUR systems.
Mishra et al., 2024	Accuracy of point-of-care capillary blood sugar measurements in critically ill patients: an observational...	Observational study of 50 adult ICU patients; simultaneous capillary GlucoCare Sense and peripheral venous laboratory glucose (glucose oxidase-peroxidase).	50 adult ICU patients; POC readings exceeded laboratory values by 24.34±12.01 mg/dL overall and by 36.82±4.84 mg/dL in shock; shock markedly reduced ISO compliance.

Table 1. Characteristics and principal findings of the 30 included studies.

3.3 RQ1: Analytical agreement and overall accuracy

The evidence does not support a single global estimate of POC meter accuracy in critically ill adults. Some modern systems performed closely to central laboratory methods in controlled or validated settings. Mitsios et al. reported a combined bias of approximately 0.71 mg/dL across 1,200 comparisons and high compliance with a stringent laboratory comparison boundary. López et al. reported $r=0.99$ and $ICC=0.99$ for StatStrip versus Cobas laboratory glucose in an ICU sample. Sheen et al. found that 98.1% and 97.2% of adult arterial measurements from two CONTOUR systems met ISO 15197:2013 criteria.

Other studies demonstrated clinically meaningful disagreement. Lonjaret et al. found that both capillary and arterial meter measurements underestimated serum laboratory glucose and failed to reach the desired ISO target. Gijzen et al. showed that five meters tested in the same ICU cohort had ISO compliance ranging from 63.3% to 99.5%, demonstrating that device technology can be as important as the clinical setting. Real-world studies likewise showed that clinically acceptable error-grid performance can coexist with substantial analytical scatter.

3.4 RQ2: Magnitude and direction of bias

Bias was heterogeneous in both direction and magnitude. Corl et al. observed POC values above laboratory glucose in patients with diabetic ketoacidosis or hyperosmolar syndrome, with mean divergence of approximately +21 mg/dL for venous POC and +30 mg/dL for capillary POC. Pidcoke et al. demonstrated systematic positive bias under anemia in a single-channel glucose oxidase meter. Pereira et al. reported positive mean differences of roughly 10-19 mg/dL depending on specimen and device, with wide Bland-Altman limits of agreement.

Conversely, Lonjaret et al. observed systematic underestimation relative to laboratory serum glucose, while Garingarao et al. reported negative bias in both normotensive and hypotensive ICU groups (-12.4 and -34.9 mg/dL, respectively). These opposing directions show why a single pooled mean bias would be uninterpretable without stratification by meter chemistry, specimen source, clinical condition, and reference method.

3.5 RQ3: Factors associated with reduced accuracy

3.5.1 Shock, hypotension and peripheral perfusion

The most consistent clinical signal concerned circulatory instability. Garingarao et al. found substantially poorer ISO and clinical error-grid performance in hypotensive patients. Mishra et al. reported a mean POC-laboratory difference of 36.82 mg/dL in shock versus 14.61 mg/dL without shock, with only 17.24% of shock readings within the target limits used by the study. Ellis et al. demonstrated that capillary measurements in vasopressor-treated cardiothoracic ICU patients were less dependable than arterial POC measurements. These findings support concern that peripheral capillary sampling is particularly vulnerable when microcirculatory flow is compromised.

3.5.2 Vasopressors and catecholamines

Catecholamine therapy and vasopressor use were repeatedly associated with poorer capillary agreement or were important prespecified modifiers. Hassen et al. observed discordant paired values in 40% of catecholamine-treated versus 29% of untreated patients. Ellis et al. focused specifically on vasopressor-treated patients and found arterial sampling more reliable than capillary sampling. Not every study found shock status itself to be a statistically decisive modifier, indicating that device, specimen, and hemodynamic context interact rather than operating as a single universal effect.

3.5.3 Hematocrit and anemia

Hematocrit effects were device-dependent but clinically important. Pidcoke et al. showed that anemia below approximately 34% hematocrit caused falsely high readings in a single-channel meter and that a hematocrit-based correction markedly reduced error. Pereira et al. also found low hematocrit associated with larger measurement differences across multiple POC configurations. Modern meters with hematocrit correction may mitigate, but do not necessarily eliminate, this source of bias.

3.5.4 Acidemia, oxygenation and glucose range

Corl et al. found markedly larger meter-laboratory differences when plasma pH was below 7.3. Several technical evaluations examined oxygen tension and meter chemistry because glucose oxidase- and dehydrogenase-based systems differ in susceptibility to oxygen and other matrix effects. Feng and Cao reported the highest inaccuracy in hypoglycemic samples, and DuBose et al. likewise found agreement poorest near hypoglycemic readings. This is clinically important because errors at low glucose concentrations can directly delay recognition or treatment of hypoglycemia.

3.5.5 Exogenous interferences

Vitamin C illustrates a context-specific exogenous interference. Smith et al. observed only 81.3% ISO 15197:2013 compliance in a small sepsis cohort receiving high-dose intravenous vitamin C, although error-grid analysis suggested limited clinical consequence in that sample. Rosengrave et al. found that Accu-Chek values met ISO criteria in 83% of paired samples and identified severe renal impairment with elevated plasma ascorbate as a condition capable of producing clinically relevant overestimation. These findings favor laboratory confirmation when an interference-prone therapy or metabolite is present.

3.6 RQ4: Compliance with analytical and clinical standards

Standards-compliance percentages varied substantially across meters and contexts. Direct ranking must be interpreted cautiously because studies used different ISO versions, thresholds, specimen types, glucose distributions, and clinical subgroups. The table below therefore preserves the standard and context attached to each percentage rather than treating the percentages as directly exchangeable performance scores.

Study	Device/specimen	Population/context	Standard	Compliance	Interpretive note
Gijzen et al., 2012	Roche Accu-Chek Inform II	ICU cohort	ISO 15197 (version reported by study)	99.5%	Highest of five meters in that study
Sheen et al., 2023	CONTOUR PLUS	Critically ill adult arterial samples	ISO 15197:2013	98.1%	Adult results reported separately
Sheen et al., 2023	CONTOUR PLUS ONE	Critically ill adult arterial samples	ISO 15197:2013	97.2%	Adult results reported separately
Mishra et al., 2024	GlucoCare Sense	Non-shock ICU subgroup	ISO 15197:2013 as applied by authors	96.43%	Marked deterioration in shock
Gijzen et al., 2012	Abbott Precision Xceed Pro	ICU cohort	ISO 15197	96.6%	Study-specific comparison
Garingarao et al., 2014	GDH-based reflectometric meter	Normotensive ICU subgroup	ISO 15197:2003	95.7%	Borderline pass under study standard
Gijzen et al., 2012	HemoCue Glu201DM	ICU cohort	ISO 15197	95.1%	Study-specific comparison
Ellis et al., 2013	Arterial POC	Cardiothoracic ICU on vasopressors	ISO tolerance used by study	94.4%	Arterial sampling outperformed capillary

Study	Device/specimen	Population/context	Standard	Compliance	Interpretive note
Gijzen et al., 2012	Nova StatStrip	ICU cohort	ISO 15197	91.0%	Study-specific comparison
Lonjaret et al., 2012	CONTOUR TS, arterial	Critically ill ICU	ISO 15197	88.7%	Still below 95% target
Ellis et al., 2013	Capillary POC	Cardiothoracic ICU on vasopressors	ISO tolerance used by study	88.3%	Peripheral perfusion concern
Rosengrave et al., 2022	Accu-Chek Inform II	Septic ICU receiving IV vitamin C	ISO 15197:2013	83.0%	Renal impairment/high ascorbate increased risk
Smith et al., 2018	POC meter	Septic ICU receiving high-dose IV vitamin C	ISO 15197:2013	81.3%	Pilot study, n=5
Garingarao et al., 2014	GDH-based reflectometric meter	Hypotensive/shock ICU subgroup	ISO 15197:2003	79.8%	Substantial loss of accuracy in hypotension
Lonjaret et al., 2012	CONTOUR TS, capillary	Critically ill ICU	ISO 15197	74.8%	Capillary less accurate than arterial
Gijzen et al., 2012	Menarini GlucoCard Memory PC	ICU cohort	ISO 15197	63.3%	Lowest of five meters in that study
Mishra et al., 2024	GlucoCare Sense	Shock ICU subgroup	ISO 15197:2013 as applied by authors	17.24%	Extremely poor agreement in shock

Table 2. Reported standards-compliance results. Percentages are study-specific and should not be pooled or directly ranked across different standards and populations.

3.7 Adapted QUADAS-2 appraisal

Most studies had low concern for the basic index-test and reference-standard domains because a commercial POC meter was directly compared with a recognized laboratory or validated reference method. Patient-selection risk was more often unclear or high when studies used convenience samples, highly selected interference cohorts, remnant blood, or very small clinical subgroups. Flow/timing was generally acceptable when measurements were simultaneous or closely paired; longer matching windows created higher risk because in-vivo glucose can change and ex-vivo glycolysis can alter laboratory values.

Study	Patient selection	Index test	Reference standard	Flow / timing
Pidcock 2010	Unclear	Low	Low	Low
Fekih Hassen 2010	Low	Low	Low	Low
Argollo 2010	Unclear	Low	Low	Low
Hopf 2011	Unclear	Low	Low	Low
Creed 2011	Unclear	Low	Low	Low
DuBose 2012	Low	Low	Low	Unclear
Feng 2012	Low	Low	Low	Low
Gijzen 2012	Low	Low	Unclear	Low
Lonjaret 2012	Low	Low	Low	Low
López 2012	Unclear	Low	Low	Low
Mitsios 2013	Low	Low	Low	Low
Corl 2013	Unclear	Low	Unclear	Low
Ellis 2013	High	Low	Low	Low
Garingarao 2014	Low	Low	Low	Low
Karon 2014	Unclear	Low	Unclear	Low
Pereira 2015	Low	Low	Low	Low
Louie 2015	Unclear	Low	Low	Low
Bhurayanontachai 2016	High	Low	Low	Low
Claerhout 2016	Unclear	Low	Low	Low
Schroeder 2016	Unclear	Low	Low	Unclear

Study	Patient selection	Index test	Reference standard	Flow / timing
Schmolze 2016	Unclear	Low	Low	Low
Mahmoodpoor 2016	Unclear	Low	Low	Low
DuBois 2017	Low	Low	Low	Low
Zhang 2017	Low	Low	Low	High
Smith 2018	High	Low	Low	High
Pilackas 2020	Low	Low	Low	Low
Alshaer 2022	Low	Low	Low	Low
Rosengrave 2022	High	Low	Low	Low
Sheen 2023	Low	Low	Low	Low
Mishra 2024	Low	Low	Low	Low

Table 3. Adapted QUADAS-2 risk-of-bias summary. Green = low risk; yellow = unclear risk; red = high risk.

Across the 120 domain judgments (30 studies × four domains), 97 were rated low risk, 17 unclear, and 6 high. High-risk judgments were concentrated in narrowly selected interference cohorts or extended timing windows, not in the existence of the laboratory comparator itself.

4. Discussion

4.1 Principal interpretation

The central finding is conditional rather than binary: bedside glucose meters can perform very well in critical care, but accuracy is not invariant across devices, specimens, physiologic states, or treatments. The best-performing contemporary systems approached central-laboratory agreement in selected settings, while other meters or the same meter under adverse conditions produced bias large enough to cross insulin-dosing or hypoglycemia thresholds. Thus, "POC versus laboratory" is not a single comparison; it is a family of device-by-specimen-by-patient-context comparisons.

The strongest recurring risk pattern concerns capillary testing under compromised peripheral perfusion. Shock, hypotension, catecholamine/vasopressor therapy, and microcirculatory disturbance plausibly alter capillary glucose and sample quality. Arterial or venous whole-blood POC testing often performed better than capillary testing, although specimen-source advantages were not universal and can be offset by contamination, line-handling, or timing problems.

The review also confirms that correlation is insufficient as an accuracy claim. Several studies reported high correlation coefficients yet clinically wide limits of agreement or poor standards compliance. For clinical governance, institutions should emphasize bias, limits of agreement, error-grid risk, and the proportion meeting prespecified accuracy criteria rather than relying on r or R^2 alone.

4.2 Implications for laboratory and healthcare operations

POC testing is fundamentally a turnaround-time versus analytical-risk trade-off. The operational value of a bedside result is greatest when rapid action is necessary, but the clinical cost of a biased result is highest near treatment thresholds. A risk-stratified verification policy is therefore more defensible than either universal reliance on POC meters or universal laboratory confirmation. High-risk triggers suggested by the evidence include hemodynamic shock or marked hypotension, escalating vasopressors, severe anemia or extreme hematocrit, acidemia, unexpected hypoglycemia, clinically discordant results, and therapies known to interfere with specific meter chemistries.

Laboratory governance should also be device-specific. The wide within-study spread in ISO compliance across meters shows that performance cannot be generalized from one platform to another. Local validation should reproduce the institution's intended specimen types and clinically relevant patient matrices, including ICU samples rather than healthy-volunteer material alone. Postmarket surveillance using paired electronic POC-laboratory results can complement prospective validation and detect drift after implementation.

4.3 Why meta-analysis was not performed

A quantitative meta-analysis was considered but rejected after inspection of the extracted outcomes. The evidence mixed mean and percentage bias, MARD, correlations, intraclass correlations, ISO 15197:2003 and 2013 criteria, CLSI/FDA thresholds, Clarke/Parkes/surveillance error grids, and different specimen/reference pairings. Many reports did not provide the

variance information required for a defensible pooled mean difference. A Bland-Altman forest plot that ignored these differences would create a visually precise but methodologically invalid pooled estimate. Structured quantitative comparison by device and context was therefore preferred.

4.4 Research gaps

Future work should prioritize prospective multicenter studies that use simultaneous paired specimens, prespecified higher-order reference methods, transparent device lot/operator information, and standardized reporting of bias with variance and limits of agreement. Studies should oversample hypoglycemia and other clinically critical ranges rather than allowing normoglycemic values to dominate aggregate accuracy. There is also a need for adequately powered analyses of shock severity, vasopressor dose, hematocrit, pH, oxygen tension, renal failure, and known electrochemical interferents, ideally with device-specific interaction estimates.

Economic and operational research is comparatively sparse. Studies that link analytical disagreement to insulin-dosing changes, repeat testing, laboratory confirmation, nursing workload, turnaround time, adverse events, and total cost would provide a stronger basis for institutional POC governance than analytical metrics alone.

4.5 Limitations

This review has several limitations. First, the formal search used PubMed, Google Scholar, and SciSpace rather than every major bibliographic database. Second, the initial title/abstract screening workflow was limited to 500 records; although the 223 previously unscreened records were subsequently assessed using the same eligibility criteria, the duplicate-matching logic used by the search platform could not be independently reproduced from the export, so the deduplication count was retained from the search log. Third, study heterogeneity prevented pooled meta-analysis. Fourth, some older or proprietary reports provided limited methodological detail, leading to conservative "unclear" QUADAS-2 ratings. Fifth, ISO compliance values across studies are not directly comparable because standards, glucose ranges, devices, and patient contexts differ.

5. Conclusion

POC blood glucose meters offer important operational advantages in critical care, but their results are not universally interchangeable with central laboratory glucose measurements. Accuracy is most dependable when the device has been validated in critically ill populations and when sampling conditions are physiologically stable. Shock, hypotension, vasopressors/catecholamines, anemia or hematocrit extremes, acidemia, glucose extremes, and selected exogenous interferents can materially increase disagreement. The evidence supports device-specific validation, ongoing quality surveillance, preferential use of more reliable specimen routes where appropriate, and targeted laboratory confirmation when clinical context or result pattern indicates elevated analytical risk.

References

- Alshaer, A., Badgheish, B. A., Alsadah, Z., Sewify, K., Alghazal, S., Alzahrani, S. A., Qadi, A. A., Alqahtani, R., Farsani, G. A., & Shilash, A. (2022). Comparing the accuracy of point-of-care with laboratory (capillary, venous, and arterial) blood glucose levels in critically ill patients with and without shock. *BMC Research Notes*. <https://doi.org/10.1186/s13104-022-06256-0>
- Argollo, A. P. B., Faustino, T. N., Faustino, T. N., & Pedreira, L. C. (2010). Valores glicêmicos oferecidos pelo glicosímetro portátil, utilizando sangue de diferentes vias de coleta: Estudo de validade. *Revista Brasileira de Terapia Intensiva*. <https://doi.org/10.1590/s0103-507x2010000400007>
- Bhurayanontachai, R. (2016). Agreement of blood glucose levels between venous blood test and arterial blood glucometer in the ICU. *Netherlands Journal of Critical Care*.
- Claerhout, H., De Prins, M., Mesotten, D., Van den Berghe, G., Mathieu, C., Van Eldere, J., & Vanstapel, F. (2016). Performance of strip-based glucose meters and cassette-based blood gas analyzer for monitoring glucose levels in a surgical intensive care setting. *Clinical Chemistry and Laboratory Medicine*. <https://doi.org/10.1515/ccm-2014-1289>
- Corl, D. E., Yin, T. S., Mills, M. E., Spencer, T. L., Greenfield, L., Beauchemin, E., Cochran, J., Suhr, L. D., Thompson, R. E., & Wisse, B. E. (2013). Evaluation of point-of-care blood glucose measurements in patients with diabetic ketoacidosis

or hyperglycemic hyperosmolar syndrome admitted to a critical care unit. *Journal of Diabetes Science and Technology*. <https://doi.org/10.1177/193229681300700516>

- Creed, G. M., Fox, T. J., & Beale, R. (2011). Point-of-care glucose monitoring in a critical care setting: Evaluation of a new maltose independent chemistry Accu-Chek INFORM II, Nova StatStrip, and Abbott Medisense PXP. *Point of Care: The Journal of Near-Patient Testing & Technology*. <https://doi.org/10.1097/poc.0b013e318207813b>
- DuBois, J. A., Slingerland, R. J., Fokkert, M. J., Roman, A., Tran, N. K., Clarke, W., Sartori, D. A., Palmieri, T. L., Malic, A., Lyon, M. E., & Lyon, A. W. (2017). Bedside glucose monitoring—is it safe? A new, regulatory-compliant risk assessment evaluation protocol in critically ill patient care settings. *Critical Care Medicine*. <https://doi.org/10.1097/ccm.0000000000002252>
- DuBose, J. J., Inaba, K., Branco, B. C., Barmparas, G., Lam, L., Teixeira, P. G. R., Belzberg, H., & Demetriades, D. (2012). Discrepancies between capillary glucose measurements and traditional laboratory assessments in both shock and non-shock states after trauma. *Journal of Surgical Research*. <https://doi.org/10.1016/j.jss.2012.04.003>
- Ellis, M. F., Benjamin, K., Cornell, M., Decker, K., Farrell, D., McGugan, L., Porter, G. P., Shearin, H., Zhao, Y., & Granger, B. B. (2013). Suitability of capillary blood glucose analysis in patients receiving vasopressors. *American Journal of Critical Care*. <https://doi.org/10.4037/ajcc2013692>
- Fekih Hassen, M., Ayed, S., Gharbi, R., Ben Sik Ali, H., Marghli, S., & Elatrous, S. (2010). Bedside capillary blood glucose measurements in critically ill patients: Influence of catecholamine therapy. *Diabetes Research and Clinical Practice*. <https://doi.org/10.1016/j.diabres.2009.09.018>
- Feng, T., & Cao, X.-Y. (2012). Evaluation of accuracy and influence factors of bedside blood glucose monitoring in critically ill patients. *Chinese Critical Care Medicine*. <https://doi.org/10.3760/cma.j.issn.1003-0603.2012.08.010>
- Garingarao, C. J. P., Buenaluz-Sedurante, M., & Jimeno, C. (2014). Accuracy of point-of-care blood glucose measurements in critically ill patients in shock. *Journal of Diabetes Science and Technology*. <https://doi.org/10.1177/1932296814538608>
- Gijzen, K., Moolenaar, D. L. J., Weusten, J. J. A. M., Pluim, H. J., & Demir, A. Y. (2012). Is there a suitable point-of-care glucose meter for tight glycemic control? Evaluation of one home-use and four hospital-use meters in an intensive care unit. *Clinical Chemistry and Laboratory Medicine*. <https://doi.org/10.1515/cclm-2012-0104>
- Hopf, S., Graf, B. M., & Gruber, M. (2011). Comparison of point-of-care testing glucose results from intensive care patients measured with network-ready devices. *Diabetes Technology & Therapeutics*. <https://doi.org/10.1089/dia.2011.0051>
- International Organization for Standardization. (2013). *ISO 15197:2013: In vitro diagnostic test systems—Requirements for blood-glucose monitoring systems for self-testing in managing diabetes mellitus*.
- Karon, B. S., Blanshan, C. T., Deobald, G. R., & Wockenfus, A. M. (2014). Retrospective evaluation of the accuracy of Roche AccuChek Inform and Nova StatStrip glucose meters when used on critically ill patients. *Diabetes Technology & Therapeutics*. <https://doi.org/10.1089/dia.2014.0074>
- Lonjaret, L., Claverie, V., Bérard, E., Riu-Poulenc, B., Geeraerts, T., Genestal, M., & Fourcade, O. (2012). Relative accuracy of arterial and capillary glucose meter measurements in critically ill patients. *Diabetes & Metabolism*. <https://doi.org/10.1016/j.diabet.2011.12.003>
- López, M., Serrano, J. L., Robles Rodríguez, J. L., & Márquez Márquez, T. (2012). Validation of a glucose meter at an intensive care unit. *Endocrinología y Nutrición: Órgano de la Sociedad Española de Endocrinología y Nutrición*. <https://doi.org/10.1016/j.endoen.2011.08.003>
- Louie, R. F., Curtis, C. M., Toffaletti, J. G., Handel, E. A., Slingerland, R. J., Fokkert, M. J., Muller, W., Weinert, S. E., Lee, D. M., & Kotagiri, S. (2015). Performance evaluation of a glucose monitoring system for point-of-care testing with the

critically ill patient population—A multicenter study. *Point of Care: The Journal of Near-Patient Testing & Technology*. <https://doi.org/10.1097/poc.000000000000045>

Mahmoodpoor, A., Hamishehkar, H., Shadvar, K., Sanaie, S., Iranpour, A., & Fattahi, V. (2016). Validity of bedside blood glucose measurement in critically ill patients with intensive insulin therapy. *Indian Journal of Critical Care Medicine*. <https://doi.org/10.4103/0972-5229.194009>

Mishra, K., Mishra, S., & Katial, T. (2024). Accuracy of point-of-care capillary blood sugar measurements in critically ill patients: An observational study. *International Journal of Critical Illness and Injury Science*. https://doi.org/10.4103/ijciis.ijciis_76_23

Mitsios, J. V., Ashby, L., Haverstick, D. M., Bruns, D. E., & Scott, M. G. (2013). Analytic evaluation of a new glucose meter system in 15 different critical care settings. *Journal of Diabetes Science and Technology*. <https://doi.org/10.1177/193229681300700518>

Page, M. J., McKenzie, J. E., Bossuyt, P. M., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>

Pereira, A. J., Corrêa, T. D., Pereira de Almeida, F., Deliberato, R. O., dos Santos Lobato, M., Akamine, N., Silva, E., & Cavalcanti, A. B. (2015). Inaccuracy of venous point-of-care glucose measurements in critically ill patients: A cross-sectional study. *PLOS ONE*. <https://doi.org/10.1371/journal.pone.0129568>

Pidcoke, H. F., Wade, C. E., Mann, E. A., Salinas, J., Cohee, B. M., Holcomb, J. B., & Wolf, S. E. (2010). Anemia causes hypoglycemia in intensive care unit patients due to error in single-channel glucometers: Methods of reducing patient risk. *Critical Care Medicine*. <https://doi.org/10.1097/ccm.0b013e3181bc826f>

Pilackas, K., El-Oshar, S., & Carter, C. (2020). Clinical reliability of point-of-care glucose testing in critically ill patients. *Journal of Diabetes Science and Technology*. <https://doi.org/10.1177/1932296819858633>

Rosengrave, P. C., Wohlrab, C., Spencer, E., Williman, J., Shaw, G., & Carr, A. C. (2022). Effect of intravenous vitamin C on arterial blood gas analyser and Accu-Chek point-of-care glucose monitoring in critically ill patients. *Critical Care and Resuscitation: Journal of the Australasian Academy of Critical Care Medicine*. <https://doi.org/10.51893/2022.2.0a7>

Schmolze, D., Horowitz, G. L., & Tolan, N. V. (2016). Using point of care glucose meters in the critically ill: Assessing meter performance in the clinical context. *Point of Care: The Journal of Near-Patient Testing & Technology*. <https://doi.org/10.1097/poc.0000000000000107>

Schroeder, L. F., Giacherio, D., Gianchandani, R., Engoren, M., & Shah, N. H. (2016). Postmarket surveillance of point-of-care glucose meters through analysis of electronic medical records. *Clinical Chemistry*. <https://doi.org/10.1373/clinchem.2015.251827>

Sheen, Y. J., Wang, J.-M., Tsai, P.-F., Lee, W.-J., Hsu, Y.-C., Wang, C.-Y., & Sheu, W. H.-H. (2023). Accuracy of point-of-care blood glucometers in neonates and critically ill adults. *Clinical Therapeutics*. <https://doi.org/10.1016/j.clinthera.2023.05.005>

Smith, K., Brown, C. S., Manning, B. M., May, T., Riker, R. R., Lerwick, P., Hayes, T. L., & Fraser, G. L. (2018). Accuracy of point-of-care blood glucose level measurements in critically ill patients with sepsis receiving high-dose intravenous vitamin C. *Pharmacotherapy*. <https://doi.org/10.1002/phar.2182>

Whiting, P. F., Rutjes, A. W. S., Westwood, M. E., et al. (2011). QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine*, 155(8), 529–536. <https://doi.org/10.7326/0003-4819-155-8-201110180-00009>

Zhang, R., Isakow, W., Kollef, M. H., & Scott, M. G. (2017). Performance of a modern glucose meter in ICU and general hospital inpatients: 3 years of real-world paired meter and central laboratory results. *Critical Care Medicine*. <https://doi.org/10.1097/ccm.0000000000002572>

Appendix A. Full-text exclusion summary

Primary full-text exclusion reason	n
Outside prespecified publication period (before 2010)	10
Wrong population/setting or adult critical-care data not separately extractable	15
Ineligible publication type / non-original empirical report	7
Wrong index/reference method or no eligible quantitative comparison	4
Duplicate, overlapping, or unverifiable record	5

The complete 41-record exclusion log is retained as a companion CSV to preserve an auditable record without overloading the manuscript.