Report on Oxevision Evidential Base and Case Study Methodologies

This report is prepared by

Professor Christl Donnelly CBE FRS FMedSci FAcadMathSci

and

Dr Maria Christodoulou CStat

Declaration of Interests and Independence

We confirm that we have no financial, professional or personal interest in Oxevision (Oxehealth/LIO), and have not undertaken consultancy work for, or received funding, from these organisations.

We were instructed to undertake this review by the Lampard Inquiry through Oxford University Innovation Limited ("OUI"), which acted solely as an administrative and contracting intermediary. OUI had no role in the design, analysis, interpretation, or conclusion of this report.

We are both employees of the University of Oxford. Oxevision is a product of Oxehealth, a spin-out from the University of Oxford, and OUI forms parts of the University's innovation infrastructure. We recognise that this institutional context may give rise to a perception of proximity and have taken this into account in our assessment.

The evidential base comprised academic literature, case studies, and other materials provided through a range of sources, including Inquiry materials and journal papers. We have taken account of the limitations inherent in this mixed evidential base, including that it was not derived from a fully systematic literature search.

We confirm that our analysis and conclusions were reached independently and have not been influenced by the instruction route or evidential sourcing described above.

In reviewing the evidential base, we also considered relevant institutional and academic contextual relations where those arose within the materials, and these did not affect our assessment or conclusion.

We understand that our overriding duty is to assist the Chair to the Lampard Inquiry on matters within our expertise. In preparing this report we have complied with that duty and have exercised our independent professional judgment at all times.

We confirm that the opinions expressed in this report are our own, represent our true and complete professional opinions on the matters to which they relate, and are based upon the information and materials available to us. Where matters fall outside our expertise, or where there are limitations in the available evidence, we have sought to identify those matters within the report.

We confirm that we have not omitted to consider any material fact or matter of which we are aware that might reasonably be expected to affect the opinions expressed.

We further confirm that neither our opinions nor the conclusions reached in this report have been influenced by any personal, financial, institutional professional or other interest.

Executive summary

This report reviews the available evidence relating to Oxevision and related Vision-Based[1] Patient Monitoring Systems (VBPMSs), with particular attention to their use in mental health inpatient settings. The evidence considered includes clinical service evaluations, economic models, technical validation studies, qualitative and ethical research, and wider digital health literature.

- The overall evidence base is limited. The studies most directly relevant to mental health inpatient settings are largely observational, before-and-after, non-randomised, or service evaluation studies. Many involve small numbers of wards, limited comparator groups, routine incident-reporting data, or local implementation contexts. These designs make it difficult to distinguish any effect of the technology from other possible explanations, including changes in patient case mix, staffing, ward practice, incident reporting, service pressures or wider safety initiatives.
- Some studies and economic models report favourable findings following implementation of Oxevision, including suggested reductions in adverse incidents, observation requirements or staffing costs. However, these findings should be interpreted cautiously. They do not provide robust evidence that Oxevision caused reductions in self-harm, ligatures, assaults, falls or other safety outcomes. Nor do they establish that reported cost savings would be realised consistently across NHS mental health services. Several economic conclusions depend on uncertain assumptions, particularly around reductions in enhanced observation hours and the transferability of local data to wider NHS settings.
- Independence is a concern. A number of key publications were funded by Oxehealth, authored by Oxehealth employees, involved Oxehealth in data collection or analysis, or relied on data supplied by Oxehealth. One relevant article was subsequently retracted because conflicts of interest were not disclosed in the required statement. These issues do not make the evidence irrelevant, but they materially reduce the weight that can be placed on strong claims about effectiveness, safety or cost-effectiveness.
- The technical validation literature provides some support for the feasibility of non-contact video-based physiological monitoring under suitable conditions. However, much of this evidence comes from settings outside mental health inpatient care, such as neonatal intensive care, haemodialysis, intensive care, and controlled volunteer studies. It demonstrates technical plausibility rather than clinical benefit in mental health inpatient wards.
- The qualitative, ethical, and policy literature identifies important unresolved issues. Patients and staff may perceive potential benefits, but concerns remain about privacy, dignity, consent, surveillance, data governance, staff monitoring, function creep, anxiety, and the impact of monitoring technologies on

[1] Here we use two similar terms, *vision-based* and *video-based*, following the usage in the reviewed evidence. Both refer to systems which utilise camera technology. The visual data collected might be viewed in real time and/or reviewed later if required.

therapeutic relationships. The available lived-experience evidence is limited, and patient outcomes and rights-based considerations have received less attention than operational and safety claims.

In conclusion, the current evidence provides only limited and cautious support for the use of Oxevision and related VBPMSs in mental health inpatient settings. It suggests that the technology is technically feasible and may have potential operational applications, but it does not provide strong independent evidence of clinical effectiveness, improved patient safety, acceptability or cost-effectiveness in mental health inpatient settings. Further independent, randomised, adequately controlled research would be needed before stronger conclusions could be drawn.

Statistical considerations in evaluating evidence and methodologies

This section sets out several methodological principles relevant to interpreting the evidence considered in this report.

Medical research studies vary in the extent to which they can support conclusions about whether an intervention has caused an observed outcome. A study should be assessed by reference to its research question, design, comparator, population, setting, outcome measures, data quality, statistical methods, and potential sources of bias. These features affect the weight that can properly be placed on its findings.

A central distinction is between association and causation. A study may show that an outcome changed after an intervention was introduced or that outcomes differed between settings using and not using an intervention. That is evidence of an association. It is not, by itself, evidence that the intervention caused the outcome. Causal interpretation requires consideration of alternative explanations, including differences in patient population, staffing, clinical practice, reporting behaviour, organisational context, and other contemporaneous changes. These alternative explanations may include confounding, where factors other than the intervention are associated with both the introduction of the intervention and the outcome being measured.

Study design is important. Randomised controlled trials reduce some forms of confounding because allocation to the intervention or comparator is determined by chance. In many service settings, randomisation may be difficult, but controlled designs can still strengthen inference if the comparator is appropriate and the groups are sufficiently similar. Before-and-after studies are weaker for causal inference because they compare outcomes across time. Changes observed after implementation may reflect the intervention, but may also reflect background trends, service changes, changes in reporting, regression to the mean, or differences in case mix.

For example, an intervention might be the introduction of alcohol-based hand sanitiser for ward staff with the aim of reducing hospital-acquired infections. In a before-and-after comparison, it could appear that the intervention caused a reduction in infection incidence. However, there could be confounders that differ between the before (no hand sanitiser) and after (hand sanitiser provided) period, for example season of the year, other policies such as cleaning protocols, or programmes to screen for infections. As a result, while there might be a substantial difference between the hospital-acquired infections before and after the provision of hand sanitiser, it might not be the hand sanitiser that *caused* the reduction in infections.

The size and setting of a study also affect interpretation. Small studies may have limited ability to estimate effects precisely or detect uncommon outcomes. Single-centre studies may be strongly influenced by local factors, including ward layout, staffing, leadership, patient mix, implementation support, and organisational culture. Findings

from such studies may be relevant outside of the study context, but their transferability to other NHS mental health settings should not be assumed.

Statistical analysis is used to describe and test patterns in data. A p-value is a measure of how compatible the observed data are with a specified null hypothesis, usually that there is no difference or no association. Statistical significance is commonly defined using a threshold such as a p-value less than 0.05, but this threshold should not be treated as a dividing line between true and false findings. Instead, p-values should be considered as part of a broader assessment of the evidence. Statistical significance does not remove bias, confounding, or measurement error. Non-significance does not prove absence of an effect.

Effect estimates should be interpreted alongside measures of uncertainty, such as confidence intervals. A confidence interval is an estimated interval within which an unknown parameter may plausibly lie within the context of the statistical model used. Wide intervals indicate imprecision. Traditionally 95% confidence intervals are used. Where a 95% confidence interval includes values consistent with little or no effect, strong conclusions about benefit should be avoided. The clinical or practical importance of an estimated effect should also be considered separately from statistical significance.

For a technology such as Oxevision, a high-quality evaluation would ideally use a clearly defined intervention and comparator, prespecified outcomes, appropriate adjustment for confounding, adequate sample size, transparent reporting of uncertainty, and sufficient follow-up. It would also examine possible unintended consequences, including effects on patient experience, privacy, dignity, consent, equality, staff workload, therapeutic relationships, and ward practice. The independence of a study is also relevant to the weight placed on it. Research that is funded, designed, analysed, or authored by parties with a commercial or institutional interest in the intervention may still provide useful evidence. However, such evidence requires particularly careful scrutiny, particularly where data access, analysis decisions, or reporting are not fully independent. Independent replication is important before strong conclusions are drawn.

Economic models require particular care. A model is not direct observation of what will happen in practice. It is a structured estimate of likely costs, savings, or outcomes based on specified inputs and assumptions. In the context of Oxevision, relevant assumptions may include ward occupancy, staffing levels, the cost of enhanced observations, incident rates, implementation costs, and the extent to which the technology changes staff activity. If the inputs are uncertain, derived from weak evidence, or not generalisable to other settings, the model outputs will also be uncertain.

The World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP)[2] has been designed "to ensure that a complete view of research is accessible to all those involved in health care decision-making". The ICTRP website states that "This will improve research transparency and will ultimately strengthen the validity and value of the scientific evidence base. The registration of all interventional trials is a scientific, ethical and moral responsibility." A key goal of the prospective registration of all clinical trials is to avoid selective reporting of results and publication bias. Publication bias can arise when studies showing evidence of a benefit from an intervention are more likely to be accepted for publication in scientific journals than studies that do not show a benefit.

These principles do not mean that limited studies or models have no value. They may identify plausible mechanisms, support feasibility, inform hypotheses and the design of future studies, and/or assist decision-making where stronger evidence is unavailable. However, the strength of any conclusion should reflect the quality, independence, and applicability of the underlying evidence.

[2] https://www.who.int/tools/clinical-trials-registry-platform

Methodology

This report is a structured evidence review. The purpose was to assess the nature, relevance, and strength of the available evidence relating to Oxevision and related Vision-Based Patient Monitoring Systems (VBPMSs), particularly in mental health inpatient settings.

Each item of evidence was reviewed by reference to its study design, setting, population, comparator, outcomes, statistical methods, reporting of uncertainty, relevance to the review question, and independence from the technology supplier. Particular attention was given to whether a study was capable of supporting causal conclusions, or whether it was better understood as feasibility, implementation, technical validation, contextual or hypothesis-generating evidence.

The review did not apply a formal numerical scoring system. Instead, evidential weight was assessed using a qualitative judgement. In broad terms:

- **Low evidential weight** was assigned where evidence was indirect, descriptive, uncontrolled, very small, single-site, primarily contextual, or subject to major limitations affecting interpretation. This included narrative commentary, conference abstracts, service descriptions, small before-and-after studies without adequate controls, and studies with limited relevance to mental health inpatient settings.
- **Low-to-moderate evidential weight** was assigned where studies were directly relevant and provided useful empirical evidence, but had important limitations such as non-randomised designs, small numbers of wards or patients, limited comparator groups, reliance on routine incident data, incomplete adjustment for confounding, limited reporting of statistical uncertainty, or material independence concerns.
- **Moderate evidential weight** was assigned where evidence was relevant and methodologically stronger for the specific question being addressed, but still insufficient to support firm or generalisable conclusions. This included some qualitative studies of acceptability and implementation, some technical validation studies for measurement feasibility, and some evidence syntheses or policy reviews where methods were more transparent and independent.
- **Moderate-to-high or high evidential weight** was reserved for evidence that was methodologically robust for the question it addressed, independently conducted, transparent in methods and reporting, and directly relevant to the interpretation of the evidence base.

The evidential weight assigned to a study should therefore be understood in relation to the specific question being considered. For example, a technical validation study may carry moderate weight for showing that a non-contact system can estimate physiological parameters under specified conditions, but very low weight for assessing whether that system reduces self-harm or improves safety in mental health inpatient wards. Similarly, an economic model may be useful for exploring possible cost implications, but it does not provide direct evidence that savings will be realised in

routine practice if its inputs are uncertain or derived from limited observational evidence.

In assessing economic evaluations, particular attention was given to the quality and source of the model inputs, the assumptions used, the transparency of the calculations, the extent of sensitivity or scenario analysis, and whether the modelled costs were supported by robust evidence. The outputs of economic models were treated as conditional estimates rather than direct observations of cost-effectiveness.

Where studies were funded, authored, analysed, or materially supported by the technology supplier, this was treated as relevant to the assessment of the evidence. Such studies were not excluded on that basis, but independence concerns reduced the weight placed on strong claims, particularly where findings had not been independently replicated. Overall conclusions were drawn by considering the evidence base as a whole, rather than relying on any single study. Greater weight was placed on evidence that was directly relevant, independently produced, methodologically transparent, appropriately controlled, and cautious in its interpretation. Strong claims about effectiveness, safety, acceptability, or cost-effectiveness were considered to require correspondingly strong evidence.

Section 1: Economic evaluations

Overview of evidence

1.1 This group of publications examines the potential economic impact of Oxevision within NHS mental health inpatient settings, including acute adult wards, older adult services and Psychiatric Intensive Care Units (PICUs). Across the studies, economic benefits are primarily modelled through reductions in enhanced observation requirements and associated staffing costs. All four publications report estimates suggesting that implementation of the technology may be associated with cost savings or positive returns on investment under the assumptions applied within the models. The findings therefore provide a consistent indication that Vision-Based Patient Monitoring Systems (VBPMSs) could offer opportunities to reduce some elements of inpatient care costs.

1.2 However, the evidential basis for these estimates is limited. The underlying effectiveness data are derived largely from observational and before-and-after studies, often involving small numbers of wards or individual NHS trusts, rather than randomised or controlled evaluations (Malcolm et al., 2022a; Malcolm et al., 2022b; York Health Economics Consortium, 2023; Buckley et al., 2024). Several of the reported savings are highly dependent on reductions in one-to-one observation hours, which in some cases were not statistically significant or were subject to considerable uncertainty. The economic models also require assumptions regarding staffing patterns, ward occupancy, service configuration and the extent to which findings from specific sites can be generalised to other settings. While sensitivity and scenario analyses were undertaken in some studies, uncertainty surrounding the underlying effectiveness estimates remains an important limitation.

1.3 A further consideration is the independence of the evidence. All four publications were either funded by Oxehealth, based on data supplied or analysed by Oxehealth, or produced by organisations commissioned by Oxehealth. In addition, some analyses were undertaken without independent access to the underlying datasets (York Health Economics Consortium, 2023). Taken together, this body of evidence may be regarded as providing low-to-moderate strength evidence regarding the potential economic implications of Oxevision. It identifies plausible mechanisms through which financial savings could be achieved, but does not provide robust independent evidence that equivalent savings would be realised consistently across NHS mental health services in routine practice.

Study summaries

1.4 Malcolm et al. (2022a) (OXHE020518 - LC/109e) present an early health economic evaluation of Oxevision in acute adult and older adult mental health inpatient settings. The model uses a cost-calculator to compare VBPMS plus standard care

with standard care alone over a 12-month time horizon. The paper estimates cost savings of £272 per acute adult patient and £4,591 per older adult patient, but these estimates rely on clinical inputs from small observational studies at a single NHS trust. The acute adult model is particularly fragile: the assumed 20% reduction in one-to-one observation hours was not statistically significant, and when one-to-one observations were excluded in a scenario analysis, the intervention was no longer cost-saving. The older adult model appears more favourable, but it still depends on limited single-trust evidence and assumptions about transferability to other settings. The model is reasonably transparent and includes scenario and sensitivity analyses, but it does not include quality-of-life outcomes or possible privacy-related disbenefits. Independence is also a concern because York Health Economics Consortium (YHEC) was funded by Oxehealth to develop the model and manuscript. Overall, this paper should be given low-to-moderate evidential weight. It is useful as an early indication of possible economic impacts, but it does not provide robust independent evidence that the claimed cost savings will be realised across NHS mental health services.

1.5 Malcolm et al. (2022b) (OXHE020516 - LC/109f) is another early economic evaluation of Oxevision in psychiatric intensive care units using a cost-calculator model. The model compares VBPMS plus standard care with standard care alone and estimates savings of £880 per patient, £72,286 per average ward per year and £5.54 million annually across the NHS. However, the clinical inputs are derived primarily from a single unpublished before-and-after study in one Psychiatric Intensive Care Unit ward, with no randomisation and no contemporaneous control group. The model's main driver is a 7% reduction in one-to-one observation hours, but this effect was not statistically significant and the confidence interval was wide. When the authors restrict the model to outcomes with statistically significant reductions, the estimated saving falls to only £14 per patient and £1,140 per average ward per year. This indicates that the base-case conclusion is highly fragile. The analysis also depends on strong assumptions about scaling from one ward to PICUs across England, and it does not include quality-of-life effects, privacy or dignity concerns, or other patient-centred outcomes. As in Malcolm et al. (2022a) independence is of concern for the same reasons. Overall, this paper should be given low evidential weight, and very low weight for any strong claim that the technology is cost-saving. Its findings are best treated as exploratory and hypothesis-generating rather than robust economic evidence.

1.6 YHEC's NHS Innovation Accelerator economic impact case study (2023) (OXHE024281) presents a modelling analysis of Oxevision in acute mental health wards, psychiatric intensive care units and older adult mental health services. The report estimates that Oxevision may generate positive returns on investment through reductions in night-time observation time and reductions in the use of bank and agency staff for continuous observations. Estimated annual incremental benefits ranged from £37,523 in acute wards to £69,181 in PICUs, with reported returns on investment between 1.41 and 2.64. However, the evidential strength of these findings is limited. The updated analysis presented in

the report was undertaken using data collated and analysed by Oxehealth, and the report explicitly states that YHEC did not have access to the underlying dataset used in the calculations. The results are therefore dependent on provider-generated data and a series of modelling assumptions regarding ward size, staffing patterns, occupancy and transferability between trusts. No confidence intervals or formal uncertainty analyses are reported. Consequently, while the report demonstrates that Oxevision can generate favourable economic estimates within the model used, it does not constitute independent evidence that equivalent savings are realised in routine NHS practice. It should therefore be assigned low evidential weight as economic modelling evidence and low weight as proof of actual cost-effectiveness.

1.7 Buckley et al. (2024) (OXHE020526 - LC/109a) also present a health economic cost-calculator model estimating the impact of implementing Oxevision in acute adult and older adult NHS mental health inpatient settings. The study is quantitative and model-based, drawing on before-and-after data from five NHS mental health trusts, together with national cost sources and cost inputs supplied by the manufacturer. The model is appropriate for exploring potential budget impact, but it is less robust as evidence of causal effectiveness because the underlying clinical data are quasi-experimental rather than randomised. Several important model inputs are based on small numbers of trusts or wards, and the results depend on assumptions about scalability to an average 16-bed ward at 90% occupancy. The paper reports substantial estimated savings, particularly through reductions in one-to-one observation hours, but does not clearly present confidence intervals, statistical significance or extensive uncertainty analysis around the main effect estimates in the article. The conclusions are therefore best understood as showing that VBPMS could be cost-saving under the assumptions used, rather than proving that the reported savings would necessarily occur across NHS mental health services. Chua (2026) (reviewed in “Section 2: Vision-based patient monitoring systems in mental health settings”) accurately suggests that a misinterpretation of relative change may lead to the substantial cost-savings calculated by Buckley et al. (2024). Independence is a material concern because the study was funded by Oxehealth, Oxehealth was involved in data collection and analysis, and the authors’ organisation was funded to develop the model and manuscript. Overall, the paper should be given low-to-moderate evidential weight: it is relevant and useful for identifying possible economic impacts, but its conclusions require cautious interpretation because of confounding, limited statistical uncertainty reporting, restricted data access and conflict-of-interest concerns.

Section 2: Vision-based patient monitoring systems in mental health settings

Overview of evidence

2.1 This body of evidence examines the use of Oxevision and related Vision-Based Patient Monitoring Systems (VBPMSs) across a range of mental health inpatient settings, including acute adult wards, psychiatric intensive care units, dementia services and high-secure forensic care. Early publications primarily described the intended role of the technology, suggesting that contactless monitoring could reduce sleep disturbance, support clinical decision-making and provide a less restrictive alternative to some forms of direct observation (Tully, Larkin and Fahy, 2015; Lloyd-Jukes et al., 2021). This rationale is supported by qualitative evidence indicating that conventional night-time observations may themselves be experienced as intrusive and sleep-disrupting, with patients reporting disturbance from staff entering rooms, use of torches, noise, reduced privacy and negative effects on emotional well-being (Veale et al., 2020). Subsequent service evaluations have reported reductions in a range of adverse outcomes, including self-harm, ligatures, assaults, rapid tranquillisation and falls following implementation of the technology (Ndebele et al., 2022; Wright and Singh, 2022; Ndebele et al., 2024; Kekic et al., 2024; Kekic, Rose, and Bayley, 2026). Across these studies, the direction of effect is broadly consistent, with most reporting reductions in incidents after implementation. However, the evidence remains largely observational and is dominated by uncontrolled or non-randomised before-and-after designs. Consequently, although the findings are suggestive of a possible association between implementation of VBPMSs and improved safety outcomes, they do not establish that the technology itself caused the observed changes.

2.2 A recurring limitation across the effectiveness literature is the potential for confounding and bias. Most studies rely on routine incident-reporting systems and involve small numbers of wards, limited control groups or single-site evaluations, making it difficult to separate the effects of the technology from changes in patient populations, staffing, service delivery, reporting practices or wider organisational initiatives (Ndebele et al., 2022; Wright and Singh, 2022; Ndebele et al., 2024). Concerns have also been raised regarding the interpretation of some reported effect sizes. Chua (2026), for example, argues that the widely cited 44% reduction in self-harm reported by Ndebele et al. (2024) may overstate the magnitude of the observed effect and recommends more robust analytical approaches. Similar methodological concerns apply to the larger multi-site evaluations conducted by Kekic et al. (2024) and Kekic, Rose, and Bayley (2026), where fixed-effect meta-analytic approaches were used despite substantial differences between wards and trusts. Independence is also a significant consideration. Many of the primary effectiveness studies were authored by, funded by, or conducted with substantial involvement from Oxehealth employees,

and one key paper was subsequently retracted because relevant conflicts of interest were not disclosed (Kekic et al., 2024). Taken together, these issues substantially reduce confidence in causal claims regarding effectiveness.

2.3 Alongside the effectiveness literature, a smaller body of qualitative, ethical and policy-focused evidence highlights important questions regarding acceptability, privacy, dignity and governance. Dewa et al. (2023) found that patients and staff identified both perceived safety benefits and concerns relating to surveillance, privacy, anxiety and understanding of the technology. Similar tensions are reflected in wider stakeholder evidence, with national survey findings suggesting that many patients and staff consider VBPMSs useful, while substantial proportions of patients report negative impacts on privacy (Nolan, 2024). Ethical and legal analyses have further highlighted concerns regarding consent, data protection, function creep, staff surveillance and the implications of algorithmic monitoring for therapeutic relationships and human rights (Gooding and Clifford, 2021). Independent reviews of the wider evidence base have reached cautious conclusions. Griffiths et al. (2024) and SQW (2025) both conclude that the available evidence remains limited, methodologically weak and affected by conflicts of interest, with relatively little attention paid to lived experience, patient outcomes or rights-based considerations. Overall, the evidence provides low-to-moderate support for the hypothesis that VBPMSs may contribute to improvements in some safety outcomes, but substantial uncertainty remains regarding the size, causality, generalisability and acceptability of these effects.

Study summaries

2.4 Tully, Larkin and Fahy (2015) (OXHE025074) provide a narrative article on new technologies used to manage risk and violence in forensic psychiatry, including GPS electronic monitoring, CCTV / body-worn video and Oxehealth-related motion sensor technology at Broadmoor Hospital. The article is relevant as historical and forensic context, but it is not an empirical evaluation. The section on Oxehealth describes collaboration with Broadmoor to develop an adapted infrared CCTV system, referred to as Oxecam, intended to provide contactless night-time monitoring without waking or disturbing patients. The article suggests potential benefits, including less restrictive monitoring, improved data for clinical decisions and reduced sleep disturbance, but provides no outcome data, statistical analysis, comparator group or formal patient acceptability findings. The authors themselves emphasise that such technologies require thorough evaluation of cost-effectiveness, patient attitudes, safety and ethical considerations. Overall, this paper should be given low weight as contextual evidence and very low weight for effectiveness.

2.5 Lloyd-Jukes et al. (2021) (OXHE025064) describe the Oxehealth VBPMS and present a conceptual framework for its deployment within mental healthcare pathways. The paper is directly relevant because it discusses both physical monitoring and non-contact physiological monitoring in adult mental health and dementia settings. However, it is not a formal evaluation study. The article

combines a description of the technology, a proposed clinical framework and a number of illustrative case examples. Although selected outcome figures are reported, including reductions in falls and reductions in room-entry observations, no comparator groups, confidence intervals, statistical testing or formal evaluation methods are presented. The paper also raises substantial independence concerns. Several authors are senior employees, shareholders, board members or founders of Oxehealth, and these conflicts are explicitly disclosed. While the paper is useful for understanding how Oxehealth conceptualises the role of vision-based monitoring within mental health services, it provides only limited evidence regarding effectiveness. It should therefore be given low-to-moderate evidential weight for implementation and feasibility, but low weight for claims relating to safety, recovery, cost-effectiveness or clinical benefit.

2.6 Ndebele et al. (2022) (OXHE020513 - LC/109g) report a brief, uncontrolled before-and-after evaluation of Oxevision in a single male PICU. The system was installed in 12 bedrooms and incident data from 31 patients were compared across 12 months before and 12 months after implementation. The authors report a 37% adjusted reduction in all assaults and a 40% reduction in rapid tranquillisation events related to assaults, but the reduction in bedroom assaults, arguably the outcome most directly linked to a bedroom-based monitoring system, was not statistically significant. The study has major methodological limitations: it involved only one ward, had no contemporaneous control group, was not randomised, and did not adjust for case mix, staffing, acuity or changes in incident reporting. The use of patient-level bootstrapping does not overcome these ward-level design limitations. Staff acceptability data were minimal, and there was no meaningful patient perspective or lived experience involvement reported. Independence is also a concern because two authors were employed by Oxehealth and an Oxehealth contributor assisted with data compilation and analysis. Overall, the paper should be given very low to low evidential weight. It is best viewed as preliminary, hypothesis-generating service evaluation evidence rather than robust evidence that the technology reduces assaults or rapid tranquillisation.

2.7 Wright and Singh (2022) (OXHE020512 - LC/109i) report a before-and-after cohort study evaluating a non-contact VBPMS in two dementia inpatient wards at Coventry and Warwickshire Partnership NHS Trust. The system was installed in 12 of 24 bedrooms, and night-time falls were compared across a 12-month baseline period and a 22-month post-installation period. The study reported a 48% reduction in night-time falls, a 68% reduction in A&E admissions and an 82% reduction in annualised moderate-severity falls. These findings are clinically relevant and the use of falls per 1000 bed days is appropriate, particularly as occupancy reportedly remained stable. However, the study has important limitations. There was no contemporaneous control group, no randomisation, and the results may be affected by changes in patient case mix, staffing, falls-prevention practice, reporting or regression to the mean. The use of one-sided p-values and patient-level bootstrapping does not resolve these design limitations.

The reported 71% reduction in enhanced observations is particularly uncertain because no p-value could be calculated and the available data covered only selected 8-month periods before and after installation. The study also does not report meaningful patient, carer or lived experience involvement, despite the privacy and dignity implications of bedroom-based monitoring for patients with dementia. Overall, this paper should be given low-to-moderate evidential weight. It provides a useful early signal of a possible association between the technology and reduced night-time falls, but it does not establish causality or generalisability.

2.8 Dewa et al. (2023) (OXHE020519 - LC/109d) report a qualitative service evaluation of Oxehealth passive remote monitoring technology in Broadmoor Hospital, a high-secure forensic psychiatric hospital. The study used semi-structured interviews with 12 patients and 12 staff members and applied thematic analysis to explore acceptability, perceived benefits, privacy concerns, understanding and practical barriers. The methodology is appropriate for exploring lived experience, staff experience and implementation issues, but it cannot determine whether the technology improves safety outcomes. The study is strengthened by meaningful patient and public involvement: patient advisors with lived experience helped shape study materials, co-produce the topic guide and contribute to early coding and theme development. The findings are balanced and identify both perceived safety benefits and significant concerns, including privacy and dignity, “Big Brother” feelings, paranoia and anxiety, poor understanding of the technology, practical problems and uncertainty about why physical checks remain necessary alongside Oxehealth. The study is limited by its small, selective, all-male patient sample, single high-secure setting, potential social desirability bias and incomplete triangulation at final theme refinement. Overall, it should be given moderate evidential weight for understanding acceptability, ethical concerns and implementation in high-secure forensic care, but low evidential weight for any claim about clinical effectiveness.

2.9 Ndebele et al. (2024) (OXHE020520 - LC/109c) report a mixed-methods, non-randomised controlled before-and-after evaluation of a VBPMS in one female and one male acute inpatient mental health ward, compared with one female and one male control ward. The study used routine Datix incident reports to compare bedroom self-harm and ligature rates before and after implementation. The authors report a 44% relative reduction in bedroom self-harm and a 48% relative reduction in bedroom ligatures compared with control wards. However, the within-intervention-ward reduction in self-harm alone was not statistically significant, and the headline effect depends on comparison with control wards in a very small, non-randomised, single-site study. The use of patient-level bootstrapping is reasonable for skewed self-harm data, but it does not adequately address ward-level clustering, confounding or the limited number of independent intervention units. The qualitative component provides some useful staff and patient perspectives, but it is based on small samples and was not formally analysed. Two authors were employed by, and had financial interests in, the technology supplier, creating a material independence concern, although this was disclosed. Overall, the study should be given low-to-moderate evidential

weight. It suggests a possible reduction in bedroom self-harm and ligature incidents following implementation, but does not provide robust causal evidence of effectiveness.

2.10 Chua (2026) (INQY024285) provides a published methodological critique of the Oxevision self-harm study reported by Ndebele et al. (2024). The commentary focuses on the interpretation of the cited "44% relative percentage change" in self-harm incidents associated with Oxevision implementation. Chua argues that this figure is derived from a ratio of relative rate changes between intervention and comparison wards and may substantially overstate the actual reduction in self-harm incidents observed in practice. By converting the reported rates into estimated incident counts, the author demonstrates that the observed change is closer to approximately one fewer incident per month on a typical ward, corresponding to an estimated reduction of around 22% rather than 44%. The commentary further argues that economic analyses relying on the 44% figure may overestimate potential savings and recommends alternative approaches, including interrupted time series analysis, to evaluate intervention effects more robustly. Although the paper is an editorial rather than an empirical study, it represents one of the most important independent methodological critiques within the evidence base and should be given substantial weight when interpreting claims regarding Oxevision's impact on self-harm outcomes.

2.11 Kekic et al. (2024) (OXHE020521 - LC/109b) report a quasi-experimental evaluation of Oxevision, across acute adult mental health wards in five NHS England mental health trusts. The study used routine incident reports to compare rates of bedroom self-harm before and after implementation, with control wards in four trusts and no control group in one trust. The study reported a pooled relative reduction in bedroom self-harm incidents of 38.9% in a fixed-effect meta-analysis, although no individual ward-level relative-risk estimate was statistically significant and a sensitivity analysis excluding two wards produced a confidence interval crossing zero. The methodology is relevant to the research question but limited by non-randomised design, reliance on routine incident reporting, variable evaluation periods, COVID-19 confounding, outlier exclusions and uncertainty about the appropriateness of fixed-effect pooling across heterogeneous wards and trusts. The article has been retracted because the authors failed to disclose, in the required conflict-of-interest statement, that they were employed by Oxehealth, the developer of the technology under evaluation. This creates a material independence concern. Overall, this is relevant only as an example of the claims advanced in support of the technology, not as robust independent evidence of effectiveness.

2.12 Kekic, Rose, and Bayley (2026) (OXHE024289) extend the retracted article from Kekic et al. (2024) by combining incident-reporting data from 29 inpatient wards across six NHS mental health trusts. The study examined changes in self-harm, falls, assaults, physical restraint and rapid tranquillisation following implementation of the monitoring platform. Pooled analyses suggested reductions across all five outcomes, with reported effect estimates ranging from

21.1% to 38.9%. However, important methodological limitations substantially affect interpretation of these findings. The study uses an observational before-and-after design, and control wards were available only for self-harm analyses. Most outcomes therefore rely on uncontrolled comparisons that are vulnerable to confounding from changes in patient populations, staffing, service delivery or wider organisational factors. In addition, ward-level estimates were pooled using fixed-effect meta-analysis despite substantial heterogeneity between trusts, ward types and patient groups. No formal assessment of heterogeneity was reported, raising concerns that uncertainty may have been underestimated and confidence intervals may appear more precise than warranted. The study is also manufacturer-led, with all authors employed by Oxehealth Limited. Consequently, while the findings suggest an association between Oxevision implementation and reduced safety incidents across multiple sites, they do not establish causality. The paper should therefore be assigned low evidential weight and interpreted as supportive but not definitive evidence of effectiveness.

2.13 Griffiths et al. (2024) (STOX006298 - SO391) provide a systematic review of surveillance-based technologies in inpatient and acute mental health settings. The review is directly relevant because it includes vision-based patient monitoring, CCTV, wearable monitoring, sensors and other surveillance technologies. The methodology is comparatively strong: the review followed PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses), was PROSPERO-registered, searched multiple academic and grey literature sources, used MMAT (Mixed Methods Appraisal Tool) quality appraisal and involved lived experience researchers. However, the review found that the underlying evidence base was weak. Thirty-two studies were included, but only five were rated high quality, with most rated low or medium quality. Conflicts of interest were common, and lived experience involvement in the primary studies was very limited. The authors did not conduct a meta-analysis because the technologies, outcomes and effect sizes were too heterogeneous. The review's conclusion is therefore appropriately cautious: there is limited and generally low-quality evidence that surveillance technologies in inpatient mental health settings achieve intended outcomes such as improved safety or reduced costs. Overall, this paper should carry high weight as an independent map and critique of the evidence base, but it does not itself provide strong evidence of effectiveness for Oxevision or related VBPMSs.

2.14 SQW (2025) (OXHE009976 - LC/60) final report to CQC (Care Quality Commission) is a directly relevant rapid evidence review of vision-based monitoring systems in inpatient mental health units. The report reviewed 68 documents, supplemented by 11 stakeholder interviews and a call for evidence. It is valuable because it examines not only intended benefits, such as improved safety, reduced sleep disturbance, reduced restrictive practice and increased efficiency, but also risks relating to privacy, consent, safeguarding, person-centred care and human rights. The report's overall conclusion is cautious. It describes the evidence base as polarised, with much evidence produced either by organisations with an interest in VBPMSs or by organisations campaigning

against its use. It also notes that lived experience evidence is limited and that much of the evidence focuses on clinical efficiency rather than patient outcomes or experiences. The report identifies limited evidence for reductions in self-harm, with one key estimate showing a relative percentage change of -44% but a confidence interval crossing no effect. Evidence on restrictive practice and economic benefit is also weak or uncertain. Overall, this report should be given moderate-to-high weight as regulatory and policy context and as a synthesis of the VBPMS evidence base, but low weight as proof of clinical effectiveness.

2.15 Nolan (2024) (OXHE020497 - LC/109x) presents an interim slide deck from an ongoing national evaluation of VBPMSs and Body-Worn Cameras in mental health inpatient wards in England. The deck is relevant as grey literature, but it is not a peer-reviewed publication or completed study report. It describes a planned seven-module evaluation including scoping review, uptake survey, governance review, intervention and control site identification, before-and-after incident-rate comparison and stakeholder evaluation. The interim findings are useful context: the scoping review reportedly found nine VBPMS papers, all single-site studies with limited sample sizes, which is consistent with the wider methodological concerns in the evidence base. The national survey reported that 40% of responding providers currently used VBPMSs, with a further four having previously used and stopped. However, the deck provides only descriptive interim findings and does not include a completed incident-rate analysis, full methods, sample denominators, confidence intervals or statistical testing. The results are mixed: while 65% of patients and 81% of staff reportedly considered VBPMSs useful, 61.3% of patients reported a negative impact on privacy. Overall, this material should be treated as low-weight interim grey literature. It is useful for context, particularly regarding national uptake, controversy and privacy concerns, but it should not be relied upon as evidence that a VBPMS improves safety or reduces incidents.

2.16 Gooding and Clifford (2021) (OXHE025065) provide a bioethical analysis of video-algorithmic patient monitoring in healthcare settings, with particular emphasis on psychiatric inpatient care. The paper is highly relevant because it directly addresses ethical, legal and human rights issues associated with technologies such as Oxevision. The authors argue that video-algorithmic monitoring occupies a space between clinical observation and surveillance, particularly where systems are used not only to monitor physiological parameters but also to analyse behaviours associated with self-harm, aggression and suicide risk. The paper highlights concerns regarding privacy, consent, data protection, function creep, staff surveillance, and the impact of monitoring technologies on therapeutic relationships. It also notes that existing legal and regulatory frameworks have largely evolved around CCTV and may not adequately address the challenges posed by algorithmic monitoring systems. Although the paper does not provide empirical evidence regarding effectiveness or safety, it offers one of the most substantial discussions of governance and rights-based concerns within the evidence base. It should therefore be assigned high evidential

weight for ethical and legal analysis, but low weight for questions relating to clinical effectiveness or patient outcomes.

2.17 Heglum et al. (2021) (OXHE020495 - LC/109z) report a measurement-validation study of a contact-free radar sensor for real-time sleep/wake classification. The study compared radar with polysomnography (PSG) and actigraphy in 12 healthy young adults recorded in a psychiatric hospital environment and 28 sleep-clinic patients recorded at home. The methodology is appropriate for validating sleep/wake classification, and the results in healthy volunteers were strong: the real-time ceiling radar model achieved 93.3% accuracy, 96.3% sensitivity, 85.4% specificity and Cohen's kappa of 0.831 against PSG. However, performance was weaker in sleep-clinic patients, where real-time radar specificity was only 53.4% and Cohen's kappa was 0.443, with very wide limits of agreement for some individual sleep-parameter estimates. The study is relevant as comparator evidence because radar offers contact-free monitoring and may raise fewer privacy concerns than camera-based systems, but it is not an Oxevision or VBPMS study. Importantly, the psychiatric-hospital dataset involved healthy volunteers rather than psychiatric inpatients, and the authors acknowledge the need for validation in an in-hospital psychiatric population. Overall, the paper should be given moderate evidential weight for radar-based sleep/wake measurement validation, but very low weight for any claim about Oxevision's clinical effectiveness, safety or acceptability in mental health services.

2.18 Veale et al. (2020) (LC78 - OXHE020533) provide a qualitative study of psychiatric inpatients' experiences of ward environments and intermittent nursing observations at night. Through semi-structured interviews with twelve patients across multiple psychiatric wards, the authors identified consistent concerns regarding sleep disruption, privacy, safety and emotional well-being. Participants reported being repeatedly disturbed by staff entering rooms, shining torches, opening doors and creating noise during routine observations. Many described feeling startled, anxious or unsettled by these practices and reported that poor sleep negatively affected their mental state. Several participants also questioned the therapeutic value of routine night-time observations and expressed concerns regarding privacy and trust. Although the study does not evaluate Oxevision or any vision-based monitoring system, it is highly relevant because it provides direct lived-experience evidence regarding the limitations of conventional observation practices. The paper should therefore be assigned moderate-to-high evidential weight for questions relating to patient experience, acceptability and the impact of observation practices on sleep and well-being, but low evidential weight for claims regarding the effectiveness of VBPMSs.

Section 3: Technical validation and implementation

Overview of evidence

3.1 This body of evidence primarily addresses the technical feasibility, measurement performance and practical implementation of non-contact video-based physiological monitoring rather than its clinical effectiveness. Across a series of technical studies conducted in neonatal intensive care units, haemodialysis services, intensive care units and controlled laboratory settings, researchers consistently demonstrate that camera-based systems can estimate physiological parameters such as heart rate, respiratory rate and, in some cases, sleep-related measures under suitable conditions (Tarassenko et al., 2014; Villarroel et al., 2014; Jorge et al., 2018; Chaichulee et al., 2019; Villarroel et al., 2019; Villarroel et al., 2020; Jorge et al., 2021; Jorge et al., 2022; Carter et al., 2023). Several studies report relatively small measurement errors when compared with reference devices, suggesting that the underlying physiological monitoring approach is technically plausible. The evidence also demonstrates continuing methodological development over time, including improvements in signal extraction, machine-learning techniques, skin segmentation, intervention detection and sleep-stage classification.

3.2 However, the technical validation literature also identifies important limitations. Many studies report performance only during periods when valid physiological signals were available, with data quality affected by patient movement, lighting conditions, clinical interventions or other environmental factors (Villarroel et al., 2014; Villarroel et al., 2020; Jorge et al., 2022). Several studies involve small samples, highly controlled conditions or populations that differ substantially from those encountered in mental health inpatient services, including preterm infants, haemodialysis patients, post-operative intensive care patients and healthy volunteers (Villarroel et al., 2014; Villarroel et al., 2019; Jorge et al., 2021; Carter et al., 2023). Importantly, none of these studies evaluates whether the technology improves patient outcomes, detects deterioration earlier, reduces adverse events, decreases staff workload or represents a cost-effective intervention. Consequently, while the technical evidence provides moderate support for the proposition that camera-based systems can estimate certain physiological parameters under appropriate conditions, it provides limited evidence regarding the benefits of such monitoring in routine clinical practice and very limited evidence regarding mental health inpatient care specifically.

3.3 The implementation literature provides some evidence that Oxevision can be introduced into healthcare and mental health settings and integrated into existing clinical workflows (Barrera et al., 2020; Lewis et al., 2021; Lewis et al., 2024; Clark et al., 2022). These studies report generally favourable findings regarding feasibility, including successful deployment alongside conventional observations, increased availability of physiological observations and positive

feedback from some staff and patients. They also highlight potential operational advantages, particularly the possibility of reducing sleep disturbance associated with conventional night-time observations and increasing access to physiological monitoring in challenging environments such as seclusion rooms (Barrera et al., 2020; Clark et al., 2022). Nevertheless, the implementation evidence is based largely on small-scale service evaluations, quality-improvement projects and observational studies without randomisation or robust comparator groups. Most studies are unable to demonstrate that improvements in monitoring processes translate into better clinical outcomes, and several involve substantial input from Oxehealth employees or company-funded investigators. Taken together, the evidence suggests that the technology is technically feasible and can be implemented in clinical settings, including mental health services, but it does not provide strong evidence that implementation results in improved safety, effectiveness or patient outcomes.

Study summaries

3.4 Tarassenko et al. (2014) (OXHE020501 - LC/109u) present an early technical methods paper on non-contact video-based vital-sign monitoring using ambient light and auto-regressive modelling. The study was conducted in haemodialysis patients who were double-monitored using conventional devices and a video camera. The paper is technically relevant because it addresses real clinical challenges, particularly artificial light flicker, and proposes a pole-cancellation method to remove non-physiological artefacts. It also shows that camera-derived heart-rate estimates can track pulse oximeter values during stable periods with minimal patient movement. However, the study is not a mental health study and does not assess clinical effectiveness, safety outcomes, patient acceptability, staff workload or cost-effectiveness. It is also mainly a method-development paper: full quantitative validation across the dialysis cohort is deferred elsewhere, and the oxygen saturation results are explicitly preliminary and calibration-dependent. The paper should therefore be given moderate evidential weight as early technical feasibility evidence, but very low evidential weight for the Inquiry's questions about Oxevision's effectiveness, safety, acceptability or economic value in mental health services.

3.5 Villarroel et al. (2014) (OXHE020500 - LC/109v) report an early feasibility study of continuous non-contact video-based vital-sign monitoring in preterm infants in a neonatal intensive care unit. The paper presents preliminary results from two infants monitored over eight daytime sessions, totalling 39.8 hours. The study used a specialised modified incubator and compared camera-derived estimates with conventional monitoring. The results are technically favourable: during valid periods, heart-rate estimates had a mean absolute error of 2.83 beats per minute against ECG-derived heart rate, and the authors show an example of a bradycardia with desaturation being tracked from video data. However, the evidential value is limited. Valid camera data were available for only 66.5% of total recording time, accurate heart-rate estimation was possible for only 53.2% of all recording time, and the analysis was retrospective and based on only two infants.

Respiratory-rate validation was complicated by noisy reference data, and oxygen saturation estimation was preliminary. The paper is not a mental health study and does not assess patient outcomes, safety incidents, staff workload, acceptability or cost-effectiveness. Overall, it should be given low-to-moderate evidential weight for early technical feasibility, but very low weight for the Inquiry's questions about Oxevision in mental health services.

3.6 Jorge et al. (2018) (OXHE025081) report a technical conference proceedings paper on camera-based respiratory-rate estimation in neonates. The authors used RGB video recordings from premature infants in a neonatal intensive care unit and developed a method combining CNN-based skin segmentation, visual shape features, blind source separation and autoregressive modelling. The key evaluation used manual respiratory counts from four infants, comprising 165 manual counts. The fused video-derived methods outperformed individual shape features and impedance pneumography in this dataset, with mean absolute error of 6.9 breaths per minute for PCA fusion and 7.5 breaths per minute for ICA fusion, compared with 10.3 breaths per minute for impedance pneumography. The methodology is appropriate for an early technical methods paper, but the evidence is narrow. The sample is very small, no external validation is reported, and the study does not evaluate real-time deployment, clinical decision-making, patient outcomes, staff workflow, privacy, acceptability or mental health settings. Overall, this paper should be given low-to-moderate weight as technical feasibility evidence and very low weight for the Inquiry's questions about Oxevision in inpatient mental health services.

3.7 Chaichulee et al. (2019) (OXHE020503 - LC/109r) report a technical machine-learning study using video recordings from 15 pre-term infants in a neonatal intensive care unit. The authors developed two convolutional neural network models to detect whether a patient was present, segment visible skin regions and identify periods of clinical intervention, so that suitable periods and regions could be used for non-contact vital-sign extraction. The methodology is appropriate for this technical aim and the reported performance is strong: patient detection accuracy was 98.8%, mean skin-segmentation intersection over union was 88.6%, and intervention detection accuracy was 94.5%. The study is strengthened by the use of real clinical video, human annotation and validation across different infant groups. However, it is not a clinical effectiveness study. It does not assess whether non-contact monitoring improves outcomes, detects deterioration earlier, reduces workload or improves safety. Its setting, neonatal incubator monitoring, is also very different from adult psychiatric inpatient care. The study should therefore be given moderate evidential weight for technical feasibility, but very low evidential weight for the Inquiry's questions about the safety, effectiveness, acceptability or economic value of Oxevision in mental health services.

3.8 Villarroel et al. (2019) (OXHE020502 - LC/109s) report a technical validation study of non-contact video-based physiological monitoring in preterm infants in a neonatal intensive care unit. The study recorded 426.6 hours of video and

reference vital-sign data from 90 sessions involving 30 preterm infants. The methodology is appropriate for measurement validation: the algorithms were developed on 15 infants and tested on a separate group of 15 infants, and the authors used Bland-Altman analysis, correlation plots, mean absolute error and reporting of the proportion of time for which estimates were available. The reported performance is reasonably strong, with a mean absolute error of 2.3 beats per minute for heart rate and 3.5 breaths per minute for respiratory rate in the test set during valid periods. However, the study is not an effectiveness study and does not assess clinical outcomes, deterioration detection, staff workload, patient safety or acceptability. Its relevance to mental health inpatient settings is indirect, because the study involved preterm infants in incubators during daytime recording under ambient light, using a specialised camera set-up. The authors also acknowledge the need for larger studies with broader representation by ethnicity, skin colour, gestational age and neonatal complications. Overall, this paper should be given moderate-to-high evidential weight for technical feasibility in the NICU (Neonatal Intensive Care Unit) setting, but very low evidential weight for the Inquiry's questions about Oxevision's safety, effectiveness, acceptability or economic value in mental health services.

3.9 Villarroel et al. (2020) (OXHE020504 - LC/109q) report a technical validation study of camera-based non-contact vital-sign monitoring in patients undergoing haemodialysis. The study recorded 84 dialysis sessions from 40 patients over one year, comprising approximately 304 hours of video. The authors compared camera-derived heart-rate and respiratory-rate estimates with reference physiological devices and reported mean absolute errors of 2.8 beats per minute for heart rate and 2.1 breaths per minute for respiratory rate under valid signal conditions. The methodology is appropriate for technical measurement comparison and is strengthened by the use of multiple reference devices, Bland-Altman analysis and detailed signal-processing methods to address motion and lighting artefacts. However, the study is not an outcomes study and does not assess real-time clinical decision-making, deterioration detection, safety incidents, patient experience or cost-effectiveness. Valid estimates were only available for selected periods, and motion, staff interaction and privacy-related interruptions affected data quality. The cohort was predominantly male and White British, and the haemodialysis setting is not directly transferable to mental health inpatient care. Overall, this paper should be given moderate evidential weight for the narrow proposition that video-based monitoring can estimate heart rate and respiratory rate in a real-world clinical setting under suitable conditions, but very low evidential weight for the Inquiry's questions about Oxevision in mental health services.

3.10 Jorge et al. (2021) (OXHE020505 - LC/109p) report a proof-of-principle physiological study using infrared video thermography to assess peripheral haemodynamic changes in 15 healthy male volunteers. Participants underwent controlled exercise and cold pressor challenges, and the authors measured changes in skin temperature across central and peripheral regions. The methodology is appropriate for a technical physiological question, and the results

show that peripheral temperature changes, particularly in the wrist and hand, can be detected following experimentally induced haemodynamic changes. However, the study is not directly relevant to Oxevision in mental health settings. It did not involve patients, psychiatric wards, clinical monitoring decisions, safety incidents, patient acceptability or outcome assessment. The sample was small, all male and healthy, and measurements were taken under controlled research conditions. Movement also impaired some measurements. Overall, this paper should be given moderate evidential weight only for the narrow proposition that non-contact infrared imaging can detect certain physiological changes under controlled conditions. It should be given very low evidential weight for the Inquiry's questions about the clinical effectiveness, safety, transferability or acceptability of Oxevision in mental health services.

3.11 Jorge et al. (2022) (OXHE020507 - LC/109m) report an observational technical validation study of non-contact camera-based physiological monitoring in post-operative ICU patients. The study compared camera-derived heart rate and respiratory rate estimates with reference values from standard ICU monitoring over 233.5 hours of recording in 15 patients. The methodology is appropriate for a feasibility and measurement-accuracy study. During periods when valid estimates were available, performance was reasonably strong, with mean absolute errors of 2.5 beats per minute for heart rate and 2.4 breaths per minute for respiratory rate. However, estimates were available for only 53.2% of valid time for heart rate and 63.1% for respiratory rate, and all analysis was conducted offline rather than being used in real-time clinical care. The cohort was small, from a single ICU, and lacked skin-tone diversity, with no participants classified as Fitzpatrick type IV or above. The study therefore provides moderate evidence that non-contact camera-based monitoring can estimate heart rate and respiratory rate in a post-operative ICU setting under valid signal conditions. It does not provide evidence that such systems improve clinical outcomes, reduce deterioration, reduce staff workload, or are transferable to mental health inpatient settings. Overall, the paper should be given moderate evidential weight for technical feasibility and measurement accuracy, but low weight for clinical effectiveness and very low weight for mental health-specific conclusions.

3.12 Carter et al. (2023) (OXHE020499 - LC/109w) report a machine-learning study using near-infrared video-derived heart rate, breathing rate and activity measures to classify sleep stages in 50 healthy adult volunteers. Sleep-stage labels were derived from expert-scored video polysomnography, and the authors used a transfer-learning approach in which a model trained on a larger contact-sensor sleep dataset was used to extract features from video-derived vital signs. The study achieved four-class sleep-stage accuracy of 73.4% and Cohen's kappa of 0.61, which the authors report as an improvement over previous video-based sleep-staging methods. The methodology is appropriate for a technical sleep-staging study, and the ablation analyses suggest that both video-derived motion and deep vital-sign features contribute to performance. However, the study is not a clinical effectiveness study and was conducted in healthy volunteers in experiment rooms, not in mental health inpatient settings. It does not assess

safety outcomes, clinical decision-making, acceptability, privacy, staff workload or cost-effectiveness. The work is also commercially connected: two authors are affiliated with Oxehealth, the work was funded by Oxehealth, and components of Oxevision Vital Signs software were used. Overall, the paper should be given moderate evidential weight for technical feasibility of video-based sleep staging, but very low evidential weight for the Inquiry's questions about Oxevision's safety, effectiveness, acceptability or economic value in mental health services.

3.13 Lewis et al. (2021) (OXHE025078) report a conference abstract describing a pilot observational study of Oxevision in respiratory medicine and orthopaedic surgery wards. Forty-one patients were monitored over an eighteen-month period, with Oxevision used to collect non-contact physiological observations and activity logs. The authors suggest that continuous monitoring identified physiological trends not evident from intermittent manual observations and that activity logs demonstrated reductions in time spent in bed during recovery. However, the abstract provides very limited methodological information, contains no statistical analysis and does not include a comparator group. As a result, it is not possible to determine whether any reported observations were attributable to the technology or reflected routine clinical recovery. The study was funded by Oxehealth and included an Oxehealth-affiliated author. Overall, the abstract should be assigned very low evidential weight and should be regarded primarily as an early feasibility and implementation report rather than evidence of effectiveness, safety or patient benefit.

3.14 Lewis et al. (2024) (OXHE020508 - LC/109k) report a short measurement-comparison study of remote vision-based monitoring of pulse and respiratory rate in six single rooms across acute respiratory and orthopaedic wards. The study compared 102 paired pulse-rate measurements and 154 paired respiratory-rate measurements from 27 patients against routine nurse observations, with a small additional validation of respiratory rate using independent review of video clips. The methodology is appropriate for preliminary measurement comparison, and the use of Bland-Altman analysis is suitable. Pulse-rate measurements were broadly similar, although nurse pulse was on average 3.91 beats per minute higher than VBPMS. The most notable finding was that nurse respiratory-rate recordings clustered around 16 to 18 breaths per minute and did not correspond well with VBPMS, while independent video review supported VBPMS respiratory-rate measurements in the selected validation sample. The study is useful evidence of feasibility and measurement plausibility, but it does not assess clinical outcomes, deterioration detection, escalation, alarm burden or workflow effects. The VBPMS also does not provide oximetry or blood pressure, limiting its role as a standalone monitoring system. Independence is a material concern because the study was funded by Oxehealth and several authors were current or former Oxehealth employees or consultants. Overall, the paper should be given moderate evidential weight for measurement feasibility, but low evidential weight for clinical effectiveness or patient safety outcomes.

3.15 Barrera et al. (2020) (OXHE020510 - LC/109j) describe the staged introduction of Oxehealth sensors for digitally assisted night-time nursing observations on an acute male psychiatric ward. The paper is best understood as an implementation and feasibility report rather than an effectiveness study. The project involved governance review, patient and public involvement, staff engagement, blind running, parallel testing against conventional observations and staged implementation. The initial findings were favourable: in 275 parallel observations over 22 patient nights, sensor-assisted observations matched conventional observations, and no sensor-related incidents were identified over an estimated 755 patient nights. However, the study was not randomised, involved only six sensor rooms on one ward, and the rooms were used for more acutely unwell patients. The exploratory findings on insomnia, length of stay and medication use are therefore highly confounded and should not be treated as evidence of clinical effectiveness. The paper includes useful patient, relative and staff feedback suggesting that conventional night observations can disturb sleep and privacy, and that digitally assisted observations may reduce disturbance. However, the qualitative component was not a formal qualitative analysis. Independence should also be treated cautiously because two authors were from Oxehealth and the project involved a commercial agreement, although the paper declared no competing interests. Overall, the paper should be given moderate evidential weight for feasibility and implementation learning, low weight for safety, and very low weight for effectiveness outcomes.

3.16 Clark et al. (2022) (OXHE020511 - LC/109h) report a proof-of-concept quality improvement project using Oxevision for non-contact physical health monitoring in a women's PICU seclusion room. The study found that the system increased the availability of pulse and breathing-rate observations compared with a historical baseline, with the authors reporting a 12.3-fold increase overall when the intervention was used alongside standard care. The methodology is appropriate for a local process-improvement project, but it is not a robust clinical effectiveness study. There was no randomisation, no contemporaneous control group and no demonstrated reduction in seclusion duration, seclusion frequency or adverse clinical outcomes. The headline increase in observations should also be interpreted carefully because non-contact vital-sign measurements are not necessarily equivalent to full physical health assessments. The system's performance varied substantially according to patient movement, with fewer readings available during high-movement periods, which may be precisely when clinical monitoring is most challenging. The paper includes some patient, carer and staff feedback, and early patient concerns shaped how the technology was explained, but this is best characterised as consultation rather than full lived experience involvement. Independence is a concern because the project was jointly supported by Oxehealth and included Oxehealth employees as authors. Overall, the paper should be given moderate evidential weight for feasibility and process improvement, but low evidential weight for clinical effectiveness or safety outcomes.

Section 4: Digital Health

Overview of evidence

4.1 The publications included in this section do not provide direct evidence regarding the effectiveness, safety, acceptability or cost-effectiveness of Oxevision or related Vision-Based Patient Monitoring Systems (VBPMSs) in mental health inpatient settings. Instead, they provide broader clinical, policy, technological, ethical and legal context relevant to the development and implementation of digital monitoring technologies in healthcare. Several publications discuss the clinical rationale for continuous or remote monitoring, particularly the limitations of intermittent observations and the potential value of longitudinal physiological data for detecting deterioration (Michard and Kalkman, 2021). Others consider the wider role of digital technologies, passive monitoring systems and data-driven approaches within mental health services and healthcare more generally (Griffin and Saunders, 2020; Morrison, 2020; WHO, 2018). Collectively, these publications help situate VBPMSs within wider trends towards digital health, artificial intelligence and remote monitoring, but they do not themselves establish that such technologies improve outcomes.

4.2 A consistent theme across the literature is that implementation challenges may be as important as the technical capabilities of the technology itself. Studies and commentaries highlight issues including trust, usability, workflow integration, staff acceptance, organisational readiness and the practical challenges of introducing novel monitoring technologies into routine care (Ede et al., 2021; Wenzel and Evans, 2019). Similar concerns are reflected in broader discussions of digital psychiatry and video-based monitoring, where questions of patient acceptability, data quality, interpretability and unintended consequences are recurrent themes (Griffin and Saunders, 2020; Climent-Pérez et al., 2020). Several authors also note that successful implementation depends on wider governance arrangements, infrastructure, workforce capability and service design rather than the monitoring technology alone (WHO, 2018; Wenzel and Evans, 2019).

4.3 Privacy, ethics, governance and accountability emerge as particularly prominent considerations. Multiple publications highlight concerns regarding surveillance, confidentiality, data security, algorithmic bias, transparency and the possibility that monitoring technologies may affect patient experience or therapeutic relationships (Griffin and Saunders, 2020; Climent-Pérez et al., 2020; Diao et al., 2022). The legal literature similarly emphasises uncertainty regarding responsibility and liability where artificial intelligence systems contribute to clinical decision-making or harm (Price, Gerke and Cohen, 2023). Across the literature there is broad agreement that digital health technologies require robust evaluation, appropriate regulation, clear governance arrangements and careful consideration of ethical implications before widespread adoption (WHO, 2018; Diao et al., 2022; Price, Gerke and Cohen, 2023).

4.4 Overall, this body of literature should be interpreted primarily as contextual evidence rather than effectiveness evidence. Its principal contribution is to identify broader opportunities, challenges and considerations associated with digital monitoring technologies in healthcare. While it provides useful background for understanding the development and implementation of Oxevision and related systems, it offers little direct evidence regarding whether VBPMSs improve safety, patient outcomes, experience of care or value for money within mental health inpatient services.

Study summaries

4.5 Michard and Kalkman (2021) (OXHE024286) provide a narrative clinical review of continuous vital-sign monitoring on hospital wards. The article is useful as general clinical context because it explains why intermittent vital-sign spot checks may miss deterioration, particularly respiratory depression, hypoxaemia, hypotension and other progressive deterioration. It also emphasises that respiratory rate is an important but often poorly measured marker of deterioration. However, the article is not a mental health study and does not evaluate Oxevision in mental health inpatient settings. Oxevision is mentioned only briefly as an example of optical sensor technology, and the authors state that they were not aware of any published clinical evaluation of such video monitoring systems on medicosurgical wards. The article also highlights implementation concerns that are relevant to VBPMS, including sensor accuracy, connectivity, alarm fatigue, nurse workload and the need for a coordinated response to deterioration. Overall, this paper should be given moderate weight as background on continuous ward monitoring, but very low weight for Oxevision-specific effectiveness, safety, acceptability or cost-effectiveness in mental health services.

4.6 Ede et al. (2021) (OXHE025082) report a qualitative study exploring intensive care staff perceptions of a prototype non-contact vital-sign monitoring system using infrared and thermal imaging. The study addresses implementation and human factors issues that are largely neglected in the wider literature. Through semi-structured interviews with nine intensive care nurses, the authors identified themes relating to trust, usability, troubleshooting, expectations of continuous monitoring and organisational adoption. Staff perceived potential benefits including reduced patient restriction, improved mobilisation and a less intrusive monitoring experience. However, participants also expressed concerns regarding technology reliability, loss of control over monitoring processes and the need to establish trust in new systems. Importantly, the study does not evaluate clinical effectiveness, patient outcomes or monitoring accuracy. Patients and relatives were not interviewed directly, meaning that views about patient experience are inferred from staff perceptions rather than lived experience evidence. Independence is a concern with one author employed by Oxehealth, and another author being a founder and shareholder of Oxehealth. The paper also states that relevant patents and commercial interests exist. These conflicts are transparently

disclosed. Overall, the paper should be assigned moderate evidential weight for understanding implementation and adoption issues, but low weight for claims regarding effectiveness, safety or clinical benefit.

4.7 Griffin and Saunders (2020) (OXHE020522 - LC/109bb) provide a peer-reviewed perspective article on the use of smartphones, wearables and passive digital monitoring to understand symptom mechanisms in psychiatry. The article is useful general background because it explains how ecological momentary assessment and passive data streams may capture longitudinal, naturalistic information about mood, behaviour, activity and cognition. It also identifies important ethical and practical concerns, including data security, patient confidentiality, missing data, interpretability, acceptability, potential preoccupation or paranoia, and the possibility that remote monitoring may increase anxiety or distress. However, it is not an empirical study and does not evaluate Oxevision or related VBPMS in mental health inpatient wards. Oxehealth is mentioned only briefly, with reliance on an Oxehealth report rather than independent peer-reviewed evidence. Overall, the article should be given very low evidential weight for Oxevision-specific questions, but may be cited as general context on the opportunities and risks of digital monitoring in psychiatry.

4.8 Morrison (2020) (OXHE024282) provides a policy and service review of digital mental health solutions operating across England and Scotland. The report maps the use of digital therapies, self-management applications, telehealth services, online support platforms and NHS digital mental health initiatives. It is relevant as contextual evidence because it documents the growing policy interest in digital mental health services and describes examples of implementation across NHS organisations. The report also notes that Oxford Health NHS Foundation Trust piloted the Oxehealth Digital Care Assistant for overnight monitoring of adult mental health inpatients. However, no outcome data or evaluation findings are presented. The report is not an empirical study, contains no original data and does not assess effectiveness, safety, cost-effectiveness or patient acceptability. Consequently, it should be assigned low evidential weight for the Inquiry's core questions, while retaining moderate value as background context regarding the adoption of digital mental health technologies within NHS services.

4.9 Climent-Pérez et al. (2020) (OXHE025079) provide a narrative review of video-based technologies used in ambient assisted living, including activity recognition, physiological monitoring, lifelogging, computer vision and deep learning systems. The review is not focused on mental health services and does not evaluate Oxevision or related vision-based monitoring in psychiatric inpatient settings. Oxehealth's OxeCam is mentioned only briefly as an example of video-based physiological monitoring technology. The review's principal relevance lies in its discussion of privacy, ethical concerns, surveillance, user acceptance and the challenges associated with deploying video-based monitoring technologies. It contains no original data, no clinical outcome analysis, no cost-effectiveness evaluation and no systematic assessment of evidence quality. Accordingly, it should be assigned very low evidential weight for the Inquiry's questions regarding

the effectiveness, safety, acceptability or economic value of Oxevision in mental health services.

4.10 Diao et al. (2022) (OXHE020509 - LC/109l) is an editorial commentary on video-based physiological monitoring, rather than an empirical study. It does not evaluate Oxevision in mental health settings and provides no original clinical, statistical or economic analysis. Its relevance is therefore limited to general context. The editorial is useful in identifying potential advantages of non-contact video-based monitoring, such as reduced disruption from contact sensors and richer longitudinal physiological data, but it also emphasises that such technologies require rigorous clinical validation and may be affected by failure modes, artefacts, signal loss, false alarms, privacy concerns, possible misuse and demographic bias, including concerns about accuracy in patients with darker skin. Overall, this publication should be given very low evidential weight for the Inquiry's substantive questions. It may be cited only as contextual opinion evidence about the broader technology field, not as evidence that Oxevision is effective, safe or cost-saving in mental health settings.

4.11 Price, Gerke and Cohen (2023) (OXHE020498 - LC/109y) provide a draft legal chapter on liability for medical AI. The chapter is not empirical research and does not evaluate Oxevision or related vision-based monitoring in mental health settings. Its relevance is therefore contextual rather than evidential. The authors identify three potential loci of liability: individual clinicians, healthcare institutions and AI developers. They emphasise that liability for medical AI remains unsettled, that causation may be difficult to establish, and that institutional decisions about selecting, testing, implementing and integrating AI systems into clinical workflow may carry legal significance. The chapter is useful for framing accountability questions around AI deployment, but it is written from a United States legal perspective. It should therefore be given low evidential weight in the clinical review, but may be cited as background legal context on the complexity of responsibility for AI-related harm.

4.12 The WHO technical brief *Digital technologies: shaping the future of primary health care* (2018) (OXHE024283) provides broad policy context on the role of digital technologies in strengthening primary healthcare. It is not an empirical study and does not evaluate Oxevision, VBPMS or inpatient mental health monitoring. Its focus is primary care, public health, electronic records, telemedicine, remote care, mobile health, digital health literacy and governance. The only apparent connection to Oxehealth is a brief reference to electronic sensors measuring vital signs and activity to assist staff monitoring patients at risk of falls, supported by an industry news item about Oxehealth accreditation rather than independent clinical evidence. The document is therefore of very low evidential value for the Inquiry's Oxevision questions. Its main relevance is contextual: it emphasises that digital health implementation requires evidence, health technology assessment, regulation, privacy and security protections, workforce capability and governance.

4.13 Wenzel and Evans (2019) (OXHE020514 - LC/109dd) provide a King's Fund policy research report on how technology may affect the NHS estate. The report is not an Oxevision or mental health inpatient study, and it does not evaluate clinical outcomes, safety incidents or cost-effectiveness. Its relevance is contextual. It cautions that technology is likely to result in a different NHS estate rather than simply a smaller one, and that assumptions about space reduction or efficiency gains need careful scrutiny. The report also emphasises that technology and estate planning have often operated in silos and should instead be integrated into wider service-change plans. This is relevant to VBPMS because any claimed benefits depend not only on the device, but on ward layout, digital infrastructure, staff workflow, governance, privacy safeguards and patient engagement. The report's strongest contribution is its emphasis on broad patient and staff engagement, including consideration of whether digital changes may disadvantage particular groups. Overall, it should be given low evidential weight for Oxevision-specific effectiveness questions, but moderate weight as implementation and planning context.

4.14 Kamel (2020) (OXHE020517 - LC/109cc) provides a narrative policy perspective on the future of health services after COVID-19 from an Egyptian viewpoint. The article discusses broad themes including health-system reform, telemedicine, digital healthcare technologies, sensors, artificial intelligence, workforce training and biomedical industry development. It is not an empirical study, does not use systematic review methods, and does not evaluate Oxevision or related vision-based monitoring in mental health settings. The brief statement that video cameras can be used in elderly home care and psychiatric hospitals to monitor high-risk situations is unsupported by independent clinical evaluation within the article and partly relies on an Oxehealth company webpage. The paper should therefore be given very low evidential weight for the Inquiry's questions. It may be used only as general background on digital health after COVID-19, not as evidence of the safety, effectiveness, acceptability or cost-effectiveness of Oxevision or related VBPMSs in mental health services.

4.15 Bremner et al. (2020) (OXHE020524 - LC/109aa) provide a broad narrative article on immersive health technologies, including virtual reality, augmented reality, machine learning and artificial intelligence. The article is not an empirical study and does not use systematic review methods. Its relevance to Oxevision is limited to a brief paragraph stating that Oxehealth had created a European medical device for contact-free pulse and respiratory-rate spot observations, and that no published academic papers were available at the time. The article then refers to an audited case study of Oxehealth use on an acute psychiatric ward, but that source appears to be Oxehealth's own material rather than independent peer-reviewed evidence. Accordingly, this paper should be given very low evidential weight for the Inquiry's questions. It is useful only as general background commentary on digital health innovation, not as evidence of the effectiveness, safety, acceptability or cost-effectiveness of Oxevision in mental health services.

4.16 Su et al. (2022) (OXHE024287) provide a narrative review examining the potential role of sixth-generation communications technologies and artificial intelligence in dementia care. The review discusses a wide range of emerging technologies, including social robots, AI-assisted communication systems, avatar robots, computer vision technologies and sensor-based monitoring systems. Oxehealth is mentioned only briefly as an example of video-based vital-sign monitoring technology within a broader discussion of future assistive technologies. The paper does not evaluate Oxevision, mental health inpatient monitoring or clinical outcomes. No original data are presented, and the analysis is largely speculative, focusing on possible future applications rather than demonstrated effectiveness. Consequently, the paper should be assigned very low evidential weight for the Inquiry's questions. Its relevance is limited to broad contextual discussion of future technology trends in dementia care.

Section 5: Unrelated

5.1 Jackson et al. (2018) (OXHE020515 - LC/109ee) analyse Twitter engagement with the hashtags #CDCGrandRounds and #VitalSigns. Although peer reviewed, the article is not relevant to the evaluation of Oxevision or vision-based monitoring. The term “Vital Signs” in this paper refers to a CDC monthly public health publication, not physiological vital signs. The study uses tweet metadata and negative binomial regression to examine whether URL links, hashtags, mentions and visual cues are associated with retweet frequency. It contains no clinical data, no patient data, no ward setting, no medical device evaluation and no assessment of safety, effectiveness, acceptability or cost-effectiveness. It should therefore be excluded from the substantive evidence review, or recorded as not relevant.

STATEMENT OF TRUTH

I confirm that I have made clear which facts and matters referred to in this report are within my own knowledge and which are not. Those that are within my own knowledge I confirm to be true. The opinions I have expressed represent my true and complete professional opinions on the matters to which they refer. I understand that proceedings for contempt of court may be brought against anyone who makes, or causes to be made, a false statement in a document verified by a statement of truth without an honest belief in its truth.

Signed:

[I/S]

Name: Professor Christl Donnelly

Date: 24 June 2026

Signed:

[I/S]

Name: Dr Maria Christodoulou

Date: 24/06/2026

Glossary

Acceptability: Acceptability is the extent to which an intervention is considered appropriate, tolerable, or satisfactory by the people affected by it.

Before-and-after study: A before-and-after study compares outcomes before and after an intervention is introduced. Such studies are sometimes called “pre-post studies”. This design can show that outcomes changed over time, but it is usually limited for causal interpretation because other changes may have occurred at the same time.

Bias[3]: Bias is the systematic (as opposed to random) deviation of the results of a study from the 'true' results, which is caused by the way the study is designed or carried out. (See also the definition of Publication bias.)

Bootstrapping[4]: Bootstrapping is a way of generating confidence intervals and the distribution of test statistics through resampling the observed data rather than through assuming a probability model for the underlying random variable. A basic bootstrap of a data set $x_1, x_2 \ldots x_n$ (where each x represents the data from a sampling unit) is a sample of size n with replacement, so that the bootstrap sample will be drawn from the original set of distinct values but not generally in the same proportions as the original data set.

Clustering: Clustering arises in situations in which observations are grouped together and may be more similar to each other than to observations in other groups. For example, patients on the same ward may be affected by the same staffing, environment, or clinical practices. Statistical analysis should take clustering into account where relevant.

Confidence interval[5]: A confidence interval gives an indication of the degree of uncertainty of an estimate and helps to describe how precise a sample estimate is. It specifies a range of values likely to contain the unknown population value. These values are defined by lower and upper limits. The width of the interval depends on the precision of the estimate and the confidence level used. Traditionally 95% confidence intervals are used. A greater standard error will result in a wider interval; the wider the interval, the less precise the estimate is.

---

[3] Definitions were sourced from the National Institute for Health and Care Excellence (NICE) glossary (https://www.nice.org.uk/glossary), accessed 18 June 2026. In some cases, these have been adapted for the requirements of this report.

[4] Definition was adapted from D Spiegelhalter, *The Art of Statistics Learning from Data*, Pelican Books, Pelican, 2020.

[5] Definition was adapted from the *Office for National Statistics Uncertainty and how we measure it for our surveys* (https://www.ons.gov.uk/methodology/methodologytopicsandstatisticalconcepts/uncertaintyandhowwemeasureit), accessed 18 June 2026.

Confounding[3]: Confounding occurs when a variable (the confounding variable) is associated with the exposure (treatment) variable and independently affects the outcome. This might result in an association being observed between the exposure and the outcome when there is none or in a true association being masked. For example, a study of heart disease may look at a group of people who exercise regularly and a group who do not exercise. If the ages of the people in the two groups are different, any difference in heart disease rates between the two groups could be because of age rather than exercise. So, age is a confounding factor.

Control group: A control group is the group of study units (whether those are people, households, or hospital wards) that does not receive the experimental treatment (or intervention). The control group should be otherwise comparable to the treatment (or intervention) group to reduce the risks of confounding.

Comparator group: This is another term for a control group (see its definition above).

Cost-calculator model: A cost-calculator model is a tool for calculating the costs of a particular set of conditions. The costs might include staff costs, administrative overheads, the costs of materials, and finance-related costs such as interest on loans if required.

Effectiveness: Effectiveness is the expected impact of an intervention under real-world conditions.

Efficacy: Efficacy is the expected impact of an intervention under ideal and controlled conditions.

Empirical study: An empirical study is based on evidence, which might be observational or experimental.

Evidential weight: A judgement about the degree of confidence that can reasonably be placed in a study or source, taking account of factors such as study design, methodological quality, risk of bias, transferability and independence.

Experimental study[3]: An experimental study meets two criteria: manipulation of a variable factor between two or more groups, and random assignment of study units (for example, patients) to those groups. A randomised controlled clinical trial is an example of an experimental study.

Grey literature: Grey literature refers to information published outside of the traditional academic publishing system. This includes government reports, newsletters, speeches and urban plans.

Heterogeneity: Heterogeneity is the amount of variation, for example between people, within the same person over time or between hospital wards. There is a spectrum from things being uniform (homogeneity) to things being extremely different from each other (heterogeneity).

Heterogeneous: See the definition of Heterogeneity: a group of things that are substantially different from each other might be described as heterogeneous.

Homogeneity: Homogeneity means uniformity; that is that the study units (whether those are people, households, or hospital wards) are the same or essentially so. There is a spectrum from things being uniform (homogeneity) to things being extremely different from each other (heterogeneity).

Homogeneous: See the definition of Homogeneity: a group of things that are very similar, even if not identical, might be described as homogeneous.

Longitudinal: Longitudinal data tracks the same study unit (whether that is a person, a household, or a hospital ward) at different points in time. Thus, there are repeated observations on the same sampling units. Similarly, a longitudinal study collects longitudinal data.

Meta-analysis: Meta analysis is the statistical analysis of a set of individual studies undertaken to describe the consistency or otherwise of their findings. It extends quantitatively beyond a literature review which describes a set of studies and synthesises the results. A **fixed-effect meta-analysis** assumes that there is one true effect (the fixed effect) underlying all of the studies considered. In contrast, a **random-effects meta-analysis** looks to describe variation (heterogeneity) in the effects between the studies considered.

Methodology: Methodology is the scientific framework for undertaking a study to answer a scientific question. It includes the study design, the population to be studied, the data collection method (for example, the diagnostic test or questionnaire), the statistical analysis, and protections of study participants including ethical review, confidentiality, and informed consent.

Mixed methods research: A mixed methods approach to research employs both quantitative and qualitative methods to estimate the magnitude of effects and to characterise the meaning and importance of these effects. As a result, multiple types of data are analysed (for example, oral testimonies, quantitative measures, and questionnaire data). The goal is to integrate the understanding of the available data so that the research utilises the strength of each data type and counterbalances the weaknesses.

MMAT (Mixed Methods Appraisal Tool): The MMAT is used to assess the methodological strengths and weaknesses of qualitative, quantitative and mixed methods studies. Criteria are tailored to study designs with quantitative studies being further subdivided into three subtypes: randomised controlled, non-randomised, and quantitative descriptive.

Model: A statistical model uses mathematics and parameters to describe the relationship between different variables. For example, the Robinson formula[6] is an early, very simple model for the ideal body weight (IBW) for men as a function of height (H) in inches:

IBW = 51.65 kg + 1.85 kg/inch (H-60 inches).

In this case the variables are ideal body weight and height, and the parameter values are 51.65kg (the intercept) and 1.85kg/inch (the slope). Models come in many different forms.

Narrative review: Narrative reviews are a form of knowledge synthesis. They are sometimes criticised for being non-systematic, but can provide interpretation, contextualisation, and critique of evidence.

Observational study[3]: An observational study is a retrospective or prospective study in which the investigator observes the natural course of events with or without control groups (for example, cohort studies and case–control studies).

P-value[3]: The p-value is a statistical measure that indicates whether or not an effect is statistically significant. For example, if a study comparing two treatments found that one seems to be more effective than the other, the p-value is the probability of obtaining these results by chance. By convention, if the p-value is less than 0.05 (that is, there is less than a 5% probability that the results occurred by chance), it is considered that there probably is a real difference between treatments. If the p-value is 0.001 or less (less than a 0.1% probability that the results occurred by chance), the result is seen as highly significant. However, a statistically significant difference is not necessarily clinically significant. For example, drug A might relieve pain and stiffness statistically significantly faster than drug B. But, if the difference in average time taken is only a few minutes, it may not be clinically significant. The p-value shows whether there is likely to be a true difference between treatments, whereas the estimate and the confidence interval describe how big the effect might be.

Peer review[3]: Peer review of a study, service or recommendation is an assessment undertaken by those with similar interests and expertise to the people who produced it to assess whether the study results are accurate, valid and clearly reported.

PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses)[7]: PRISMA is a guideline designed to improve the reporting of systematic reviews. It provides guidance to authors on how to report the purpose, methods and results of systematic reviews and meta-analyses.

---

[6] Robinson JD et al. Determination of Ideal Body Weight for Drug Dosage Calculations. *American Journal of Hospital Pharmacy* 40: 1016-1019, 1983.

[7] Definition was sourced from PRISMA (https://www.prisma-statement.org/), accessed 18 June 2026.

PROSPERO[8]: PROSPERO is an international registry of systematic reviews. The goal is to promote transparency, thereby reducing publication biases and unintended research duplication.

Publication bias: Publication bias can arise when studies showing evidence of an effect, for example a clear benefit, of an intervention are more likely to be accepted for publication in scientific journals than studies that do not show that effect.

Quantitative research[3]: Quantitative research generates numerical data or data that can be converted into numbers. An example is research using clinical trials. Another example is the Census of England and Wales, which counts people and households. It might involve questions such as "How many people visit their GP each year?" or "What proportion of children have had this vaccine?".

Qualitative research[3]: Qualitative research explores people's beliefs, experiences, attitudes, behaviour, and interactions. It asks questions about how and why. For example, why people want to stop smoking, rather than asking how many people have tried to stop. It generates non-numerical data, such as a person's description of their pain rather than a measure of pain.

Quasi-experimental study[3]: An experimental study meets two criteria: manipulation of a variable factor between two or more groups, and random assignment of study units (for example, patients) to those groups. A quasi-experimental study uses the first criterion, but patients are not randomly assigned to groups. This means a researcher cannot draw conclusions about "cause and effect". This design is frequently used when it is not feasible, or not ethical, to conduct a randomised controlled trial.

Randomisation[3]: Randomisation is the assignment of study units (whether those are people, households, or hospital wards) in a research study to different groups without taking any similarities or differences between them into account. It could involve using a random numbers table or a computer-generated random sequence. As a result, each individual (or each group in the case of cluster randomisation) has the same chance of having each intervention.

Randomised: See the definition of Randomisation.

Relative risk[3]: The probability of an event occurring in the treatment group compared with the probability of the same event occurring in the control group as a ratio is known as relative risk. If both groups face the same level of risk, the relative risk is 1. If the treatment group had a relative risk of 2, subjects in that group would be twice as likely to have the event happen. A relative risk of less than 1 means the outcome is less likely in the treatment group.

---

[8] Definition was sourced from PROSPERO (https://www.crd.york.ac.uk/prospero/), accessed 18 June 2026.

Scenario analysis: A scenario analysis is a method used to evaluate possible outcomes by comparing different scenarios (perhaps baseline, best-case and worst-case).

Sensitivity analysis: A sensitivity analysis examines how consistent the results of an analysis are if the assumptions, input data, and/or model structure are varied.

Service evaluation: A service evaluation is an assessment of a service to evaluate how it performs. For example, healthcare performance might be considered from the point of view of patients, the point of view of staff, and/or in terms of clinical outcomes.

Significance: See the definition of Statistical significance.

Skewed: In statistics, skewness refers to the degree of asymmetry in a distribution of data. A symmetric distribution (such as the normal distribution, which has a bell-shaped curve) has zero skew. When a distribution is more concentrated to one side, it is described as skewed, meaning one tail is longer or more extended than the other.

Statistical significance: A statistically significant result is one in which the observed effect is unlikely to have occurred by random chance. The p-value is a statistical measure used to assess this, indicating the probability of obtaining the observed result (or a more extreme result) if there were no true effect. The conventional threshold for statistical significance is a p-value of less than 0.05. However, statistical significance does not prove that an effect is real or meaningful. Additionally, studies with small sample sizes may fail to detect real effects, as they may lack sufficient statistical power.

Standard error: A standard error is a measure of the precision of a parameter estimate, such as the sample mean (average). The standard error depends on both the variation observed within the study data and the sample size. As the sample size increases, the standard error becomes smaller.

Standard deviation[3]: The standard deviation is a measure of the heterogeneity or variability of a set of measurements. It is often used with the mean (average) of the measurements to describe numerical data.

Transferability: Transferability is the extent to which study results will apply to other populations, settings, and conditions. If, for example, some groups of patients were excluded from a study, then a key question is whether the results of the study will be transferable to them.

Theme refinement: In qualitative data analysis, information is first considered to seek an initial understanding of the topic. Initial themes are identified and information is coded according to these themes. As similar codes are grouped together and codes for larger topics are split, themes are refined iteratively. The reliability and validity of the themes are challenged and confirmed by comparison with further data with the goal of obtaining robust findings.

Triangulation[3]**:** Triangulation involves the use of multiple different research methods in combination, principally used as a check of validity. The more the different methods produce similar results, the more valid the findings are.

Uncertainty analysis: Confidence intervals are a form of statistical uncertainty analysis because they quantify the level of uncertainty associated with a parameter estimate. These intervals could be obtained using bootstrapping, a resampling method applied to the observed data to explore the uncertainty in parameter estimates. Sensitivity analysis is another form of uncertainty analysis, in which aspects of the analysis (such as input data, underlying assumptions, or model structure) are varied to assess the robustness of the results. All of these approaches aim to quantify the degree of certainty that can be placed in the findings.

References

Barrera, A., Gee, C., Wood, A., Gibson, O., Bayley, D. and Geddes, J. (2020) 'Introducing artificial intelligence in acute psychiatric inpatient care: qualitative study of its use to conduct nursing observations', *Evidence-Based Mental Health*, 23, pp. 34–38.

Bremner, R., Gibbs, A. and Mitchell, A.R.J. (2020) 'The era of immersive health technology', *EMJ Innovations*, 4(1), pp. 40–47.

Buckley, C., Malcolm, R. and Hanlon, J. (2024) 'Economic impact of a vision-based patient monitoring system across five NHS mental health trusts', *PLOS Digital Health*, 3(9), e0000559.

Carter, J., Jorge, J., Venugopal, B., Gibson, O. and Tarassenko, L. (2023) 'Deep learning-enabled sleep staging from vital signs and activity measured using a near-infrared video camera', *arXiv preprint*, arXiv:2306.03711v1.

Chaichulee, S., Villarroel, M., Jorge, J., Arteta, C., McCormick, K., Zisserman, A. and Tarassenko, L. (2019) 'Cardio-respiratory signal extraction from video camera data for continuous non-contact vital sign monitoring using deep learning', *Physiological Measurement*, 40, 115001.

Chua, K.C. (2026) 'Reducing incidents of self-harm in mental health wards: a tale of two rates', *Journal of Mental Health*, 35(3), pp. 293–294.

Clark, H., Edwards, A., Davies, R., Bolade, A., Leaton, R., Rathouse, R., Easterling, M., Adeduro, R., Green, M., Kapfunde, W., Olawoyin, O., Vallianatou, K., Bayley, D., Gibson, O., Wood, C. and Sethi, F. (2022) 'Non-contact physical health monitoring in mental health seclusion', *Journal of Psychiatric Intensive Care*, 18(1), pp. 31–37.

Climent-Pérez, P., Spinsante, S., Mihailidis, A. and Florez-Revuelta, F. (2020) 'A review on video-based active and assisted living technologies for automated lifelogging', *Expert Systems with Applications*, 139, 112847.

Dewa, L.H., Broyd, J., Hira, R., Dudley, A., Hafferty, J.D., Bates, R. and Aylin, P. (2023) 'A service evaluation of passive remote monitoring technology for patients in a high-secure forensic psychiatric hospital: a qualitative study', *BMC Psychiatry*, 23, 946.

Diao, J.A., Marwaha, J.S. and Kvedar, J.C. (2022) 'Video-based physiologic monitoring: promising applications for the ICU and beyond', *npj Digital Medicine*, 5, 26.

Ede, J., Vollam, S., Darbyshire, J.L., Gibson, O., Tarassenko, L. and Watkinson, P. (2021) 'Non-contact vital sign monitoring of patients in an intensive care unit: a human factors analysis of staff expectations', *Applied Ergonomics*, 90, 103149.

Gooding, P.M. and Clifford, D.M. (2021) 'Semi-automated care: video-algorithmic patient monitoring and surveillance in care settings', *Bioethical Inquiry*, 18, pp. 541–546.

Griffin, B. and Saunders, K.E.A. (2020) 'Smartphones and wearables as a method for understanding symptom mechanisms', *Frontiers in Psychiatry*, 10, article 949.

Griffiths, J., Saunders, K.E.A., Foye, U., Greenburgh, A., Regan, C., Cooper, R., Powell, R., Thomas, E.M., Brennan, G., Rojas-García, A., Lloyd-Evans, B., Johnson, S. and Simpson, A. (2024) 'The use and impact of surveillance-based technology initiatives in inpatient and acute mental health settings: a systematic review', *BMC Medicine*, 22, 364.

Heglum, H.S.A., Kallestad, H., Vethe, D., Langsrud, K., Sand, T. and Engstrøm, M. (2021) 'Distinguishing sleep from wake with a radar sensor: a contact-free real-time sleep monitor', *Sleep*, 44(8), zsab060.

Jackson, A.M., Mullican, L.A., Yin, J., Tse, Z.T.H., Liang, H., Fu, K.W., Ahweyevu, J.O., Jenkins, J.J. III, Saroha, N. and Fung, I.C.H. (2018) '#CDCGrandRounds and #VitalSigns: a Twitter analysis', *Annals of Global Health*, 84(4), pp. 710–716.

Jorge, J., Harford, M., Villarroel, M., Chaichulee, S., Davidson, S., Finnegan, E., Clark, S.H., Young, J.D., Watkinson, P.J. and Tarassenko, L. (2021) 'Non-contact assessment of peripheral artery haemodynamics using infrared video thermography', *IEEE Transactions on Biomedical Engineering*, 68(1), pp. 276–288.

Jorge, J., Villarroel, M., Chaichulee, S., McCormick, K. and Tarassenko, L. (2018) 'Data fusion for improved camera-based detection of respiration in neonates', *Proceedings of SPIE*, 10501, 1050112.

Jorge, J., Villarroel, M., Tomlinson, H., Gibson, O., Darbyshire, J.L., Ede, J., Harford, M., Young, J.D., Tarassenko, L. and Watkinson, P. (2022) 'Non-contact physiological monitoring of post-operative patients in the intensive care unit', *npj Digital Medicine*, 5, article 4.

Kamel, M.I. (2020) 'A view of the health services after COVID-19: an Egyptian perspective', *Alexandria Journal of Medicine*, 56(1), pp. 118–129.

Kekic, M., Rose, A., Baker, C. and Bayley, D. (2024) 'RETRACTED: Reduced self-harm on acute mental health wards following the implementation of a vision-based patient monitoring system: evidence from five NHS trusts', *Journal of Psychiatric and Mental Health Nursing*, 31, pp. e1–e9.

Kekic, M., Rose, A. and Bayley, D. (2026) 'Enhancing safety in mental health hospitals: a multi-site evaluation of a contactless patient monitoring platform', *Mental Health and Digital Technologies*, 3(1), pp. 32–45.

Lewis, A., Gandhi, V., Venugopal, B., Bowen, J., Polkey, M. and Gibson, O. (2021) 'Non contact patient monitoring and management system (Oxevision) in patient management', *Physiotherapy*, 114(Suppl. 1), p. e13.

Lewis, A., Venugopal, B., Gandhi, V., Gibson, O., Swanton, L., Green, M., Bowen, J. and Polkey, M.I. (2024) 'Remote vision-based digital patient monitoring of pulse and respiratory rates in acute medical wards', *Thorax*, 79, pp. 363–365.

Lloyd-Jukes, H., Gibson, O.J., Wrench, T., Odunlade, A. and Tarassenko, L. (2021) 'Vision-based patient monitoring and management in mental health settings', *Journal of Clinical Engineering*, 46(1), pp. 36–43.

Malcolm, R., Shore, J., Stainthorpe, A., Ndebele, F. and Wright, K. (2022a) 'Economic evaluation of a vision-based patient monitoring and management system in an acute adult and an older adult mental health hospital in England', *Journal of Medical Economics*, 25(1), pp. 1207–1217.

Malcolm, R., Shore, J., Stainthorpe, A., Ndebele, F. and Wright, K. (2022b) 'Economic evaluation of a vision-based patient monitoring and management system in addition to standard care for adults admitted to psychiatric intensive care units in England', *Journal of Medical Economics*, 25(1), pp. 1101–1109.

Michard, F. and Kalkman, C.J. (2021) 'Rethinking patient surveillance on hospital wards', *Anesthesiology*, 135(3), pp. 531–540.

Morrison, C. (2020) *Review of Current Use of Digital Solutions for Mental Health*. Glasgow: Digital Health & Care Institute, University of Strathclyde.

Ndebele, F., Wright, K., Gandhi, V. and Bayley, D. (2022) 'Non-contact health monitoring to support care in a psychiatric intensive care unit', *Journal of Psychiatric Intensive Care*, 18(2), pp. 95–99.

Ndebele, F., Wright, K., Gandhi, V. and Bayley, D. (2024) 'Non-contact monitoring to support care in acute inpatient mental health', *Journal of Mental Health*, 33(3), pp. 320–325.

Nolan, F. (2024) *A Multi-Site Evaluation of the Use of Vision Based Monitoring Systems & Body Worn Cameras in Mental Health Inpatient Wards in England*. Presentation, North London Mental Health Partnership and Anglia Ruskin University, 14 June.

Price, W.N. II, Gerke, S. and Cohen, I.G. (2023) 'Liability for use of artificial intelligence in medicine'. In: Solaiman, B. and Cohen, I.G. (eds.) *Research Handbook on Health, AI and the Law*. Cheltenham: Edward Elgar Publishing.

SQW (2025) *Exploring Evidence Regarding Vision-Based Monitoring in Inpatient Mental Health Units: Final Report to CQC*. July 2025.

Su, Z., Bentley, B.L., McDonnell, D., Ahmad, J., He, J., Shi, F., Takeuchi, K., Cheshmehzangi, A. and da Veiga, C.P. (2022) '6G and artificial intelligence technologies for dementia care: literature review and practical analysis', *Journal of Medical Internet Research*, 24(4), e30503.

Tarassenko, L., Villarroel, M., Guazzi, A., Jorge, J., Clifton, D.A. and Pugh, C. (2014) 'Non-contact video-based vital sign monitoring using ambient light and auto-regressive models', *Physiological Measurement*, 35, pp. 807–831.

Tully, J., Larkin, F. and Fahy, T. (2015) 'New technologies in the management of risk and violence in forensic settings', *CNS Spectrums*, 20, pp. 287–294.

Veale D, Ali S, Papageorgiou A, Gournay K. 2020. The psychiatric ward environment and nursing observations at night: A qualitative study. *Journal of Psychiatric and Mental Health Nursing*, 27(4), 342-351.

Villarroel, M., Chaichulee, S., Jorge, J., Davis, S., Green, G., Arteta, C., Zisserman, A., McCormick, K., Watkinson, P. and Tarassenko, L. (2019) 'Non-contact physiological monitoring of preterm infants in the neonatal intensive care unit', *npj Digital Medicine*, 2, 128.

Villarroel, M., Guazzi, A., Jorge, J., Davis, S., Watkinson, P., Green, G., Shenvi, A., McCormick, K. and Tarassenko, L. (2014) 'Continuous non-contact vital sign monitoring in neonatal intensive care unit', *Healthcare Technology Letters*, 1(3), pp. 87–91.

Villarroel, M., Jorge, J., Meredith, D., Sutherland, S., Pugh, C. and Tarassenko, L. (2020) 'Non-contact vital-sign monitoring of patients undergoing haemodialysis treatment', *Scientific Reports*, 10, 18529.

Wenzel, L. and Evans, H. (2019) *Clicks and Mortar: Technology and the NHS Estate*. London: The King's Fund.

World Health Organization (2018) *Digital Technologies: Shaping the Future of Primary Health Care*. Technical Series on Primary Health Care. Geneva: World Health Organization.

Wright, K. and Singh, S. (2022) 'Reducing falls in dementia inpatients using vision-based technology', *Journal of Patient Safety*, 18(3), pp. 177–181.

York Health Economics Consortium (2023) *NHS Innovation Accelerator Economic Impact Case Study: Oxevision*. Final version, January 2023.