Public Services Delivery
Scotland

NHS SCOTLAND

Built environment Experimental studies Critical Appraisal Tool (BECAT)

Guidance document

ASR603-002
Version 1 - August 2026

NHS Scotland Assure

Contents

Cite this document: Public Services Delivery Scotland, NHS Scotland Assure, Healthcare Scientist Technical Team. Built environment Experimental studies Critical Appraisal Tool (BECAT): Guidance document. 2026

Research Team

Healthcare Scientist Technical Team

Public Services Delivery Scotland (PSD Scotland), NHS Scotland Assure, Research and Innovation Services

- Dr Joanne Kinloch
- Dr Laura Rodriguez-Labajos
- Matthew Northcote
- Donald Bunyan
- Dr Lorena Jimenez-Sanchez
- Dr Sreedevi P. Unnikrishnan

Acknowledgements

We would like to acknowledge the members of the Delphi panel for contributing their time, expertise and insight to this study. We also wish to thank colleagues who provided research, programme and administrative support. Their contributions were invaluable in shaping and delivering this work.

Delphi panel

- Dr Keith Baker, Research Fellow
    - Glasgow Caledonian University, School of Science and Engineering
- Craig Webster, Healthcare Associated Infection System for Controlling Risk in the Built Environment (HAI-SCRIBE) Manager
    - Public Services Delivery Scotland (PSD Scotland), NHS Scotland Assure, Research and Innovation Services
- Fiona Strachan, Supply Chain Manager
    - NHS Lothian, Estates and Facilities, Facilities Management Supply Chain
- Gordon Ramage, Professor of Infection Prevention and Control
    - Glasgow Caledonian University, Research Centre for Health
- Abbie Houston, Healthcare Scientist (Health Protection)
    - PSD Scotland, NHS Scotland Assure, Antimicrobial Resistance and Healthcare Associated Infection (ARHAI) Scotland
- Catherine Stokoe, Infection Control Manager
    - NHS Highland, Infection Prevention and Control
- Susan Grant, Principal Architect
    - PSD Scotland, NHS Scotland Assure, Property, Sustainability and Capital Planning
- Eliza Kowalska De Souza, Accommodation and Strategic Asset Management System (SAMS) Manager
    - NHS Lothian, Estates and Facilities, Land and Property
- George Loumakis, Senior Lecturer in Energy
    - Glasgow Caledonian University, School of Science and Engineering
- W Gordon Mackay, Professor of Microbiology
    - University of the West of Scotland, Faculty of Computing, Science and Engineering

- Prof. Michael Gormley, Professor of Public Health and Environmental Engineering
    - Heriot-Watt University, School of Energy, Geoscience, Infrastructure and Society (EGIS), Institute for Sustainable Built Environment (ISBE)
- Susannah Cameron, Advanced Nurse Specialist, Infection Prevention and Control
    - NHS Highland Argyll & Bute Health and Social Care Partnership

Project support

- Sarah Murdoch, Senior User Researcher
    - PSD Scotland, Programme Management Services
- Dr Heather Gray, Lead Health Services Researcher
    - Healthcare Improvement Scotland, Evidence and Digital Directorate, Research and Information Service
- Sarah Gibson, Portfolio Programme Manager
    - PSD Scotland, NHS Scotland Assure, Programme Management Office
- Dr Geraldine O'Brien, Research Manager
    - PSD Scotland, NHS Scotland Assure, Research and Innovation Services
- NHS Scotland Assure Learning Network
- National Service Scotland (NSS) Innovation and Research Governance Group

Approvals

- Research governance approval granted by the NSS Innovation and Research Governance Group on 2 September 2025.
- Data Protection Impact Assessment (DPIA) 9596 approved on 23 November 2025.

1. Introduction

1.1. The Built environment Experimental studies Critical Appraisal Tool (BECAT) has been developed by the NHS Scotland Assure Healthcare Scientist Technical Team to support the structured appraisal of methodological quality of non-human built environment experimental studies.

1.2. The tool addresses a gap in existing critical appraisal tools. Unlike existing tools, which focus on human subjects, this tool is suitable for assessing studies with non-human subjects, such as physical systems, materials, or environments. To use this tool, the studies should be defined as:

- non-human: outcomes relate to building systems or objects. Studies with human outcomes (for example, patient infection, satisfaction, or well-being outcomes) are out of scope
- experimental: involves a deliberate intervention, exposure, or change to a built environment feature or condition, material, system, or process. Non-experimental designs (for example, audits or surveys) are out of scope
- built environment focused: a physical system, structure, space, or component under study (for example, ventilation, layout or materials). Studies focused solely on human outcomes are out of scope

1.3. BECAT is applicable across three primary study types:

- field (in situ): studies in real-world or near real-world settings, involving full- or partial-scale systems under operational or field-like conditions
- laboratory testing (in vitro): controlled experiments on materials or components in laboratory settings, allowing precise manipulation of variables
- COmputerised Simulation/ MOdelling (COSMO): studies using simulated, computational models, digital software (for example, computational fluid dynamics (CFD), or finite element analysis) to examine systems under controlled, modelled conditions

Purpose and scope

1.4. BECAT is designed to:

- assess methodological quality and reporting transparency of studies
- identify potential sources of bias
- support consistent decision-making in evidence synthesis
- improve comparability of study methodological quality within literature reviews
- provide structured guidance for reviewers

How to use BECAT

1.5. Users should:

- apply the Minimum Quality Criteria to screen initial eligibility
- where studies meet the screening criteria, complete all checklist items using ‘Yes’, ‘No’, ‘Unclear’, or ‘Not Applicable (N/A)’
- use the accompanying guidance to support consistent interpretation of items

1.6. Guidance and examples are provided to illustrate how each item should be assessed against a study.

1.7. Appraiser Notes are included throughout the tool to reflect feedback captured during the Delphi process that remains unresolved and to highlight areas requiring caution when rating. These notes are intended to support appraisers where there is potential for unintentional bias, inconsistent or subjective judgement, broad study scope, topic-specific expertise required (for example, from a subject matter expert) or challenges arising from built environment complexity or variability. In this context, complexity refers to the difficulty of isolating individual factors within interconnected built environment systems, while variability refers to differences between real-world settings, conditions, materials, or processes that may be difficult to replicate exactly.

1.8. In some of these cases, the use of ‘N/A’ or the removal of an item from the checklist may be appropriate.

1.9. For details on study identification and tool development, see Appendix A and Appendix B

2. Built environment Experimental studies Critical Appraisal Tool (BECAT)

2.1. This section presents the critical appraisal tool, including domain-level descriptions, items, topic areas, associated guidance, and examples. As this is a working tool currently undergoing testing and refinement, some areas may require further development. Review teams using this tool should develop a shared understanding of how each item applies to their topic area to support consistent appraisal.

Minimum quality criteria - screening domain (2 items)

2.2. This domain is a preliminary check that assesses whether a study meets the minimum study quality criteria required for further critical appraisal.

2.3. If a study is rated 'No' or 'Unclear' on any of the two screening items, it should not progress to full appraisal.

2.4. 'Not Applicable (N/ A)' may be used where justified (for example, exploratory, novel, or iterative studies), to avoid unfairly penalising complex built environment research.

Aims or objectives (Item S.1)

Item S.1: Are there clear research questions, aims or objectives?

Why it matters:

2.5. A clearly defined aim ensures the study has a focused purpose and provides the basis for evaluating whether the design, methods, and data are appropriate.

What to look for:

- statement of specific and measurable aims
- in some cases, this may follow a structured research framework, such as Population[1], Intervention or exposure, Comparator, and Outcomes (PICO)
- formulation of hypotheses for quantitative studies, where applicable
- a clearly stated and justified purpose for exploratory studies

[1] For non-human studies, 'Population' will be the subject under study.

- fails if: the purpose is missing, overly vague, or provides no meaningful direction for the study

When assessing this item, consider:

- whether broad study scope is appropriately justified
- the complexity of real-world built environment systems
- potential for inconsistent ratings between appraisers
- requirement for topic-specific or methodological expertise

Examples:

2.6. **Good:** A study aims to measure airborne particle dispersion under mechanical versus natural ventilation in controlled full-scale test rooms.

2.7. **Poor:** A study states the objective is to test the effect of air quality in buildings. It is vague, and provides no further details, research questions or aims.

Study coherence (item S.2)

Item S.2: Do the collected data answer the research questions, aims or objectives?

Why it matters:

2.8. Data should align with the study aims to ensure findings are meaningful and interpretable.

What to look for:

- collected data that directly addresses what the study intends to investigate
- a clear explanation of why relevant data are not collected (for example, practical constraints, safety limits, or technology availability), acknowledging the primary influences on limitations of the study
- sufficient partial alignment between data and aim when accounting for real-world built environment constraints
- clarity about the study's practical purpose, acknowledging how built environment factors can influence physical, environmental, or system-level performance outcomes (for example, air quality, surface contamination, thermal conditions, or infection-control processes) in complex, interacting ways

- fails if: major gaps exist between aims and collected data without explanation, or the data are incapable of informing the stated aims

When assessing this item, consider:

- the complexity of real-world built environment systems
- requirement for topic-specific or methodological expertise
- potential for subjective judgement

Examples:

2.9. **Good:** In a simulation study assessing how ventilation outlet placement in an intensive care unit (ICU) affects the clearance rate of airborne pathogens from a patient's breathing zone, the researchers collected the precise time in seconds it takes for a tracer gas concentration to drop below 1% under different vent layouts using a high-fidelity computational fluid dynamics simulation of a standard ICU room. The data collected directly mapped to the ventilation design, eliminating external noises (such as staff movement or open doors).

2.10. **Poor:** In a simulation study assessing how ventilation outlet placement in an ICU affects the clearance rate of airborne pathogens from a patient's breathing zone, the researchers only collected data on the total electricity consumed by the ventilation system during the simulation. This data failed to align with the study's objective, as it did not measure pathogen clearance.

Domain 1: Study integrity (4 items)

2.11. This domain assesses whether the study clearly defines what is being examined, is transparent about funding or possible conflicts of interest, and considers other conditions or influences that could affect findings.

Built environment feature or aspect of interest (item 1.1)

Item 1.1: Is the built environment feature or aspect being studied clearly defined?

Why it matters:

2.12. A clear definition of the feature under study is essential for understanding what is being measured and how findings should be interpreted.

What to look for:

- a feature, aspect, or process it is investigating (for example, airflow behaviour, moisture accumulation, surface contamination, or noise propagation)
- a definition of a feature (quantitative, qualitative, or conceptual), that is clear enough to understand what is being studied
- the relative definitions or contextual framing of a feature (for example, ‘elevated humidity for this construction type’) where direct thresholds vary by building type or context
- the relationship between the feature and its wider environment (for example, ventilation, materials, or occupancy), acknowledging that built environment features rarely operate in isolation
- a justification of how the feature is defined for exploratory or emerging topics
- the real-world relevance of a feature where applicable, for example, whether the feature reflects conditions typically encountered in healthcare, residential, or institutional buildings

When assessing this item, consider:

- whether broad study scope is appropriately justified
- the complexity of real-world built environment systems
- requirement for topic-specific or methodological expertise

Examples:

2.13. **Good:** For a study assessing the effect of podular emergency department layout design on spatial efficiency, the exposure variable is defined as a digital three-dimensional podular emergency department layout spanning a fixed 1,200 square metres (m^2) gross floor area, quantified by three self-contained clusters, each featuring a decentralised nurse desk with a mathematically fixed 4.2 metres (m) average travel radius to 8 surrounding beds, and an 85% direct visual field.

2.14. Poor (lack of quantifiable metrics, just using descriptive adjectives): A study simulates a modern podular emergency department layout to see if decentralised pod area improves layout efficiency, but provides no measurable layout parameters, dimensions, visibility metrics, or spatial relationships.

Sources of bias (item 1.2)

Item 1.2: Are potential sources of bias from study design or methods identified and addressed?

Why it matters:

2.15. Identifying and mitigating bias improves the credibility and internal validity of study findings.

What to look for:

- the main anticipated sources of bias and explanations on how they were minimised or accounted for. An exhaustive list is not expected
- potential bias arising from measurement procedures, operator decisions, selection of test conditions, model assumptions, or analysis choices
- mitigation of bias in non-human studies, which may rely on standardised procedures, consistent setup, or independent verification rather than blinding

When assessing this item, consider:

- that not all sources of bias can be fully identified or reported
- the complexity of real-world built environment systems
- requirement for topic-specific or methodological expertise
- potential for subjective judgement

Examples:

Good: A study identifies bias from measurement variability and uses repeated measurements and standardised procedures to minimise it.

Poor: A study fails to discuss potential bias despite clear design limitations.

Confounding control (item 1.3)

Item 1.3: Are other factors outside the main variable of interest identified and addressed?

Why it matters:

2.16. Other influencing factors can affect outcomes and may lead to misleading conclusions if not considered.

What to look for:

- factors such as environmental conditions, material properties, equipment variation, or operational schedules, ensuring these variables are identified and either controlled, monitored, or justified
- the covariates or other secondary factors relevant to the outcome
- a focus on major factors rather than exhaustive lists, including an explanation of the author's decisions regarding included or omitted data

When assessing this item, consider:

- the complexity of real-world built environment systems
- requirement for topic-specific or methodological expertise
- potential for subjective judgement

Examples:

2.17. **Good:** In a study examining how water chemistry affects metal leaching from stainless-steel pipes, researchers identified pipe surface finish as a confounder, since rougher pipes can release more metals regardless of water chemistry. To control for this, they randomly distributed rough and smooth pipes across all water conditions so any differences in metal leaching could be attributed to water chemistry. They also identified stagnation time as a moderator, meaning the effect of water chemistry on metal leaching changes over time. By taking repeated measurements of metal concentrations in the pipes, they found that water chemistry had the strongest effect on metal release in the first week after installation, with the effect fading as the pipe's protective inner layer settled.

2.18. **Poor:** In a study on how water chemistry affects metal leaching from stainless-steel pipes, there was no mention or control of confounders such as pipe surface finish, even though rougher pipes may independently release more metals. The researchers also failed to

account for factors such as stagnation time. They did not take measurements over time, making it unclear whether changes in metal levels were truly due to water chemistry. As a result, the findings are difficult to interpret because multiple uncontrolled factors could explain the observed differences in metal release.

Funding disclosure (item 1.4)

Item 1.4: Are funding sources and conflicts of interest declared?

Why it matters:

2.19. Transparency about funding and involvement helps reviewers assess potential external influences on the study.

What to look for:

- a clear funding and conflict-of-interest statement
- description of the type of support provided, where relevant (for example, financial support, materials, or facilities)
- disclosure of relevant partners (for example, manufacturers or collaborators)
- recognition that commercial funding alone does not necessarily imply bias
- statement of the funder's role in study design, conduct, analysis, and conclusions, where relevant

When assessing this item, consider:

- potential for subjective judgement

Examples:

2.20. **Good:** A study discloses that panels were supplied by Manufacturer X while the design and analysis were conducted independently.

2.21. **Good:** A study indicates that the testing facility was provided in kind by Organisation Y.

2.22. **Poor:** A study provides no statement of funding or conflicts of interest when there are concerns over involvement.

Domain 2: Experimental setup (4 items)

2.23. This domain evaluates whether the study environment, conditions, materials, and procedures are described clearly enough to understand how the experiment was conducted and how results were generated.

Study environment (item 2.1)

Item 2.1: Is the environment adequately described and with sufficient detail?

Why it matters:

2.24. A clear description of the study setting supports understanding of the physical context and assessment of the relevance of findings to real-world conditions.

What to look for:

- description of the study setting: size, layout, boundaries, key features, and any characteristics important to the measured outcomes (for example, orientation, construction type, occupancy, or ventilation strategy)
- a consideration of whether limited environmental detail affects interpretation of the findings, noting that this item may not always apply to laboratory or COmputerised Simulation/ MOdelling (COSMO) studies
- inclusion of visual information (for example, plans, diagrams, or photos), where appropriate.
- description of features relevant to the outcomes (for example, ventilation strategy, or zoning)

When assessing this item, consider:

- the complexity of real-world built environment systems
- potential for inconsistent ratings between appraisers

Examples:

2.25. **Good:** A study employs a mock-up of a 20 square metres (m²) single-patient room with 2.4 metres (m) high ceilings, constructed with plasterboard partitions, ceiling-mounted supply diffuser, and controlled occupancy simulation. Conditions included limited sunlight, external

temperatures between 5–25 degrees Celsius (°C), and external and indoor humidity levels of about 40–90% and 40–60%, respectively.

2.26. **Poor:** A study describes the test environment only as a room having ‘a window and a ventilation system’, with no details on size, layout, or system type.

Environmental conditions (item 2.2)

Item 2.2: Are environmental conditions described and quantified with appropriate detail for interpreting the results?

Why it matters:

2.27. Environmental conditions can strongly influence outcomes in built environment studies, reporting them supports the interpretation and comparison of results.

What to look for:

- measurement and reporting of relevant environmental parameters (for example, temperature, airflow, or humidity)
- inclusion of variables that materially affect the study, such as wind speed or direction, solar gain, background noise levels, or occupancy where relevant
- explanation of practical limitations that prevented full environmental real-world applicability or measurement

When assessing this item, consider:

- the complexity of real-world built environment systems

Examples:

2.28. **Good:** A study reports indoor temperature (21 ± 1°C), outdoor temperature (5–8°C), wind speed (4–6 metres per second (m/s)), prevailing from south-west, and corresponding ventilation rates measured during all test runs.

2.29. **Poor:** A study states that the environment was set up to replicate ‘typical environmental conditions’ but reported no measured values.

Test materials and systems (item 2.3)

Item 2.3: Are materials, systems, or components under study described with sufficient technical detail?

Why it matters:

2.30. A detailed description of materials and systems is essential for understanding performance and enabling reproducibility.

What to look for:

- identification and description of the materials, systems, or components under study
- provision of relevant specifications, such as dimensions, material type, density, treatments, adhesives, coatings, fasteners, construction methods, or system configuration
- inclusion of supporting detail provided through text and or visuals (for example, drawings or schematics)
- reference to specifications or trade names for known or standardised products
- documentation of digital software details (type and version) and input models for COmputerised Simulation/ MOdelling (COSMO) studies
- clear statement of assumptions where full details are unavailable (for example, existing structures, or proprietary systems)

When assessing this item, consider:

- the complexity of real-world built environment systems
- requirement for topic-specific or methodological expertise
- potential for subjective judgement

Examples:

2.31. **Good:** A study specifies a cross laminated timber wall panel composed of five-layer spruce, 120 millimetres (mm) thick, with polyurethane adhesive, and surface-sealed with water-based coating, connected to a steel frame using 8 mm self-tapping screws at 150 mm spacing.

2.32. **Poor:** A study states that timber panels were tested, but provides no information on type, thickness, or configuration.

Intervention implementation (item 2.4)

Item 2.4: Are intervention methods and procedures described clearly and in sufficient detail?

Why it matters:

2.33. Clear procedural description ensures that the intervention can be understood, evaluated, and potentially replicated.

What to look for:

- description of how the intervention was implemented, including steps or sequence, parameter values (for example, intensity, duration, or distance), equipment settings, and number of runs or repetitions

Examples:

2.34. **Good:** For an ultraviolet-C (UV-C) disinfection study, the intervention is described as lamp intensity (20 milliwatts per square centimetre (mW/cm²)), exposure duration (15 minutes), distance to target (2 metres (m)), room airflow state ('fan on'), three repeated cycles per condition, and comparison to manufacturer-recommended exposure times.

2.35. **Poor:** For a UV-C disinfection study, the intervention is described as a UV-C lamp (approximately 15–25 mW/cm²) used at a distance of about 1–3 m for 15–20 minutes per session in a room with the fan on.

Domain 3: Measurement approach (4 items)

2.36. This domain evaluates how well the study measures what it aims to measure, including the clarity of data collection procedures, the suitability and transparency of measurement tools, and the use of quality-control and validation processes.

Data collection methods (item 3.1)

Item 3.1: Are data collection procedures systematic and described in sufficient detail?

Why it matters:

2.37. Clear and structured data collection ensures that results are reliable, consistent, and reproducible.

What to look for:

- description of how data were collected, including sampling frequency, number of measurements, positioning, duration, and how data were logged
- evidence of a systematic or standardised approach
- information on accuracy, precision, calibration checks, where appropriate, and whether methods align with recognised standards
- description of any constraints (for example, field-based limitations or seasonal variation) that affect the approach

When assessing this item, consider:

- potential for inconsistent ratings between appraisers

Examples:

2.38. **Good:** For an indoor air quality study, carbon dioxide (CO_2) is measured every 5 minutes using calibrated non-dispersive infrared sensors (±50 parts per million accuracy) placed at child head height (1.0 metre above finished floor level) in four classrooms over three school days (36 hours).

2.39. **Poor:** For an indoor air quality study, CO_2 is measured periodically in classrooms using sensors, with no information on placement, calibration, or duration.

Instrumentation description (item 3.2)

Item 3.2: Are measurement tools and equipment described in enough detail to understand their capabilities and limitations?

Why it matters:

2.40. Understanding the performance and limitations of measurement tools is essential for interpreting results and assessing reliability.

What to look for:

- specification of instruments with model, manufacturer, measurement range, accuracy, and relevant settings (for example, sampling rate, or filter type)
- details of name, version, and configuration parameters for software or simulation tools
- information on calibration status, maintenance logs, or verification procedures if available

Examples:

2.41. **Good:** A study describes a Class 1 sound level meter (International Electrotechnical Commission (IEC) 61672 compliant), with a range of 20–150 decibels (dB), A-weighting, 1-second logging interval, and calibration performed two months prior.

2.42. **Poor:** A study states that noise levels were measured with a sound meter but provides no further details.

Instrumentation justification (item 3.3)

Item 3.3: Is the choice of measurement tools explained and appropriate for the study?

Why it matters:

2.43. Justification ensures the selected methods are suitable for the outcomes being measured.

What to look for:

- justification including accuracy needs, compatibility with the environment (lab, field or simulation), compliance with standards, or prior validation

- application of measurement tools' including sensors, models, simulation tools (for example, computational fluid dynamics (CFD)), scoring frameworks, surveys, and observer protocols
- explanation of why the chosen method was selected, where multiple options exist
- recognition of tool limitations where relevant

When assessing this item, consider:

- the complexity of real-world built environment systems

Examples:

2.44. **Good:** A study justifies the selection of a computational fluids dynamics (CFD) turbulence model (k–ε) due to its prior validation in similar indoor airflow studies and its match with the spatial scale required for ward-level simulations.

2.45. **Poor:** A study uses a simulation model with no explanation of why it was chosen or whether it is appropriate.

Validation (item 3.4)

Item 3.4: Are appropriate data quality control or validation measures used to ensure data reliability?

Why it matters:

2.46. Validation helps ensure that the data produced are accurate, reliable, and suitable for analysis.

What to look for:

- evidence of validation or quality control measures, such as:
  - calibration before and after measurement
  - use of duplicate or parallel measurements
  - comparison with reference or 'gold standard' methods
  - internal or external controls
  - justification of the number of repetitions
- reporting of detection limits, uncertainty, or error where relevant
- description of how missing, outlier, or noisy data were handled

Examples:

2.47. **Good:** A study validates its data by recording temperature and humidity using duplicate sensors and cross-checking the recorded values against a calibrated reference instrument.

2.48. **Poor:** A study reports data with no mention of calibration, replication, or quality checks.

Domain 4: Outcomes and analysis (3 items)

2.49. This domain evaluates whether a study clearly defines its outcomes, measures them with appropriate, defensible methods, and analyses its data with transparent, methodologically rigorous approaches.

Outcome definition (item 4.1)

Item 4.1: Are study outcomes clearly defined, measurable, and described using appropriate units or criteria?

Why it matters:

2.50. Clear outcome definition ensures that results are interpretable, comparable, and directly linked to the study aim.

What to look for:

- explicit description of what is being measured (the outcome), including:
  - clarification of whether outcomes are quantitative, qualitative, or mixed methods
- reporting in absolute or relative terms, depending on study context
- key components or indicators used for multidimensional outcomes (for example, ‘thermal comfort’)
- differentiation between the outcome definition from earlier items (for example, distinguishing a built environment feature definition from what is actually being assessed)

Examples:

2.51. **Good:** A study defines its primary outcome measure as the peak fine particulate matter ($PM_{2.5}$) concentration (micrograms per cubic meter (µg/ m^3)) 1 hour after intervention.

2.52. **Poor:** A study reports that air quality improved but does not define specific outcome metrics.

Measurement suitability (item 4.2)

Item 4.2: Are the measurement methods appropriate and fit-for-purpose for the outcomes being assessed?

Why it matters:

2.53. Even when tools are well described (Domain 3), they must also be appropriate for capturing the intended outcome.

What to look for:

- clear alignment between the defined outcome and the measurement method used
- evidence that the method:
  - directly captures the outcome of interest
  - is valid for the type of measurement
  - is sensitive enough to detect expected changes
- a brief explanation of why the chosen method is most appropriate where multiple approaches are possible

When assessing this item, consider:

- the complexity of real-world built environment systems
- requirement for topic-specific or methodological expertise

Examples:

2.54. **Good:** A study uses light metres to calculate glare indices when assessing visual comfort.

2.55. **Poor:** A study uses window size alone as a proxy for glare without measuring light levels.

Analytical method (item 4.3)

Item 4.3: Are the analytical methods appropriate for the data, clearly justified, and described with enough detail for interpretation?

Why it matters:

2.56. Appropriate analysis ensures that results are valid, interpretable, and reproducible.

What to look for:

- description of the analytical approach, including:
  - specific methodology used, such as statistical methods, simulation analysis, or qualitative analysis
- alignment between the analysis and:
  - type of data (for example, continuous, categorical, or model outputs)
  - study design and objectives
- reporting of key methodological elements, such as:
  - model or test selection
  - assumptions (for example, normality, or independence)
  - software names and version numbers
- explanation of how data issues were handled, including:
  - missing data
  - outliers
  - censored or bounded values
- reporting of hypothesis testing or model validation approaches, where relevant

When assessing this item, consider:

- potential for inconsistent ratings between appraisers

Examples:

2.57. Good (statistical): A study applies a generalised linear model after confirming a Poisson distribution and performs the analysis in R version 4.2.

2.58. Good (simulation): A study analyses computational fluid dynamics (CFD) outputs using validated flow visualisation techniques and compares the results with measured data.

2.59. Good (qualitative): A study conducts a thematic analysis independently using two reviewers with a consensus resolution process.

2.60. **Poor:** A study states that data were analysed using statistical software with no further detail.

Domain 5: Reporting and transparency (5 items)

2.61. This domain evaluates how clearly and openly a study reports its results, limitations, and contextual benchmarks, and whether it provides enough detail and access for others to understand, interpret, and potentially replicate its findings.

Results presentation (item 5.1)

Item 5.1: Are results clearly presented and interpreted in a way that aligns with the study aims?

Why it matters:

2.62. Clear reporting helps understand what was found, how the findings relate to the study aims, and how much confidence can be placed in the conclusions.

What to look for:

- appropriate presentation of results (for example, using clear tables, figures, narrative descriptions, and, where relevant, statistical summaries)
- clear linkage between results and the stated aims, objectives, or hypotheses
- reporting of key patterns, magnitudes, and uncertainties

Examples:

2.63. **Good:** A study reports its results as a 43% reduction in mean peak fine particulate matter ($PM_{2.5}$) concentration ($p < 0.01$) after intervention, supported by a bar chart with confidence intervals.

2.64. **Poor:** A study describes results only as ‘improved’ or ‘reduced’ without numerical values, figures, or explanation.

Study limitations (item 5.2)

Item 5.2: Are study limitations acknowledged?

Why it matters:

2.65. Transparent reporting of limitations helps readers judge the strength, relevance, and applicability of the findings.

What to look for:

- acknowledgement of the main limitations affecting interpretation of results
- limitations relating to scope, sample, environmental representativeness, assumptions, measurement constraints, or generalisability. Only the main limitations need to be acknowledged
- provision of context for limitations (for example, ‘winter not tested due to equipment constraints. Results apply to summer conditions only’)

When assessing this item, consider:

- potential for subjective judgement

Examples:

2.66. **Good:** A study states that airflow patterns may differ under winter conditions because heating was not in use during the test period.

2.67. **Poor:** A study includes no discussion of limitations despite clear methodological constraints.

Results benchmarking (item 5.3)

Item 5.3: Are results contextualised against relevant standards, guidelines, or prior research?

Why it matters:

2.68. Contextualising findings against external benchmarks supports interpretation and helps determine how meaningful or applicable the results are.

What to look for:

- clear identification, dating, and description of benchmarks used (including limitations of those standards)
- consideration of whether benchmarks are outdated or based on different calculation methods when interpreting results
- inclusion of appropriate benchmarks for comparative referencing, such as regulatory thresholds, industry standards, prior empirical findings, or typical ranges in real buildings
- explanation of methodological or contextual differences where these affect comparability

When assessing this item, consider:

- potential for inconsistent ratings between appraisers

Examples:

2.69. **Good:** A study compares ventilation rates with World Health Organization (WHO) or American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) guidance, acknowledging differences in methods.

2.70. **Poor:** A study describes results as ‘acceptable’ or ‘within expected range’ without identifying the reference standard.

Reproducibility (item 5.4)

Item 5.4: Is enough methodological detail provided for others to reproduce the study?

Why it matters:

2.71. Reproducibility depends on transparent reporting of the methods, parameters, settings, and analytical steps used.

What to look for:

- inclusion of sufficient detail on procedures, parameters, settings, materials, and analytical steps
- clear statement of key parameters even when methods reference established protocols
- provision of enough contextual and procedural information in field or real-world studies to support partial reproducibility
- justification for instances where full reproducibility is unfeasible due to complex, real-world settings

When assessing this item, consider:

- the complexity of real-world built environment systems
- do not penalise real-world variability where it reflects realistic conditions rather than poor study design
- potential for subjective judgement

Examples:

2.72. **Good:** A computational fluid dynamics (CFD) study reports boundary conditions, mesh settings, solver details, and software version, or an acoustic study specifies microphone height, source location, mounting method, and test sequence.

2.73. **Poor:** A study describes methods only in broad terms, with no parameter values, procedural detail, or settings.

Data accessibility (item 5.5)

Item 5.5: Are data, protocols, or models accessible, or is access clearly explained and justified?

Why it matters:

2.74. Transparency is strengthened when data, protocols, or models are shared or when limitations on access are clearly explained.

What to look for:

- raw data, processed data, protocols, or models being shared in repositories where possible
- justification and description of what can be shared where full sharing is impossible due to privacy, security, or commercial restrictions
- specification of accessibility channels, such as data repositories, supplementary files, open-source code, model parameters, or metadata

Examples:

2.75. **Good:** A study deposits raw temperature data in an institutional repository with a digital object identifier (DOI).

2.76. **Good:** A study shares computational fluid dynamics (CFD) model files with proprietary geometry removed and replaced by anonymised mesh files or makes its protocol available on request under a non-disclosure agreement.

2.77. **Poor:** A study omits a data availability statement and provides no explanation of whether materials can be accessed.

Domain 6: Transferability (3 items)

2.78. This domain captures methodological considerations specific to the three main subtypes of non-human experimental studies in the built environment. These items should be rated only where relevant. The study types cover laboratory testing (in vitro), field (in situ), and COmputerised Simulation/ MOdelling (COSMO), as previously introduced in section 1.2.

2.79. If an item does not apply to the study design, rate it Not Applicable (N/A).

Laboratory conditions in laboratory testing (in vitro) studies (item 6.1)

Item 6.1: Are laboratory environmental conditions controlled, reported, and appropriately representative of the intended real-world application?

Why it matters:

2.80. In laboratory studies, controlled conditions support internal consistency, while clear reporting helps judge how far the findings may apply beyond the test environment.

What to look for:

- justification for instances where laboratory conditions differ substantially from real-world conditions, allowing this item to be marked as ‘Not Applicable (N/A)’ if the study has limited relevance beyond a highly specific laboratory context
- explanation of how conditions relate to, or differ from, the intended real-world application
- accreditation or compliance with recognised laboratory standards, where relevant

When assessing this item, consider:

- the complexity of real-world built environment systems
- do not penalise real-world variability where it reflects realistic conditions rather than poor study design

Examples:

2.81. **Good:** A study maintains an acoustic chamber at 23 degrees Celsius (°C) and 50% relative humidity to approximate typical indoor conditions.

2.82. **Poor:** A study states that testing was ‘carried out under laboratory conditions,’ with no further description.

Field context description in field (in situ) studies (item 6.2)

Item 6.2: Is the real-world context described in sufficient detail?

Why it matters:

2.83. Field and real-world studies are shaped by contextual factors that affect interpretation, transferability, and relevance.

What to look for:

- description of the real-world setting sufficient to understand how findings may generalise
- inclusion of contextual drivers (for example, wind, shading, or occupancy schedules) where relevant
- evidence of built environment-specific expertise in the interpretation of field context
- inclusion of visual or spatial documentation (for example, plans, photos, or diagrams) where possible

When assessing this item, consider:

- requirement for topic-specific or methodological expertise
- potential for subjective judgement

Examples:

2.84. **Good:** An airflow study describes the field site as a three-storey office building (built 1990), with 25 staff, occupancy from 8am to 6pm, south-facing windows, and prevailing south-west wind during monitoring.

2.85. **Poor:** A study reports results from a ‘real building’ with no information on building type, occupancy, layout, or operating conditions.

Model validation in COmputerised Simulation/ MOdelling (COSMO) studies (item 6.3)

Item 6.3: Are computational or physical models validated against empirical data or credible benchmarks, and is their applicability to the intended setting explained?

Why it matters:

2.86. Validation helps establish that a model is credible and suitable for representing the system or conditions of interest.

What to look for:

- comparison of models with empirical measurements, controlled tests, or well-established benchmarks
- inclusion of a validation statement explaining how well the model represents the real-world space or scenario of interest
- acknowledgement and justification where validation is limited or indirect

When assessing this item, consider:

- the complexity of real-world built environment systems
- do not penalise real-world variability where it reflects realistic conditions rather than poor study design

Examples:

2.87. **Good:** A study validates its computational fluid dynamics (CFD) model against tracer-gas decay measurements in a full-scale test room, with a discussion of how well the room reflects the target setting.

2.88. **Poor:** A study presents simulation results without any comparison to real measurements, tests, or recognised benchmarks.

Appendix A Tool development and validation

A.1 Built environment Experimental studies Critical Appraisal Tool (BECAT) was developed to address the lack of a standardised critical appraisal tool for non-human experimental studies in the built environment, which has limited consistent evaluation of evidence in this field.

A.2 The tool was developed using a multi-stage, iterative methodology:

- scoping review of existing critical appraisal tools
- framework development based on extracted appraisal criteria
- mapping of built environment studies to refine domains and item relevance
- draft tool development
- Delphi consensus process to evaluate the clarity and importance of items

A.3 The Delphi study involved multiple rounds of expert consultation, using structured feedback and quantitative consensus measures to refine and validate the tool content.

A.4 Planned future work to refine the BECAT includes:

- inter-rater reliability testing
- further validation when applied to literature reviews
- refinement based on user feedback

A.5 BECAT currently includes:

- 25 items
- seven domains covering screening, methodological quality, reporting, and transferability

Table A.1 - Built environment Experimental Critical Appraisal Tool (BECAT) Checklist Domains and Descriptions

Domain	Description
Minimum quality criteria	Screening domain
Study integrity	Internal validity and study design
Experimental setup	Design and execution of the experiment
Measurement approach	Data collection methods and tools
Outcomes and analysis	Analytical methods and interpretation
Reporting and transparency	Completeness and clarity of reporting
Transferability	Context considerations for subtypes

Appendix B Study design identification guidance

B.1 Experimental studies in the built environment involve deliberate manipulation of one or more environmental variables to evaluate their effect on a defined outcome.

B.2 These studies:

- do not directly involve human participants
- focus on physical, environmental, or system-level outcomes such as:
  - airflow and ventilation performance
  - microbial growth
  - material performance or durability
  - environmental conditions (for example, temperature, lighting, or acoustics)

B.3 A study is likely to be included if it demonstrates:

- active manipulation of one or more variables
- a defined intervention, exposure, or condition
- a comparison between conditions (for example, control versus intervention, before and after, or scenario testing)
- implementation in laboratory, simulated, or real-world environments

B.4 Eligible designs may include:

- controlled laboratory experiments
- quasi-experimental designs (for example, pre–post)
- repeated-measures or scenario-based testing
- computational simulation or modelling studies

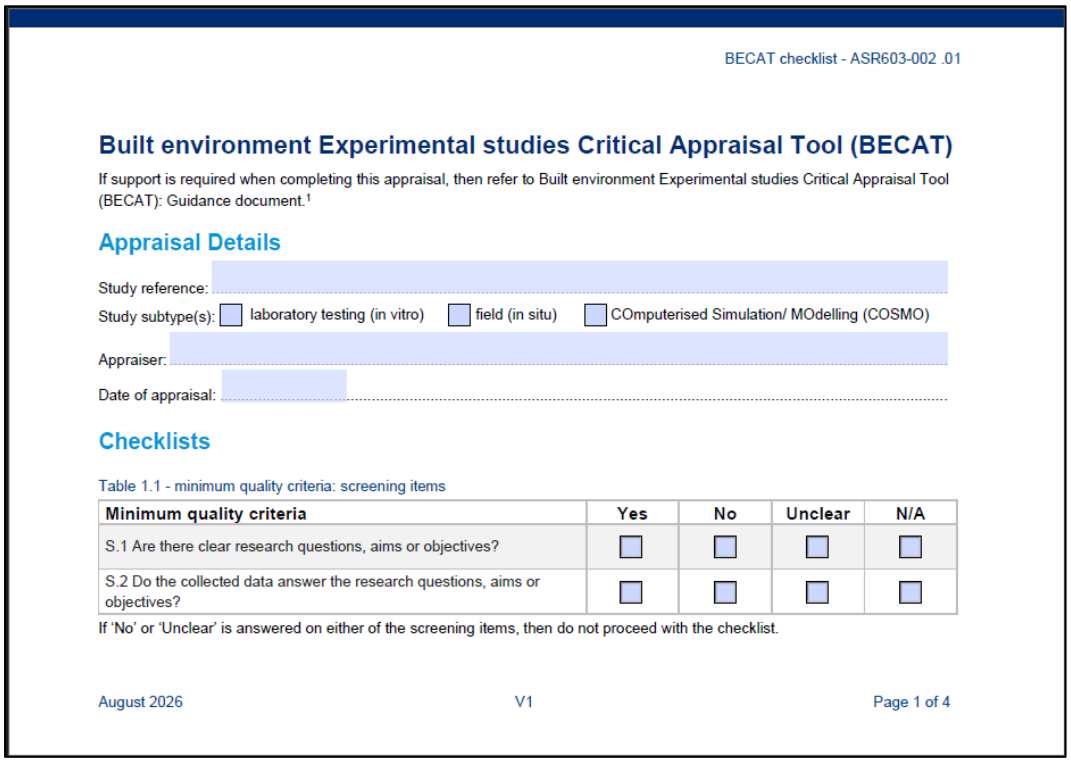
B.5 Studies are typically excluded if they:

- do not manipulate variables (purely observational)
- focus solely on human outcomes (studies that assess both human and non-human outcomes are eligible for inclusion; however, appraisal using this tool should be limited to the non-human outcomes)
- lack a defined intervention or comparison

Appendix C Built environment experimental studies critical appraisal tool (BECAT)

C.1 The Built environment Experimental studies Critical Appraisal Tool (BECAT) is available as a fillable checklist; click the following link to access the landing page for Built environment Experimental studies Critical Appraisal Tool (BECAT): Guidance and Checklist and open the file 'ASR602-002.001'.

Figure C.1 Built environment Experimental studies Critical Appraisal Tool (BECAT): Checklist Page 1 of 4



BECAT checklist - ASR603-002 .01

Built environment Experimental studies Critical Appraisal Tool (BECAT)

If support is required when completing this appraisal, then refer to Built environment Experimental studies Critical Appraisal Tool (BECAT): Guidance document.[1]

Appraisal Details

Study reference:

Study subtype(s): ☐ laboratory testing (in vitro) ☐ field (in situ) ☐ COmputerised Simulation/ MOdelling (COSMO)

Appraiser:

Date of appraisal:

Checklists

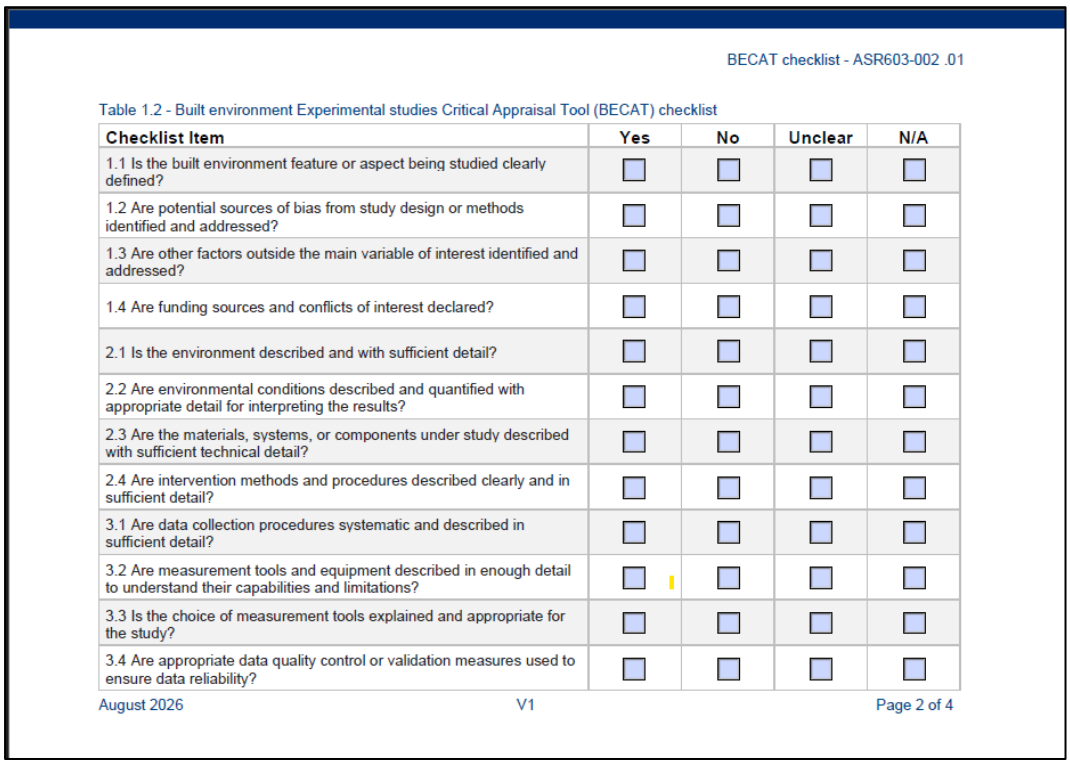
Table 1.1 - minimum quality criteria: screening items

Minimum quality criteria	Yes	No	Unclear	N/A
S.1 Are there clear research questions, aims or objectives?	☐	☐	☐	☐
S.2 Do the collected data answer the research questions, aims or objectives?	☐	☐	☐	☐

If 'No' or 'Unclear' is answered on either of the screening items, then do not proceed with the checklist.

August 2026 V1 Page 1 of 4

Figure C.2 - Built environment Experimental studies Critical Appraisal Tool (BECAT): Checklist Page 2 of 4



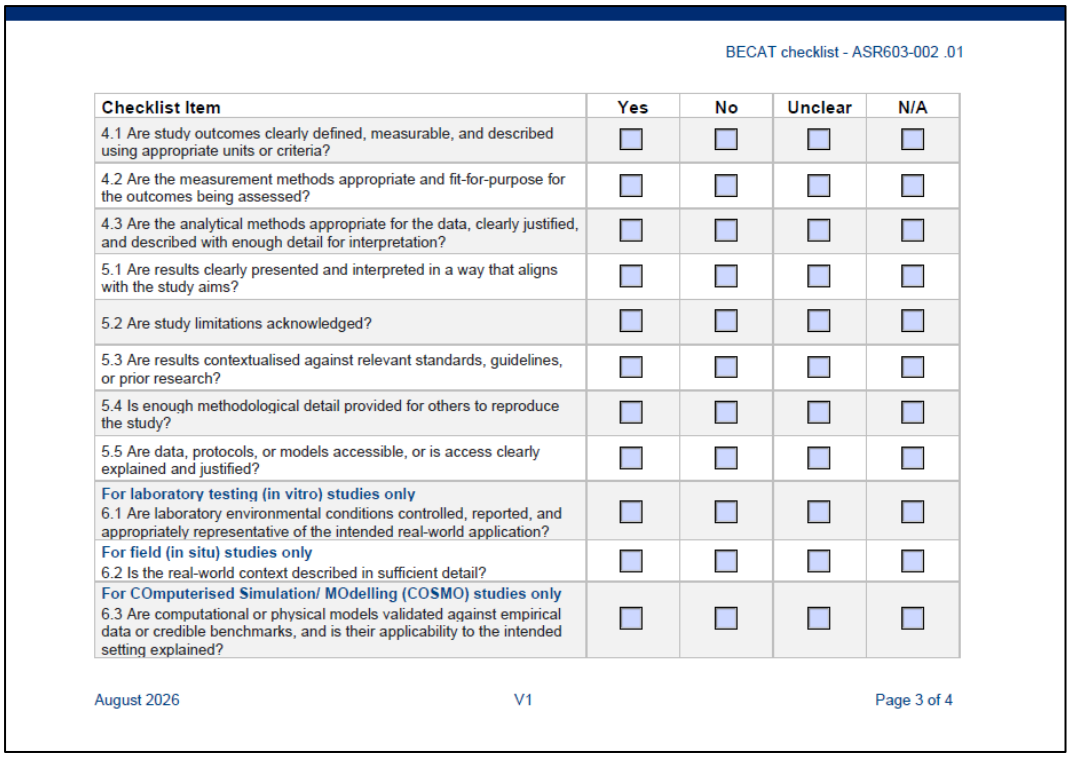
BECAT checklist - ASR603-002 .01

Table 1.2 - Built environment Experimental studies Critical Appraisal Tool (BECAT) checklist

Checklist Item	Yes	No	Unclear	N/A
1.1 Is the built environment feature or aspect being studied clearly defined?	☐	☐	☐	☐
1.2 Are potential sources of bias from study design or methods identified and addressed?	☐	☐	☐	☐
1.3 Are other factors outside the main variable of interest identified and addressed?	☐	☐	☐	☐
1.4 Are funding sources and conflicts of interest declared?	☐	☐	☐	☐
2.1 Is the environment described and with sufficient detail?	☐	☐	☐	☐
2.2 Are environmental conditions described and quantified with appropriate detail for interpreting the results?	☐	☐	☐	☐
2.3 Are the materials, systems, or components under study described with sufficient technical detail?	☐	☐	☐	☐
2.4 Are intervention methods and procedures described clearly and in sufficient detail?	☐	☐	☐	☐
3.1 Are data collection procedures systematic and described in sufficient detail?	☐	☐	☐	☐
3.2 Are measurement tools and equipment described in enough detail to understand their capabilities and limitations?	☐	☐	☐	☐
3.3 Is the choice of measurement tools explained and appropriate for the study?	☐	☐	☐	☐
3.4 Are appropriate data quality control or validation measures used to ensure data reliability?	☐	☐	☐	☐

August 2026 V1 Page 2 of 4

Figure C.3 - Built environment Experimental studies Critical Appraisal Tool (BECAT): Checklist Page 3 of 4



BECAT checklist - ASR603-002 .01

Checklist Item	Yes	No	Unclear	N/A
4.1 Are study outcomes clearly defined, measurable, and described using appropriate units or criteria?	☐	☐	☐	☐
4.2 Are the measurement methods appropriate and fit-for-purpose for the outcomes being assessed?	☐	☐	☐	☐
4.3 Are the analytical methods appropriate for the data, clearly justified, and described with enough detail for interpretation?	☐	☐	☐	☐
5.1 Are results clearly presented and interpreted in a way that aligns with the study aims?	☐	☐	☐	☐
5.2 Are study limitations acknowledged?	☐	☐	☐	☐
5.3 Are results contextualised against relevant standards, guidelines, or prior research?	☐	☐	☐	☐
5.4 Is enough methodological detail provided for others to reproduce the study?	☐	☐	☐	☐
5.5 Are data, protocols, or models accessible, or is access clearly explained and justified?	☐	☐	☐	☐
For laboratory testing (in vitro) studies only 6.1 Are laboratory environmental conditions controlled, reported, and appropriately representative of the intended real-world application?	☐	☐	☐	☐
For field (in situ) studies only 6.2 Is the real-world context described in sufficient detail?	☐	☐	☐	☐
For COmputerised Simulation/ MOdelling (COSMO) studies only 6.3 Are computational or physical models validated against empirical data or credible benchmarks, and is their applicability to the intended setting explained?	☐	☐	☐	☐

August 2026 V1 Page 3 of 4

Figure C.4 - Built environment Experimental studies Critical Appraisal Tool (BECAT): Checklist Page 4 of 4

BECAT checklist - ASR603-002 .01

Overall appraisal

- ☐ very minor methodological limitations or risk of bias
- ☐ minor methodological limitations or risk of bias
- ☐ moderate methodological limitations or risk of bias
- ☐ serious methodological limitations or risk of bias
- ☐ serious methodological concerns: failed screening questions

[1] Public Services Delivery Scotland, NHS Scotland Assure, Healthcare Scientist Technical Team. Built environment Experimental studies Critical Appraisal Tool (BECAT): Guidance document. 2026

August 2026 V1 Page 4 of 4

Abbreviations

ARHAI: Antimicrobial Resistance and Healthcare Associated Infection

ASHRAE: American Society of Heating, Refrigerating and Air-Conditioning Engineers

BECAT: Built environment Experimental studies Critical Appraisal Tool

CO_2: carbon dioxide

CFD: Computational Fluid Dynamics

COSMO: COmputerised Simulation/ Modelling

dB: decibels

DOI: Digital Object Identifier

DPIA: Data Protection Impact Assessment

EGIS: Energy, Geoscience, Infrastructure and Society

HAI-SCRIBE: Healthcare Associated Infection System for Controlling Risk in the Built Environment

ICU: Intensive Care Unit

IEC: International Electrotechnical Commission

ISBE: Institute for Sustainable Built Environment

µg/ m^3: micrograms per cubic meter

mW/cm^2: milliwatts per square centimetre

NSS: National Services Scotland

PICO: Population, Intervention or exposure, Comparator, and Outcomes

PSD Scotland: Public Service Delivery Scotland

SAMS: Strategic Asset Management System

UV-C: ultraviolet-C

WHO: World Health Organization

DPIA: Data Protection Impact Assessment.

IEC: International Electrotechnical Commission.