AI DECISION SUPPORT FOR SYMPTOM MANAGEMENT IN NURSING HOMES

Usage framework and requirements for AI-supported non-urgent symptom management

Tania Braga

Supervisor: Mitha Rachel Jose

PhD I Senior Lecturer I Research Expert, Laurea University of Applied Sciences

Date: 28/02/2026

Co-funded by the European Union

AI DECISION SUPPORT FOR SYMPTOM MANAGEMENT IN NURSING HOMES

Usage framework and requirements for AI-supported non-urgent symptom management

Tania Braga

Supervisor: Mitha Rachel Jose

PhD I Senior Lecturer I Research Expert, Laurea University of Applied Sciences

Date: 28/02/2026

Co-funded by the European Union

Dedication / Acknowledgements

I would like to thank the interview participants for the thesis, who kindly shared their time at a period when time is increasingly difficult to manage.

I would also like to thank the nursing homes managers for their openness and cooperation in adapting daily work to enable the investigative effort, and for their positivity regarding its purpose and relevance. Your encouragement made this process so much more enjoyable.

I am grateful to the ManagiDiTH programme and the European Union for funding this education. Managing digital transformation in healthcare is extremely important and relevant with real potential to improve lives. It is essential to move forward, optimising resources and using available knowledge to innovate in ways to strengthen healthcare for the benefit of those who need it.

A special thanks to my thesis supervisor, Prof. Mitha Jose, for her guidance and availability.

Thank you to my family and friends who supported me in times when energy and enthusiasm lingered ever so slightly.

Finally, and most importantly, thank you, my beautiful son, Noah, for your patience, kindness and understanding. I could not have done it without you.

Abstract

Norwegian nursing homes face an increased work burden due to population ageing and insufficient staffing. Frontline staff manage frequent non-urgent symptom changes and decisions about assessment, follow-up, handover, and escalation.

The thesis examined whether and how an AI system could support the non-urgent symptom change management and escalation. A scoping review and qualitative study design were employed. For the qualitative study, data were collected in Stavanger, Norway, at three nursing homes using individual interviews (n = 15) and one focus group that used a case-based discussion and a storyboard/prototype walkthrough and were analysed using thematic analysis.

Results show the decision-making experience was described as a process rather than a single decision, with varying support needs across different roles. Concise assessment support and easily communicable guidance, as well as criteria for escalation, were important for healthcare professionals, whereas prioritisation, concise summaries, and next steps in accordance with clinical roles were important for nurses. The physician perspective highlighted the need for high-quality information from the bedside healthcare professionals and trend summaries, recognising that there is rarely a single generic answer for complex residents.

The study concludes that for an AI CDSS to be safe and usable, it must provide "actionable explainability", fit into the workflow, be fast, integrated with the Electronic Patient Journal (EPJ), provide role-specific outputs, and consider residents' baselines. Finally, it requires clear governance to ensure AI serves as a "second opinion" under human clinical accountability.

Keywords: *Artificial intelligence*, *clinical decision support*, *nursing homes*, *non-urgent symptom changes*, *escalation/clinical deterioration*, symptom management.

JEL codes: I18 (Government Policy • Regulation • Public Health) and O33 (Technological Change: Choices and Consequences • Diffusion Processes.

Index

Dedication / Acknowledgements ii

Abstract iv

Glossary xiii

Introduction 1

Background and context 1

Problem statement and gap 1

Aim and objectives 3

Research Questions 3

Scope and limitations 4

Proposed Approach 4

Thesis structure 4

Use of AI tools 5

Chapter 1 - Literature Review 6

1.1 Search strategy and selection 6

1.2 Results: descriptive mapping of included studies (RQ1) 7

1.2.1 Analytical lens and extraction categories (RQ1a-d) 8

1.2.2 Operational definitions and descriptive mapping results 8

1.3 Results linked to RQ1 (c-d): Evidence maturity and clinical translation gap 10

1.4 Summary of gaps and rationale for qualitative study and synthesis output 11

1.5 Thematic synthesis aligned with RQ1 (a-d) 13

Chapter 2 - Analytical Framework 17

2.1 Symptom management as a decision episode: from information to action 17

2.2 What “decision support” looks like in NH/LTC — and where it does not match non-urgent escalation 18

2.2.1 Risk stratification/prediction (anticipatory decisions) 18

2.2.2 Detection and monitoring (change recognition) 18

2.2.3 Action-oriented recommendations (explicit recommendations associated with touchpoints) 18
2.3 Workflow touchpoints 19
2.4 Data sources: feasibility constraints and design drivers 19
2.5 Fit and accountability 20
2.5.1 Task–Technology Fit (TTF) as a practical fit lens (used qualitatively) 20
2.5.2 Explainability and accountability requirements across roles and workflow touchpoints 20
2.6 Consolidated analytical framework used in the thesis 21
2.7 Operationalisation for the qualitative analysis 22
2.8 Summary of the analytical framework 23
Chapter 3 - Methodology 24
3.1 Setting 24
3.2 Participants and sampling 24
3.3 Data collection methods 26
3.3.1 Participant materials and session procedure 26
3.3.2 Semi-structured individual interviews 26
3.3.3 Focus group 27
3.3.4 Case-based discussions 27
3.3.5 Storyboard walkthrough 27
3.3.6 Interview guide development and traceability 27
3.4 Transcription process 28
3.5 Data management 28
3.6 Data Analysis (Attride-Stirling Thematic Networks) 29
3.7 Rigor and trustworthiness (COREQ-aligned) 29
3.8 Ethics and GDPR 30
3.9 Use of AI tools 30
Chapter 4 - Findings 31
4.1 Data set and analytic output 31

4.2 Global theme 32
4.2.1 Turning scattered observations into a shared, actionable escalation story under time pressure and fragmented systems 32
4.3 Organising theme 1: From bedside noticing to “relevant signs” and next actions 32
4.3.1 What participants described (RQ2) 32
4.3.2 What this implies for requirements/safeguards (RQ3) 33
4.4 Organising theme 2: Workflow friction and fragmented digital ecosystems shape what “usable” means 34
4.4.1 What participants described (RQ2) 34
4.4.2 What this implies for requirements/safeguards (RQ3) 34
4.5 Organising theme 3: Role-dependent thresholds — what “support” means differs across HFAs, nurses, and the physician 34
4.5.1 What participants described (RQ2) 34
4.5.2 What this implies for requirements/safeguards (RQ3) 35
4.6 Organising theme 4: Trust, safety, and “explainability” are practical and patient-specific 35
4.6.1 What participants described (RQ2) 35
4.6.2 What this implies for requirements/safeguards (RQ3) 36
4.7 Thematic network summary (Attride-Stirling output) 36
4.7.1 Global theme 37
4.7.2 Organising themes: 37
Chapter 5 - Discussion 38
5.1 Purpose of the discussion 38
5.2 Summary of key findings 38
5.3 Summary of the findings in relation to RQ2 and RQ3 38
5.3.1 Organising theme 1- From bedside noticing to “relevant signs” and next actions 38
5.3.2 Organising theme 2 -Workflow friction and fragmented digital ecosystems shape usability (RQ2; informs RQ3 implementation conditions) 39
5.3.3 Organising theme 3: Role-dependent thresholds define what “support” means (RQ2; informs RQ3 role-adapted requirements) 40

5.3.4 Organising theme 4: Trust, safety and explainability are practical and patient-specific (RQ2/RQ3; informs safeguards and explanation needs) 41
5.3.5 Link back to the scoping review (RQ1): what the findings explain or extend 42
5.4 Implications for design and implementation 42
5.5 Strengths, limitations and reflexivity 42
5.6 Conclusion of the Discussion 43
Conclusion 44
Direct answers to each research question 44
RQ1: What does the existing evidence show, and why is there a clinical translation gap? 44
RQ2: How are non-urgent symptom changes currently managed and escalated in nursing homes, and what makes it difficult? 44
RQ3: What requirements must AI decision support meet to be safe, usable, and trustworthy in this context? 45
Contributions of the thesis 46
Practical Contributions: Requirements for the Frontline 46
Theoretical Contributions: Reframing "Explainability" 47
Methodological Contributions: Bridging Theory and Practice 47
Practical recommendations 47
Limitations 49
Future scope 49
Closing statement 49
Bibliographic References and Sources 51
Annexes 57
Annex A – Study Sourcing Log 57
Annex B – Included studies table (n=44) 63
Annex C – Prisma 68
Annex C1 — PRISMA flow 68
Annex C2 — Full-text exclusions table and duplicates 68

Annex D - Quality Rubric 70
Annex E - Recruitment texts 72
Email to Participants 72
Email to Nursing Home Managers 72
Sms to Participants 73
Annex F - Participant Information Documents 74
Individual interview 74
Focus group 76
Doctor Interview 78
Annex G – Interview Consent forms 80
Individual Interview 80
Focus Group 81
Annex H — Privacy statement (Personvernerklæring) 82
Annex I — Demographic/context form 83
Nurse and HFA 83
Doctor 83
Annex J – Interview Scripts 85
Individual interview script – HFA and Nurse 85
Individual interview script – Doctor 87
Focus group script 89
Annex K - Prototype/storyboard walkthrough screenshots 92
Annex L - Key gaps identified in the literature (NH/LTC symptom-related decision support) 95

Index of tables and figures

Index of Figures

Figure 1.1 - Evidence maturity versus innovation maturity of AI/CDSS for symptom-related decision support in NH/LTC (n = 44) 11

Figure 2.1 Workflow-aligned analytical framework for AI decision support in non-urgent symptom escalation 22

Index of Tables

Table 1.1 - Study type distribution across included studies (n = 44) 7

Table 1.2 - Distribution of included studies by thematic cluster (A–D) (n = 44) 9

Table 4.1 - Participant background characteristics (tick box form, n = 19) 31

Glossary

Abbreviation	Full Name
AI	Artificial Intelligence
AUROC	Area Under the Receiver Operating Characteristic curve
CDSS	Clinical Decision Support System
COPD	Chronic Obstructive Pulmonary Disease (See: KOLS)
CosDoc	A specific Electronic Patient Journal (EPJ) system
EPJ	*Elektronisk pasientjournal* (Electronic Patient Journal / EPR)
EPR	Electronic Patient Record
HFA	Healthcare Assistant in Norway (*Helsefagarbeider*)[1]
HCP	Healthcare Professional
ICT	Information and Communication Technology
IKOS	*Informasjon Koordinasjon og Styring* (Digital coordination system)
ISBAR	Airway, Breathing, Circulation, Disability, and Exposure (a systematic clinical assessment framework)
KOLS	*Kronisk obstruktiv lungesykdom* (Norwegian for COPD)
LLM	Large Language Model
LTC	Long-Term Care
MD	Medical Doctor / Physician
NEWS	National Early Warning Score
NH	Nursing Home
NLP	Natural Language Processing
RN	Registered Nurse (Sykepleier)
RQ	Research Question
XAI	Explainable Artificial Intelligence

[1] A helsefagarbeider is a qualified vocational care worker in the Norwegian health and care system. It is a formal and licensed healthcare profession that requires specific formal education and training.

Introduction

Background and context

Nursing homes care for older adults and are increasingly shaped by complex multimorbidity, frailty and high symptom burden. This trend is well documented internationally, including Norway, England, USA and Australia (Australian Institute of Health and Welfare [AIHW], 2024; Barker, 2021; Barstad & Hammersland, 2025; National Academies of Sciences, Engineering, and Medicine [NASEM], 2022). The proportion of long-term nursing home residents who need extensive assistance is increasing in Norway; specifically, the percentage increased from 76% in 2009 to 89% in 2024. The increasing trend in the acuity levels of nursing home residents and the shortage of health professionals require the efficient use of the available human resources and the competencies that they represent. (Helsepersonellkommisjonen, 2023).

In everyday nursing home practice, nurses and healthcare assistants (HFAs) must detect and respond to subtle changes in residents' condition (Bagchus, 2024). This study is shaped by significant workloads and limited or inconsistent availability of medical support (Helsepersonellkommisjonen, 2023). As a result, escalation decisions are often based on unstructured to structured observation/scoring systems (e.g., NEWS) and communication between professionals (Lapp et al., 2022), rather than on embedded computerized decision support systems (Helsepersonellkommisjonen, 2023 and Abdellatif et al., 2021).

It has been found that the adoption of AI-based decision support systems based on AI in the municipal care sector is limited, and most of these systems are still in the pilot or development stages and are not implemented for non-urgent symptom management (Helsedirektoratet, 2021; Linqvist Leonardsen et al., 2025).

Digitalizing decision support has the potential to improve consistency and safety in nursing-home care (Abdellatif et al., 2021 and Lapp et al., 2022), but usability and workflow fit remain decisive for its adoption and professional trust (Fossum et al., 2011 and Olakotan et al., 2025).

Problem statement and gap

Adoption and integration of information and communication technology (ICT) in the health care sector have historically been slower than in other sectors. Even though there are obvious potential benefits for the quality, efficiency, and safety of health care services (Kruse et al., 2017; Lukkien et al., 2024). New studies suggest that unfulfilled integration requirements, such as workflow integration, professional trust and, progressive implementation are the reason behind ICT implementation delays (and specifically AI-based technologies) in long-term care (LTC) services (Lukkien et al., 2024). While the systemic challenges are well understood, there is a significant knowledge gap regarding the

requirements for a successful ICT integration in the workflow routines of health care staff, including health care assistants and nurses.

This integration need is further supported by the high workload and workforce shortages experienced in health care services. This requires more efficient detection and management of changes in residents' health status (Helsepersonellkommisjonen, 2023; Bagchus, 2024). Clinical Decision Support Systems (CDSS) are systems based on computers that offer healthcare professionals evidence-based knowledge, as well as data specific to the patient, in order to optimize healthcare, improve safety, and increase efficiency. They seek to offer the patient useful and applicable recommendations during the time of treatment in various environments (Garg et al., 2005; Kawamoto et al., 2005) but working around the common obstacles in implementing these systems that include time, training, and organizational issues, among others, is a challenge (Meunier et al., 2023; Berkhout et al., 2025).

In nursing homes settings and by nurses in general, studies and evaluations of decision support have been described as relatively uncommon compared with other clinical settings, despite clear relevance to prevention, symptom recognition, and care planning (Abdellatif et al., 2020; Abdellatif et al., 2021; Thompson et al., 2021 and Rony et al., 2025).

Usability and workflow integration are important mentioned factors. Difficulty of accessing via the EHR (Electronic Health Records) and if requires too many steps, may increase staffs' adoption resistance, even when they recognize its usefulness (Liberati et al., 2017; Olakotan et al., 2025; and Abdellatif et al., 2021).

These challenges make nursing homes a particularly important setting for studying not only what decision support can do, but how it should fit into real work.

Recent developments in artificial intelligence (AI), specifically, machine learning predictions models as well as large language models (LLMs), have revived the interest in supporting decisions in long-term care. There still exists, however, a considerable difference between the state of the art in AI and actual nursing home practice. The literature review shows that the scope of current work related to AI in long-term care includes: predicting falls, infection, or deterioration; AI assisted deprescribing; agitation recognition in dementia; and application of LLMs to electronic health records text, but mostly with a focus that is not centred in integration, explainability, ethics, or nursing home staff needs (Abdellatif et al., 2021; Lapp et al., 2022).

In case of nursing homes, "success" is not solely dependent of algorithmic accuracy, yet it should be ensured that the decision support is linked to the actual decisions being made in the nursing home, the type of information the staff possesses, and the location in the work process at which it would be feasible to use the support without adding to the staff's burdens (Fossum et al., 2011). The current studies in the nursing homes concerning CDSS emphasize the importance of training and support,

leadership, and compatibility of the nursing home work tasks and documentation practices in relation to the nursing home's adoption and perceived usefulness (Olakotan et al., 2025). This implies that the approach to the application of AI in the decision support of symptom management should not be solely technical in nature.

Even symptom evaluation and escalation in nursing facilities may be improved through integrating structured assessment methods such as ABCDE with professional skills, experience, and national and local clinical recommendations, this integration effort is a high level cognitive process and may be a source of variability in care quality (Fossum et al., 2011; Standing, 2008). Hence, there is a clear benefit in implementing support systems that assist with a uniform evaluation process and make relevant data available at the point of care, provided they address critical issues of workflow fitness, professional trust, and patient safety (Abdellatif et al., 2021; Lukkien et al., 2024; Olakotan et al., 2025).

Aim and objectives

The overall purpose of the thesis is to explore how AI based decision support could safely support nurses and healthcare assistants (HFAs – "helsefagarbeidere") with non-urgent symptom recognition, assessment and escalation in nursing homes. The thesis also seeks to identify the design and workflow requirements such support must meet to be useful, safe, and trustworthy in reality.

The thesis focuses on non-urgent symptom situations. These have sufficient time for structured assessment, documentation, and escalation decisions and are the most frequent occurrences in nursing homes. Non-urgent symptoms also represent a realistic and ethically appropriate starting point for AI-supported decision-making.

Urgent and life-threatening symptom changes will not be considered within this assessment. Urgent clinical situations need to be handled via emergency procedures and time sensitive escalation pathways, which require a separate safety case, evaluation process, and associated regulatory burden compared to the non-urgent decisions considered here.

To gather requirements anchored in actual clinical workflows, this research employed a prototype/storyboard walkthrough as a design probe. This served to provide participants with a concrete example of AI-based decision support eliciting specific requirements regarding functionality and usability while identifying potential concerns and guidelines specific to workflows. This approach ensured the practical relevance of the study, triangulating findings from the scoping review and qualitative interviews with a context specific workflow analysis.

Research Questions

RQ1 (Scoping review): What AI-CDSS approaches for symptom based decision support in nursing homes have been reported, and what do these studies reveal about three factors. Firstly, what kind of

decision tasks are supported, then, which data were sources used, transparency and explainability, and finally, barriers and facilitators to implementation?

RQ2 (Qualitative): How do nurses and healthcare assistants (HFAs) in Norwegian nursing homes identify, evaluate, document, and escalate non-urgent symptom changes, and where do they experience information gaps or practical challenges in the workplace?

RQ3 (Synthesis/output): Which requirements and safeguards are necessary for AI-based decision support to be perceived as usable, reliable, and trustworthy by frontline staff?

Scope and limitations

The work presented in the thesis concentrates on non-urgent symptoms recognition, assessment, and escalation in nursing homes. The focus serves as a first step toward developing a set of requirements cantered on workflow for safe and useful AI based decision support. The work does not develop, test, nor validate an artificial intelligence system, nor does it measure clinical outcomes or effects at patient level. Therefore, this study is limited to identifying requirements and developing a framework.

Technical development of testing an AI system; testing of clinical outcomes and patient impact are beyond this research's scope.

Proposed Approach

The thesis approaches the three needs:

1. A rising clinical acuity and complexity among care home residents, including in Norway, thereby increasing the complexity of decision-making for care home staff.
2. The historical implementation gap to achieve adoption in routine clinical practice, that decision-support systems need to be user friendly and compatible with the workflow.
3. The emerging opportunity to leverage AI capabilities for safe, transparent, and practical decision support at the point of care.

The major contribution in the thesis is the development of a practical framework and set of requirements for AI-based non-urgent symptom evaluation and escalation in the nursing home environment, as has been identified through the scoping review and qualitative outcomes. This synthesis informed the design of a prototype/walkthrough storyboard, which functioned as a catalyst for requirement gathering.

Thesis structure

The structure of the thesis is as follows:

The introduction describes background, problem area, aim of the thesis, research questions, scope and expected contribution. Chapter 1 presents the scoping review, with a summary of the available evidence based on AI and decision support methods in the context of nursing homes and residential

aged care. Chapter 2 describes the theoretical framework developed to align the clinical decision points, staff workflow points, and technological design needs. Chapter 3 describes the research method, a qualitative approach in Norwegian nursing homes, with a prototype/walkthrough method for design requirements elicitation. Chapter 4 presents the findings with the developed framework aligned with the workflow and design requirements. Chapter 5 discusses the findings, literature review, practice, recommendations, limitations, and future research directions. The conclusion summarizes the thesis contributions, answers research questions, and gives recommendations.

Use of AI tools

AI-based tools (ChatGPT), were employed to facilitate brainstorming, organization, and refinement of language in the process of writing the thesis. The author maintains full responsibility for the academic assessment, interpretation and finalization of the text. These outputs by the AI were considered as draft support and reviewed using peer-reviewed literature and official sources. Confidential/personal level data from participants was also not used in the AI tools.

Chapter 1 - Literature Review

This scoping review was carried out to systematically map and understand the existing literature related to the area of clinical decision support systems (CDSS) and decision support based on artificial intelligence within nursing homes. The aim was not to assess the effectiveness of these interventions, but rather to analyse the types of interventions researched, the clinical decision addressed, the data sources utilised, and the types of outcome measurement employed, particularly regarding considerations based on workflow for the safe application of the technology.

The reason for conducting a scoping review rather than an effectiveness review was the heterogeneity of the literature, which includes a wide range of technology types (rule based CDSS, model based approaches, through to the more recent LLM/NLP approaches), diverse clinical outcomes (such as falls, infections, medication optimisation, and dementia/agitation), and various study designs.

The review aimed to answer the following question: RQ1:What AI/CDSS approaches for symptom-related decision support in nursing homes and residential elderly care have been reported, and what do these studies report regarding (a) decision tasks supported, (b) data sources, (c) transparency/explainability, and (d) implementation barriers/facilitators?

The results of the scoping review are presented in two ways. First, a descriptive summary describes the essential details of the studies included in the review. Second, a thematic synthesis (a–d) provides an overview of the findings relevant to the development of AI decision support requirements, focusing on their applicability to the Norwegian nursing context.

1.1 Search strategy and selection

The search strategy involved a scoping review methodology. To augment the traditional database search, AI tools (ChatGPT and Grok) were used for initial discovery and keyword identification. This created a preliminary set of peer-reviewed articles relevant to AI/CDSS in long-term care. Every candidate article was then verified and retrieved manually by the author to ensure source integrity.

Most papers were identified using specific Google/Google scholar searches by title, author and keyword. This was supplemented by systematic searches within university library databases (Laurea and ISCTE), including access to major repositories such as PubMed, CINAHL and Elsevier. For every paper identified, the abstract was examined; when appropriate, the paper was then closely read in full form. Papers were included in the review if they were focused on the nursing home or long-term, care setting (or were generally applicable) and if they described decision support in the context of AI/CDSS for symptom recognition, assessment, monitoring, escalation or closely related decision tasks such as falls, infection, deterioration, medication optimization, or dementia symptom support.

To facilitate transparency and traceability, a study sourcing log (annex A)[2] was kept for each of the studies that were retrieved, and a description was kept as to how the article was identified and which platform was used. In addition, a structured data extraction was performed for all articles that were finally included, and these factors were taken into consideration: theme, decision point category, research goal, AI method (if applicable), data source, methodology, key results and limitations. Finally, a quality rubric assessed the relevance of each article. It had binary criteria set as 0 and 1 and was employed to identify studies that were most relevant for the thesis, including NH/LTC setting, "transferability" of the studies to the Norwegian context, decision point defined, workflow touchpoint, external validation, and explainability (XAI).

1.2 Results: descriptive mapping of included studies (RQ1)

A total of 44 peer-reviewed studies were included and mapped in this scoping review[3], mapping the evidence using the RQ1 lens (decision tasks, data sources, explainability, implementation). As reflected in the mapped literature, decision support for symptoms in NH and LTC settings include a variety of technologies, from non-AI CDSSs, such as rule-based algorithms, clinical scoring, structured checklists, and decision support embedded in computerized workflow, to AI driven decision support (such as predictive modelling using machine learning and text analysis assisted by NLP/LLM). This wide spectrum of decision support applications in NH and LTC reflects the overlap in terminology, as well as the differences in operationalization of "decision support" in NH and LTC studies.

In terms of technological diversity, besides the differences already mentioned, there is also diversity of type of evidence presented in the literature. It appears that there is a range of study types, including model development/validation, usability studies, qualitative studies and cross-sectional surveys. Table 1.1 presents an overview of the types of studies included within this literature, and it appears that there is a preponderance of development, validation and mapping at the review level.

Table 1.1 - Study type distribution across included studies (n = 44)[4]

Study type (primary category)	N
Model development/validation (ML, NLP, Prediction)	21
Review (scoping/systematic/narrative)	10
Survey/cross-sectional	7
Qualitative (Interviews, Focus Groups)	2
Usability/simulation evaluation	2
Trial/controlled evaluation (RCT/Pilot)	2

[2] The Annex A is a log at study-level (44 included studies), so it is not a classic database search-run log.
[3] See Annex B for the full list of included studies.
[4] Each paper was assigned one primary category based on its main purpose.

1.2.1 Analytical lens and extraction categories (RQ1a-d)

To ensure uniform comparison across diverse studies, the scoping review utilized a standardized analytical lens with four components across all studies (RQ1a-d). Each study was coded with respect to:

a) Decision tasks supported: The types of decision work supported by the tool (e.g., change detection, prioritization of risks, determination of next steps in assessment, escalation decisions), as well as the specific decision point (e.g., monitor, assess now, escalate) if provided.

b) Data sources: The data that the system relies on (e.g., assessment/scoring data, vital signs or observations, medications and comorbidities, clinical documentation and electronic health record data, data streams).

c) Transparency/explainability: The presence or absence of explainability in the study, with a focus on the distinction between technical interpretability and explanations supporting frontline reasoning.

d) Implementation factors: The presence of barriers and facilitators to implementation in the real practice (e.g., workflow fit, usability, training burden, integration, documentation, governance, and accountability).

The coding also includes workflow touchpoints (routine and periodic, event-driven, continuous and passive) as well as user roles (e.g., health facility assistant, nurse, physician), as these types of decision support in home healthcare and long-term care are highly dependent on workflow and user roles. These are operationalised in the sections below.

1.2.2 Operational definitions and descriptive mapping results

The workflow touchpoints were conceptualized as a “practical window” of tool integration with routine work activities. In the literature searched, three types of touchpoint patterns were identified: routine and periodic touchpoints, which were integrated with organized activities (such as planned reviews and routine monitoring activities); event-driven touchpoints, which were triggered by incidents and changes (such as emerging symptoms and changes in vital signs); and continuous/passive touchpoints, which were background processes, often involving monitoring technology, that produced signals or notifications once a threshold was reached.

Based on RQ1(a), decision tasks were operationally defined as the cognitive and operational work the tool is meant to accomplish (e.g., “change detection”, “prioritizing risks,” “decisions of what to assess next”, or “decisions to escalate”). A decision point was defined as “the specific clinical decision supported by the tool” (e.g., “continue to monitor”, “assess now”, or “escalate to nurse/doctor”). Although most studies addressed important domains of NH/LTC care (e.g., falls, infection risk, decline, medication management, and behavioural symptoms related to dementia), fewer papers clearly

identified the specific decision point the tool supported in actual practice. A significant majority (82%) of the papers identified at least one workflow touchpoint, whereas only 68% specifically identified which decision the tool supported.

To answer RQ1 (b), studies were categorized based on the type of data used to inform decision support, such as clinical data, assessment and scoring data, vital signs and observations, medication and comorbidities data, and, to a lesser extent, data streams.

For RQ1 (c), the literature was coded on whether, and how, transparency and explainability were reported, and whether model interpretability was considered in terms of technical model explanation, such as feature importance, or through explanations for users that facilitate bedside thinking. While this section provides the descriptive mapping of the literature, these findings are further discussed in 1.3 (evidence maturity/translation gap) and synthesised in section 1.5. (Thematic synthesis).

Regarding RQ1 (d), the literature was coded for reported barriers to and facilitators of implementation of the system in terms of workflow compatibility, usability, staff training needs, governance, and other contextual factors pertinent to NH/LTC settings. (The synthesis of barriers to and facilitators of implementation is provided in Section 1.5 Thematic synthesis). Finally, given the reliance of tool use within the context of NH/LTC upon the recipients of the output, the studies were also categorized by the type of user. Within the literature reviewed, the type of user most commonly associated with assessment and decision tools is the nurse. The HFA (helsefagarbeider) member is most commonly associated with the monitoring and detection points involving escalation, and the physician is most commonly associated with the recommendation tools.

To facilitate consistent reporting across the thesis, the studies were also classified into four thematic categories which have been used consistently throughout the thesis: (A) Usability/Responsibility AI and workflow and Implementation Issues, (B) Risk Prediction/Early Warning Systems (covering infection, deterioration/mortality, and fall risk, (C) Dementia/Behaviour/Continence monitoring, and (D) Medication Safety/Deprescribing. For the studies included (n=44), Theme A consisted of 17 studies, Theme B consisted of 19 studies, Theme C consisted of 7 studies, and Theme D consisted of 1 study. These numbers are shown in the table below.

Table 1.2 - Distribution of included studies by thematic cluster (A–D) (n = 44)

Theme	Description	N
A	Usability / responsible AI / workflow & implementation	**17**
B	Infection/deterioration/mortality risk (incl. early warning/prediction)	**19**
C	Dementia / behaviour / continence monitoring	**7**
D	Medication safety / deprescribing	**1**
Total		**44**

1.3 Results linked to RQ1 (c-d): Evidence maturity and clinical translation gap

The analysed literature shows that we are mostly in a stage of foundational exploration of AI CDSS. There is a growing body of literature, but the majority still focus on the conceptualisation, such as feasibility or usability testing, of the tool/system or testing in controlled environments. Fewer have tested in real world environments, namely on everyday workflows of a nursing home. Also, only 5% (2/44) of the included studies were found to have performed an external validation or assessment at the level of a randomised controlled trial (RCT). In addition, only 32% (14/44) of the included studies assessed outcomes beyond the technical performance of the model (e.g., workflow impact or practical utility), which limited the ability to draw definitive conclusions about its usefulness and implementability in the nursing homes and long-term care (NH/LTC) setting.

Regarding transparency and explainability (RQ1c), 48% (21/44) included some form of explanation or transparency. However, this was often presented as technical interpretability (e.g., feature importance or model interpretability) rather than explanations designed to support frontline judgement, reasoning, and escalation decisions. This distinction is important in NH/LTC, where symptom changes are often subtle and contextual. In this environment, assessment typically depends on clinical observations, baseline function, clinical history, and team communication rather than a single linear indicator.

Regarding implementation factors and workflow touchpoints (RQ1d), common themes emerged concerning data availability, documentation practices, integration into daily workflow, training, usability, and governance issues such as action ownership. A practical implementation requirement is alignment between tool output and the role, authority, and accountability of the intended user. In NH/LTC, observations and initial detection are often performed by HFAs and nurses, while escalation decisions and treatment changes are typically the responsibility of nurses and physicians. The system that lack clear specification of the targeted decision point and user may therefore face adoption and safety challenges, even when predictive performance appears strong.

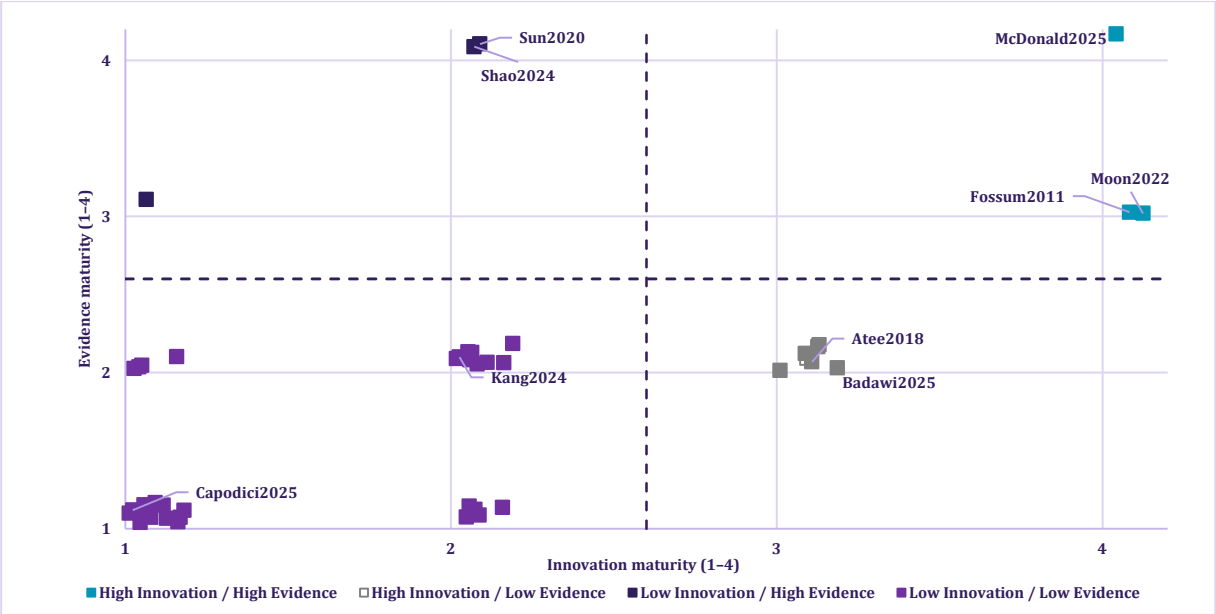


Figure 1.1 - Evidence maturity versus innovation maturity of AI/CDSS for symptom-related decision support in NH/LTC (n = 44)[5]
Source: Author's own figure (2026)

Figure 1.1 contrasts technical innovation with evidence maturity and translation to clinical practice across the 44 included studies. The mapped evidence indicates high innovation in AI CDSS for decision support related to symptom management in NH/LTC, but a comparatively immature evidence base, characterised by limited external studies and a persistent gap between algorithmic capability and practically implementable decision support. This gap provides the rationale for the subsequent thematic synthesis aligned with RQ1 (a–d) and for RQ2 and RQ3, which examine how non-urgent symptom changes are managed, from recognition to action or escalation in Norwegian nursing homes and how design requirements and safeguards can be specified in a usable and trustworthy framework for practice.

1.4 Summary of gaps and rationale for qualitative study and synthesis output

As shown in this scoping review, publications on AI CDSS in NH/LTC are increasing, but there remains a clear gap between what is technically possible and what is ready to implement in everyday nursing home practice. Many studies describe solutions at a high level, but less consistently specify the concrete decision points, roles, and workflow touchpoints where a tool would actually be used. In this review, 80% (35/44) of papers named at least one workflow touchpoint, while only 66% (29/44) clearly defined the specific decision being supported. This matters because symptom changes in NH/LTC are often subtle and highly context dependent, and responsibility for noticing and escalating them is shared across HFAs, nurses, and physicians.

[5] Only selected key studies are labelled to improve readability; all included studies are represented by points.

The evidence base is also still immature in terms of clinical translation. Only 5% (2/44) of included studies reported external validation or evaluation through a clinical trial, and only 32% (14/44) assessed outcomes beyond technical performance (e.g., workflow impact or practical utility). In addition, although 43% (19/44) reported some form of transparency or explainability, this was most often technical interpretability rather than explanation designed to support frontline judgement and escalation reasoning. Together, these limitations reduce the ability to draw firm conclusions about usefulness and adoption in routine nursing home practice.

These gaps make an in-depth qualitative study (RQ2) necessary to understand how non-urgent symptom changes are actually recognized, assessed, documented, and escalated in Norwegian nursing homes, and where information gaps and practical difficulties emerge within the routine daily workflows. This is especially important given the information fragmentation happening at the point of decision. Fragmentation occurs through staff, systems, local documentation routines, staffing patterns and escalation guidelines. Such variability is very difficult to capture on a linear performance report. This workflow relevance is shared by a Norwegian study of a CDSS integrated into an HER system within Norwegian nursing homes. The study found use of the CDSS encountered barriers such as insufficient training, resistance to computer use and integration/usability challenges (Fossum et al., 2011). On the contrary, when decision support is related to a clearly defined touchpoint with clear role responsibilities, such as support integrated into structured medication reviews, evaluation and utilization become more feasible, given that the “who acts and when” is well specified (McDonald et al., 2025). In addition, the lack of congruence between many data driven tools and practice is exemplified in prediction studies. These are often based on periodic assessment cycles and time horizons, such as 90 days, and may not align with changes in symptoms, such as in the middle of a shift, to drive escalation decisions in a nursing home setting (Boyce et al., 2022).

To strengthen the synthesis output (RQ3) and elicit concrete requirements and safeguards that fit the workflow, the thesis also uses a prototype/storyboard walkthrough as a design probe. A simplified escalation support concept (e.g., symptom cues, suggested observations, escalation criteria, and next steps specific for each role) allows participants to assess usability, fit with routines, and what would be needed for the tool to feel safe, reliable, and trustworthy in practice. Elicitation through a walkthrough is consistent with prior NH decision support evaluation approaches that used structured usability assessment and staff feedback to identify practical workflow and training constraints before broader implementation (Fossum et al., 2011).

The next section (1.5) synthesises the literature thematically in relation to RQ1 (a–d). The subsequent chapters then build on this evidence base by examining symptom escalation as it is carried out in practice (RQ2) and by specifying design requirements and safeguards for AI-supported decision-making in Norwegian nursing homes (RQ3).

Annex B summarises each included study's decision focus, workflow touchpoints, and findings relevant to implementation to support traceability from evidence to design. The framework and design requirements presented in Chapter 5.4 (Implications for design and implementation) and Chapter 6.3 (Practical recommendations) were developed by synthesising these insights from individual studies with the qualitative findings from interviews and the focus group.

1.5 Thematic synthesis aligned with RQ1 (a-d)

This section integrates the concepts identified in Section 1.3 and further developed in Section 1.4. The thematic synthesis combines the findings of the included studies with respect to Research Question 1(a-d): (a) supported decision tasks, (b) data sources, (c) transparency and explainability, and (d) implementation barriers and facilitators. The text follows a practical workflow perspective, focusing on who is expected to perform the action, when in the workflow, and what information and outputs the action is based on, because the vast majority of papers mention at least one point of contact with the workflow in their text (82%, 36/44). Fewer studies clearly define the specific clinical decision the system is intended to support (68%, 30/44).

In the thesis, decision support refers to information or structuring of a clinical decision (e.g., what to observe, what documentation is necessary). Computerized decision support systems (CDSS) include rule-based systems (e.g., scoring, checklists, guidelines) and artificial intelligence systems (e.g., predictive modeling, NLP). The systems' outputs must inform action at a specific point in the workflow.

(a) Decision tasks supported and types of AI/CDSS approaches

Across the included studies, decision support clusters around four recurring decision tasks: (1) change detection, (2) prioritising risks, (3) deciding what to assess next, and (4) decisions to escalate.

Change detection and assessment structuring are often operationalised through structured documentation, screening logic, or embedded prompts that aim to help staff notice deviations from baseline and collect relevant observations. Fossum et al. (2011) examined a nursing home CDSS delivered through the electronic health records, aimed at identifying and preventing common nursing home problems, as well as provide recommendations linked to care interventions. This type of support system is more closely related to the normal flow of clinical practice because it focuses on identification, documentation, and subsequent assessment, rather than prediction.

A major trend that emerged was the prioritization of risk, particularly within the more AI focused literature, where prediction of probability over a defined time horizon is used (for example, falls risk or deterioration/mortality risk). Boyce et al. (2022) created a model to predict falls, utilizing routinely collected long-term care data (minimum data set and drug therapy), which predicts falls over a defined future time interval. Susnjak et al. (2025) also utilized a prediction model to estimate survival over six

months after residential aged care facility admission, used for palliative care decisions. This type of prediction is important, but often used for periodic monitoring or planning, as opposed to the "right now" decisions that need to be made when a subtle non-urgent change is noticed during the course of a shift.

Escalation decisions play a key role in the thesis, although they have not been as clearly defined throughout the literature in relation to explicit decision logic. There have been many tools that can clearly indicate a risk but fail to clearly define the next step, leaving the practitioner unclear on whether to continue to monitor or escalate immediately (e.g., whether to continue to monitor or to escalate immediately). This is particularly relevant to non-urgent changes in symptoms, where decisions to escalate may be based on a series of small pieces of information and basic knowledge.

Lastly, recommendations can be most clearly seen in the area of medication optimization. McDonald et al. (2025) conducted a study on electronic deprescribing decision support and medication review, clearly defining a decision support that is directly linked to a clinical review and decision. Within the mapped literature, this area of medication safety and deprescribing appears to be a minority (Theme D, n = 1), so it represents a good but rare example of CDSS workflow alignment within nursing homes/long-term care settings.

(b) Data sources employed

The analysed data sources included in the studies follow the same categorization used during the data extraction phase. The sources include clinical data, assessment and scoring data, vital signs/observations, medication/comorbidity, and to a lesser extent, continuous data streams. Feasibility is greatly affected by what is normally available in nursing home (NH) and long-term care (LTC) settings.

The structured data sources and resident characteristics are commonly used because of their standardized form and availability. Features like admission features, comorbidities, cognitive impairment, and nutritional status are specifically designed to inform long-term clinical planning (Susnjak et al., 2024). For example, Boyce et al. (2022), combined Long Term Care Minimum Data Set (MDS 3.0) in combination with drug therapy, demonstrating the practicality of normally collected electronic data for predictive modelling in NH settings.

Decision support within Electronic Health Records (EHR) includes both embedded structured documentation and models that utilize clinical data and documentation patterns. A recurring practical theme is that the effectiveness of any decision support tool is only as strong as the quality of its input. Inconsistent documentation practices, high workloads, or hybrid paper and electronic routines can undermine these systems if key information is missing or delayed.

(c) Transparency, explainability, safety and accountability considerations

While repeatedly identified as a prerequisite to trust and adoption, its form ranges from model interpretability to explanations for clinician use. In the mapping, 43% (19/44) of the studies employed some form of transparency/explainability. However, this was most commonly in the form of model interpretability rather than explanations to inform clinician reasoning and responsibility in practice.

The study by Susnjak et al. (2025), presents a good example of a model that takes a transparency oriented approach through calibration, interpretability at both cohort and patient level, as well as the employment of reporting standards such as TRIPOD. Explainability here was framed as supporting clinical decision-making and ethical responsibility. At the same time, the authors highlight the importance of workflow integration alongside pragmatic evaluation.

From the viewpoint of safety as well as responsibility, the question seems not to just be whether the model can be explained, but also who is expected to take action as well as what action is expected to be safe at that role level. In Norwegian nursing homes, both HFAs and nurses are often the first to notice changes in direct care settings. Both can take immediate actions with low risk, such as encouraging fluids, assisting with toileting, observing, as well as monitoring vital signs. The nurses have clinical responsibility also for more complex actions, for assessment and escalation to a physician.

(d) Implementation barriers, facilitators and workflow touchpoints

The implementation evidence consistently points to the importance of workflow integration, usability, training, and documentation realities. This is well demonstrated in the evaluation carried out by Fossum et al. (2011) in a Norwegian nursing home. The facilitators were ease of use and favorable environments, while the barriers were lack of training, opposition to computer use, and lack of integration with electronic health records. The evaluation also points to the difference between those comfortable with technology and those less comfortable. This suggests that clinical decision support systems (CDSS) with good intentions can fail if the workflow touchpoint is not well chosen, if the interface is not user friendly, or if there is a lack of training.

The mapping also points to the reality that while a significant proportion of the papers have identified a workflow touchpoint (82%), fewer have identified the actual decision being supported (68%). The implication here is that ambiguity in “who does what, when” increases safety risks as well as implementation difficulties in nursing homes and long-term care settings.

Strongest examples of workflow alignment appear to be with interventions that are deliberately linked to established routines. In McDonald et al. (2025), deprescribing decision support is linked to medication reviews, with demonstrable effects. In contrast, many prediction models are anchored to

periodic assessment cycles (e.g., MDS-based prediction windows), which may align well with scheduled review and planning decisions, but do not necessarily map directly onto event-driven, mid shift escalation decisions that arise from subtle symptom changes in daily care.

Synthesis implications for the thesis

The mapped literature supports the need to: (1) specify decision points (especially escalation), (2) align outputs with real roles and accountability in nursing homes (frontline staff and nurses, with escalation to physicians when needed), (3) design for documentation and workflow realities in NH/LTC, and (4) treat explainability, safety and accountability as dependent on role and touchpoint requirements rather than generic add-ons (figure with summary of "Key gaps identified in the literature" on Annex L).

Chapter 2 - Analytical Framework

The thesis explores how AI-based decision support can be safely used to support nurses and healthcare assistants (HFAs) in recognizing, assessing, and escalating non-urgent symptoms within nursing homes, and what requirements this decision support must meet to be usable, trustworthy, and safe to be used in practice.

Chapter 1 gives an overview of AI/CDSS approaches used within nursing homes/long-term care (NH/LTC) and points out two challenges that are of particular relevance to the thesis. First, many studies provide an overview of AI/CDSS approaches without clearly indicating which decision point is being addressed (monitoring, immediate assessment, escalation, etc.). Second, even when a tool is judged to be technically viable, its actual success is to a large extent dependent on workflow compatibility, role clarity, documentation requirements, training requirements, and system integration (Fossum et al., 2011; Abdellatif et al., 2021; Lukkien et al., 2024; Olakotan et al., 2025).

Chapter 2 presents this mapping as an analytical framework to be used for the empirical part of the thesis (qualitative interviews, focus group interview, prototype walkthrough), as well as for the final synthesis (Research Question 3). The ambition is not to apply a general "technology adoption model," but to design a practical approach to analyse symptom related decision-making within nursing homes and identify when decision support can be safely used.

Scope note: urgent and time-critical symptoms are excluded from the thesis as this is handled by emergency plans that require a separate safety evaluation approach.

2.1 Symptom management as a decision episode: from information to action

A core assumption in the thesis is that non-urgent symptom management in nursing homes can be understood as decision episodes that unfold through everyday care. These episodes are often based on subtle changes, incomplete information, and shared responsibility between staff. In practice, staff move through steps like:

1. Notice - Identifying deviations from baseline (small deviation from baseline, new complaint, behaviour change, reduced intake, altered mobility, etc.)
2. Interpret - interpretation based on baseline knowledge, context, and judgement
3. Choice - deciding on a next step (monitor, assess now, act, document, escalate)
4. Action/communication - act and communicate (act, document, hand over, escalate)

This matters because AI/CDSS rarely "makes the decision". Instead, it changes what information is available, how it is presented, and what it suggests as a next step—thereby influencing decisions and escalation pathways.

To stay consistent with Chapter 1, the framework uses the same operational concepts throughout the thesis:

- Decision task: the type of cognitive/operational work being supported (e.g., change detection, prioritising risk, deciding what to assess next, deciding to escalate).
- Decision point: the specific decision that should become clearer at the point of use (e.g., continue monitoring, assess now, escalate to nurse/physician).
- Workflow touchpoint: the practical moment where decision support could realistically be used (routine/periodic, event-driven, continuous/passive).

These concepts keep the analysis grounded in real workflow moments rather than discussing models/outputs only in the abstract.

2.2 What "decision support" looks like in NH/LTC — and where it does not match non-urgent escalation

In Chapter 1, the studies included were grouped around four themes that will be used throughout the thesis: (A) usability/responsibility and workflow/implementation issues, (B) risk prediction and early warning, (C) dementia/behavior/continence monitoring, and (D) medication safety/deprescribing. As can be noted, across these themes, decision support can also be characterized by its function to the user, which will be discussed further to highlight areas for non-urgent escalation. Based on the literature that the thesis mapped, there are three main functional types of decision support in nursing homes/long-term care (NH/LTC):

2.2.1 Risk stratification/prediction (anticipatory decisions)

There are many AI systems that seek to forecast risks within a given time frame (e.g., fall risk, deterioration, death). As this can be very useful in prevention and anticipation, a problem that the thesis highlights repeatedly is that the results of these predictions do not always indicate a course of action that can be taken by the nursing staff in the event that there are subtle changes in the patients' conditions mid shift (Boyce et al., 2022; Susnjak et al., 2025).

2.2.2 Detection and monitoring (change recognition)

A second cluster aims to identify signals of deterioration by monitoring or by considering documentation as a signal source (e.g., documenting notes). While this can help identify issues earlier, it also prompts operational considerations that might be more important than the model itself: who receives the signal, how often, when, and what defines "actionable" information, as well as what defines "the next step" as safe (Fossum et al., 2011).

2.2.3 Action-oriented recommendations (explicit recommendations associated with touchpoints)

Recommendation based decision support is perhaps most obvious when it is associated with a workflow. For example, medication optimization/deprescribing is associated with scheduled reviews.

The pattern is highly relevant to the thesis because it illustrates an important principle: decision support is more “usable and safe” when it is associated with an “established clinical activity with clear roles and follow-up” (McDonald et al., 2025).

This study considers non-urgent symptom change management as a set of ad-hoc/event-driven decisions occurring during the daily routine. The healthcare professional notices the clinical change (symptom or sign) and faces the decision of whether to observe further without immediate action, evaluate with a new clinical assessment (tests/measurements), or escalate to a nurse or physician.

2.3 Workflow touchpoints

A fundamental concept of this research is the workflow touchpoint, which has been defined as the "practical 'window' of opportunity wherein a tool could be used without adding impractical steps."

Three touchpoint patterns are identified throughout this research:

- Routine/periodic touchpoints: planned activities, such as reviewing medication, scheduling, and periodic rounds.
- Event-driven touchpoints: those initiated by observed change, symptoms, deterioration concern, or an event.
- Continuous/passive touchpoints: continuous monitoring that produces alerts upon identification of defined thresholds or patterns.

Escalation of non-urgent symptoms is event-driven but also influences the other touchpoints, as symptoms can be managed through periodic routines or require immediate attention and communication. This framework helps identify where decision support should be situated and what type of output is most realistic at that specific moment.

2.4 Data sources: feasibility constraints and design drivers

Data sources were classified through the scoping review, and this classification was used because the feasibility of implementing AI in these settings depends heavily on data availability. This research has followed the same rationale, with design requirements being influenced by the availability of the data.

Data sources, as a constraint and design driver, are considered through the same classification system used in Chapter 1. Data sources are considered as follows:

- Structured assessments/scoring
- Vital signs and observations
- Medications and comorbidities
- Clinical/EHR documentation, including free text
- Data streams (if available)

A fundamental concept, as identified throughout the design, is the pros and cons of the data source type. Structured data supports scalability but may not account for subtle mid shift changes. On

the other hand, richer EHR or observation based inputs, can be more sensitive to daily changes but are heavily influenced by documentation quality.

2.5 Fit and accountability

The thesis does not treat implementation as a simple "acceptance" problem. Instead, it treats implementation as a problem of fit and accountability: whether the tool fits the real task, at a real touchpoint, for a real user role, with a clear and safe next step (Fossum et al., 2011; Abdellatif et al., 2021; Lukkien et al., 2024).

2.5.1 Task–Technology Fit (TTF) as a practical fit lens (used qualitatively)

Fossum et al. (2011) used Task–Technology Fit (TTF) to examine whether an EHR-integrated CDSS matched nursing personnel's needs in nursing homes. That logic is directly relevant here because the thesis focuses on whether decision support can fit the time pressured, distributed reality of symptom recognition and escalation in nursing homes.

In the thesis, TTF is not measured quantitatively. Instead, it is used as a sensitising lens during analysis and synthesis—so the work stays focused on practical questions such as: is the right data available at the moment the decision happens; can staff find baseline and recent changes quickly; does the output fit documentation routines and handovers; is it usable under time pressure; what training is needed; and does it reduce fragmentation or create extra work (Fossum et al., 2011).

2.5.2 Explainability and accountability requirements across roles and workflow touchpoints

Chapter 1 has already shown that explainability is often conceptualized as technical interpretability, rather than as an explanation that supports frontline reasoning and accountability within everyday decision-making situations. For the thesis, explainability is conceptualized as a safety requirement that must be compatible with:

(1) The decision point (monitoring, immediate assessment, escalation);

(2) The workflow touchpoint (routine, event-driven change);

(3) The user's role within that particular situation.

This is especially relevant to nursing homes within Norway, as nursing staff are required to be involved both as frontline caregivers and as clinical decision-makers. Both nursing staff and health and functional assistants (HFAs) may be the first to detect changes and take actions that are low risk, while nursing staff are responsible for clinical assessments and escalating to physicians when necessary. Thus, it is conceptualized as having accountability as a focus on who is to act, what actions are safe to take at that juncture, and what explanations are required to justify those actions.

Finally, besides fit, this framework conceptualizes safeguards as necessary conditions for safe usage. For the thesis, safeguards are:

- Explanations matched to user roles (not just technical interpretability);
- Ownership and accountability (specifying who is to act and what action is anticipated);
- Governance (specifying escalation triggers, who to contact, and when);
- Feasibility of data and documentation (specifying what to document and where).

2.6 Consolidated analytical framework used in the thesis

Building on the above, the thesis uses a four layer analytical framework to guide interview and walkthrough prompts, qualitative coding, and synthesis of requirements.

Layer 1: Decision episode (information to action chain)

Notice → interpret → decide next step → act/communicate

Layer 2: Workflow embedding (touchpoint pattern)

Routine/periodic vs event-driven vs continuous/passive

Layer 3: Decision-support function and decision point

Function type: prediction/stratification; detection/monitoring; recommendation

Decision point: monitor; assess now; escalate; initiate low risk action; schedule review

Layer 4: Conditions for safe use (fit + safeguards)

Fit (informed by TTF): access, timeliness, compatibility, usability, training, reliability (Fossum et al., 2011)

Safeguards: explanation needs; ownership/accountability; governance and escalation rules; data feasibility; documentation expectations

Such fit requirements and safeguards are detailed in Section 2.5 and used to guide qualitative coding in Section 2.7.

The framework described herein is shown in Figure 2.1. This figure illustrates the decision episode sequence and the concentric layers of fit and safeguards.

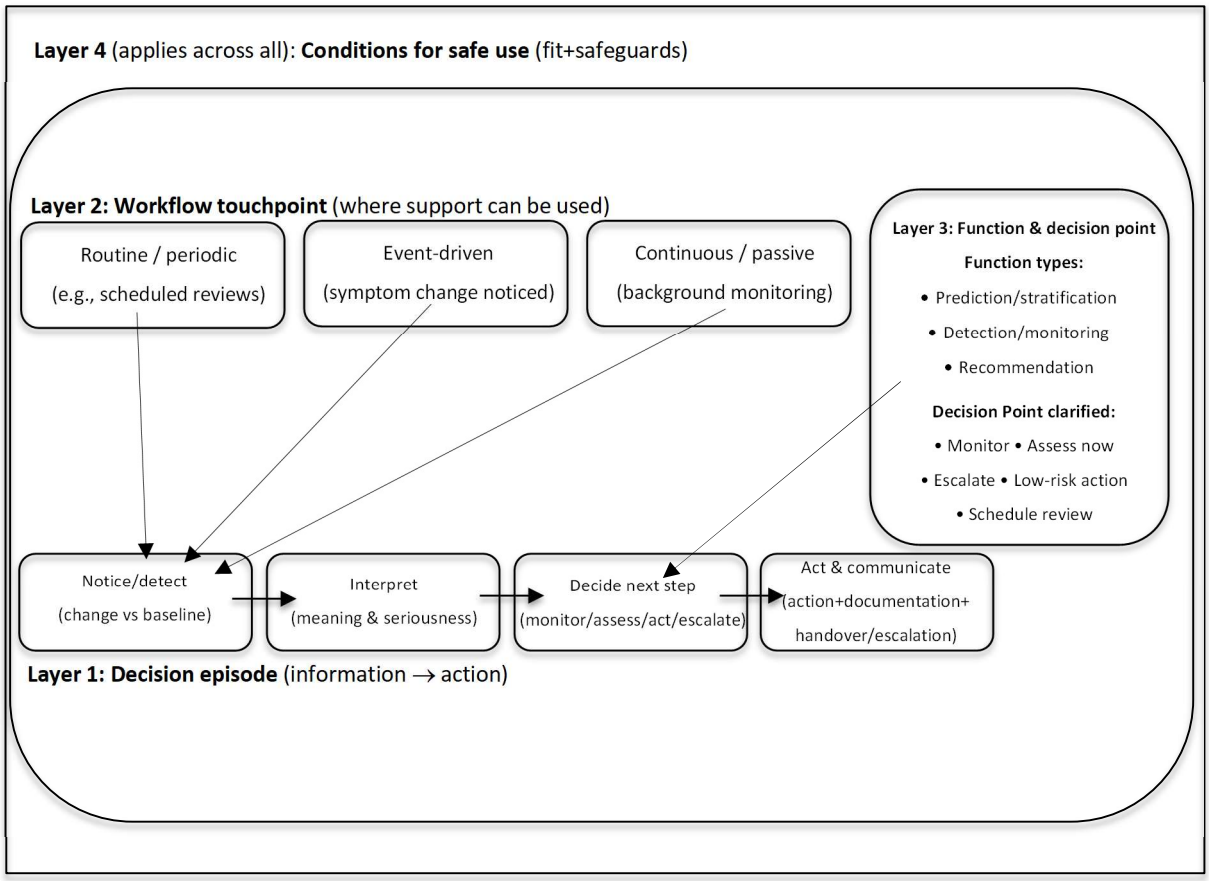


Figure 2.1 Workflow-aligned analytical framework for AI decision support in non-urgent symptom escalation

2.7 Operationalisation for the qualitative analysis

In the qualitative phase, findings will be analysed by mapping decision episodes (real or typical examples of symptom recognition and escalation described in interviews, the focus group, and the walkthrough) to the framework above. To ensure consistency across different interviews and settings, each clinical situation described by participants was summarized and analyzed using a common set of coding questions such as:

- Trigger/notice: what changed, and how was baseline considered?
- Interpretation: how was seriousness/non-urgency judged? what cues mattered?
- Decision point: monitor, assess now, low risk action, document, escalate — and why?
- Action/communication: who was informed, how, and when? what was documented?
- Touchpoint: routine/periodic, event-driven, or continuous/passive
- Role/accountability: who noticed, who acted, who escalated, who had follow-up responsibility?
- Data feasibility: what data would the tool need, and is it realistically available in practice?
- Explanation requirement: what explanation would make support trustworthy in that moment?

- Safety/governance: what checks, escalation rules, training, or documentation steps are needed?

This approach applies the same analytical lens used in the scoping review to the clinical descriptions provided by staff (RQ2), ultimately helping to define the requirements and safeguards for the final framework (RQ3).

2.8 Summary of the analytical framework

Chapter 2 presents a practice oriented analytical framework for understanding non-urgent symptom management as decision episodes embedded in real nursing home workflow. It links the scoping review mapping (decision tasks, data sources, transparency/explainability, implementation factors) to workflow touchpoints, decision points, and distributed accountability. Task–Technology Fit is used as a practical lens to define "fit" and requirements for safe use without claiming quantitative measurement (Fossum et al., 2011). This framework is then used to structure qualitative analysis and the synthesis of requirements and safeguards.

Chapter 3 - Methodology

This study was conducted with a qualitative approach, specifically within public Norwegian nursing homes in Stavanger, Norway. For this study, a multi site approach was employed, using individual interviews and a single focus group at three different nursing home locations. The study explores how registered nurses and HFAs/Helsefagarbeidere identify, assess, document, and escalate non-urgent changes in patient symptoms, and what workflow and safety requirements exist for an AI-based decision support system.

The choice of a qualitative approach was considered suitable because the aim was not to develop or test an AI system. The study focused on an understanding of actual decision-making processes and workflow, and on eliciting requirements for a safe and useful decision support system grounded in real everyday nursing home routines.

The study's objectives were explored with an integrated approach that includes:

1. a scoping review (RQ1),
2. a qualitative study conducted within nursing homes (RQ2),
3. synthesis and development of a practical framework and requirements (RQ3).

To make the requirements more tangible and grounded in workflow, a walkthrough with a prototype/storyboard was included in the qualitative study design, which was employed as a design probe. The walkthrough was employed to discuss potential locations for decision support, information that should be included, and what safety aspects should be included, but was not employed for effectiveness or model evaluation purposes. Situations involving urgent or acute symptom changes were beyond the scope, as these relate to time critical emergency situations, which require a different safety and evaluation approach.

3.1 Setting

The study setting comprised public nursing homes in the Stavanger area of Norway. Nursing homes often face critical time pressures characterized by poor documentation and limited medical support at the nursing home. Fluctuations of non-urgent symptoms are common phenomena that require evaluation, documentation, and escalation decisions. To the author’s best knowledge, there are no existing AI tools currently used in nursing homes across Norway that can support decisions for non-urgent fluctuations of symptoms at the nursing home level.

3.2 Participants and sampling

The study participants were sourced from three nursing homes, namely, NH1, NH2, and NH3, which were anonymized for this study. The sample comprised nurses and healthcare and affiliated staff (HFA) who were employed at nursing homes in Stavanger, Norway. These staff play a crucial role in

monitoring symptom change, taking initial assessment actions, documenting observations, and making escalation decisions, which answers RQ2.

A Purposive Sampling Approach

This sampling approach was adopted because it enabled a comprehensive sampling frame that considered variation across:

- Role, i.e., nurses and HFAs
- Years of experience in nursing home care
- Unit type, i.e., somatic and dementia units
- Level of responsibility for symptom assessment and escalation

The Study Sample

The study sample comprised:

- Sites: Nursing Homes 1, 2, and 3
- Sample size: n = 19
- Sample composition: 11 HFAs, 7 RNs, and 1 MD
- Data collection: 15 individual interviews and 1 focus group with 4 participants
- Inclusion criteria: employed at a nursing home/long-term care facility; a minimum of 1 year's experience at a nursing home

The sample was predominantly comprised of nurses, but with the inclusion of a MD, a comprehensive understanding was gained on how non-urgent symptoms develop, commencing with an HFA's initial observations, followed by a nurse's assessment, and concluding with a MD's decision-making process.

Recruitment strategies were adapted according to the conditions of each site:

- NH1: The leadership was contacted via email and declined data collection on site due to time constraints. The participants were approached individually by the researcher. The selection of participants was done by the researcher based on the criteria.
- NH2: The leadership was approached in person and then via email. The selection of participants was done by the leaders in consultation with the researcher based on the criteria.
- NH3: The leadership was approached via referral from another nursing home contact. The participants were then contacted via email by the researcher. The selection of participants was done by the leader according to the criteria.

The participants were given the information sheet and consent to participate at the beginning of the session (just before the interview or focus group). The participants were free to withdraw at any time without any negative effects.

3.3 Data collection methods

3.3.1 Participant materials and session procedure

At the start of each session, participants were provided with:

- Participant information sheet
- Consent form
- Privacy statement (personvernerklæring)
- Data Management Plan (DMP)

At the start of each session, the researcher:

1. confirmed the participant had read the information sheet and privacy statement (and provided a short verbal summary),
2. answered questions,
3. collected signed consent,
4. asked the participant to complete a short tick box context form (Annex I, 1–2 minutes),
5. then started the interview/focus group.

Audio recording of the participants' responses was obtained from 18 out of the total of 19 participants. The consent to use the recording device was denied by one participant. Handwritten notes were taken during the interview and focus group discussion. Where consent to use the recording device was denied, analysis of the data was done using the comprehensive notes taken by the researcher. The notes were also used to supplement the interview transcript in one instance when the recording of the interview was incomplete.

3.3.2 Semi-structured individual interviews

The individual interviews focused on participants' first hand experiences of non-urgent symptom change events, with specific emphasis on:

- Initial detection and interpretation of symptom change
- Information used versus information lacking
- Documentation habits and practical limitations
- Escalation criteria and ownership of decision-making
- Associated risks and opportunities of AI-based decision support, as well as trustworthiness

Interview duration: Up to 45 minutes, including greeting, introduction, interview, and closure. The planned interview content was approximately 35 minutes, with an added buffer of 5-10 minutes to account for time constraints on the participants.

Language: All interviews were carried out in Norwegian, with the first interview (P01) being conducted in English.

3.3.3 Focus group

A focus group method was used to explore norms and differences in evaluation, action, and escalation. The method also allowed us to test whether the requirements we had gathered in interviews are feasible across staff. The method was appropriate to discuss workflow alignment, documentation burdens, and governance conditions relevant to decision support.

Focus Group Duration: The maximum duration of the focus group would be 60 minutes. The time includes consent and closing remarks. The focus group will be in Norwegian.

3.3.4 Case-based discussions

Case-based discussions were incorporated into interviews and a focus group by using brief written scenarios. Participants described what they would do and how they would evaluate or monitor, assess, document, or escalate.

3.3.5 Storyboard walkthrough

A straightforward approach to a storyboard walkthrough was adopted to elicit requirements based on the workflow and to identify safety aspects. This approach was useful for achieving RQ3 by converting "opportunities" related to "AI" into design and safety requirements.

Walkthrough procedure:

- Intro (2 minutes): the tool is conceptual; the goal is requirements and safety conditions, not testing performance.
- Scenario (one vignette): a short non-urgent symptom change scenario relevant to nursing home work. In most cases, the same scenario was used as the case-based discussion (see 3.3.4).
- Walkthrough (10–12 minutes): the researcher presented the prototype/storyboard and asked questions along the way to validate understanding of workflow, information needs, and intended use. Participants then explored the prototype/storyboard.
- Requirements + safeguards (5–10 minutes): what the tool must show, what risks exist, and what safeguards are needed (kept short to fit the session time limits).

3.3.6 Interview guide development and traceability

The interview and focus group guides were directly derived from the scoping review coding framework and evidence gaps. The scoping review had coded studies by decision task, data sources, transparency and explainability, and barriers and facilitators to implementation. Additionally, workflow touchpoints, decision points, and user groups had also been coded.

As decision point had not always been clearly stated in the literature and as explainability had often been conflated with technical interpretability rather than bedside explanation, the qualitative

study guides were designed to uncover (1) realistic decision flows and information gaps (RQ2) and (2) requirements and safeguards for different user groups and workflow touchpoints related to usability, reliability, and trustworthiness (RQ3).

3.4 Transcription process

The following process was followed to transcribe interviews and the focus group:

1. Audio files were pseudonymized at the file level before processing.
2. Then, those pseudonymized audio files were stored on university cloud storage (student OneDrive) for backup purposed only.
3. The pseudonymized audio files were also saved to researcher's personal computer.
4. Transcription was performed using a local instance of the WhisperAI model running on the researcher's encrypted hardware, ensuring no data was transmitted to external cloud servers.
5. Transcription checked while listening to audio.
6. Clean transcript created (removed small talk and technical explanations not necessary for meaning; corrected obvious transcription errors while keeping meaning).
7. Anonymisation performed during cleaning.
8. Final transcripts stored.

The codes used during transcription were:

Code	Meaning
"short small talk and technical check removed"	remove off topic small talk or technical instructions/explanations
"INAUDIBLE min:sec]"	mark unclear or inaudible words
"Part of the theme removed – not relevant for the interview purpose"	remove off topic or not relevant information
"ANONYMIZED: name/place removed"	remove personal names or locations
"AUDIO STOPPED min:ss–min:sec"	mark where audio stopped
"NOTES summarized from handwritten notes"	mark when transcript is based on handwritten notes only

During the transcription, the focus was placed on the semantic content of the words, not their exact orthography in Norwegian. Minor changes to spelling and grammar were made, provided that they did not affect the content, in order to improve readability.

3.5 Data management

The audio recordings and transcripts were handled based on the principles of the GDPR, which highlighted data minimization, secure storage, restricted access, and deletion of data, as outlined in the DMP. The participants were given an information sheet, consent form, and privacy statement.

The audio recordings were stored on the researcher's mobile device, with transcription of the data completed within two weeks. Handwritten notes were included in the transcripts when the audio was

unclear or missing. The transcripts were anonymized by using nursing home codes, with all personal identifiers, such as the names of the staff, being removed. The names of the residents were not mentioned during the interviews. Only the researcher had access to the audio recordings and transcripts.

3.6 Data Analysis (Attride-Stirling Thematic Networks)

Data analysis was conducted in accordance with the thematic networks framework developed by Attride-Stirling (Attride-Stirling, 2001), using the data analysis framework presented in Chapter 2, which includes Decision Flow, Workflow Touchpoints, Decision Points, and Safeguards.

Data analysis was conducted in the following order:

1. Familiarisation: Interview data was studied to create initial impressions.
2. Initial coding: Relevant data related to Research Questions 2 & 3 were coded across the interview data (i.e., sentences and paragraphs).
3. Code consolidation: Similar initial codes were consolidated, their meanings clarified, and a code book created.
4. Basic themes: Initial codes were combined to create basic themes, which represent recurring, basic patterns.
5. Organising themes: Basic themes were combined to create organising themes, which represent broader categories of workflow and types of requirements.
6. Global theme: Organising themes were integrated to create a global theme that encapsulates the major overarching theme.
7. Review and refinement: The data analysis was conducted with a review of the interview data to avoid overgeneralisation.
8. Quotation selection: Quotations were selected to reflect common patterns while prioritizing recurring patterns while acknowledging unique perspectives.

Key documentation has been maintained to ensure transparency in the data analysis, which includes the final interview and focus group guides, vignette and storyboard materials, a list of codes with their definitions, and a theme development table.

3.7 Rigor and trustworthiness (COREQ-aligned)

To assure rigor and trustworthiness, the individual interviews and the focus group followed the COREQ's guidelines. The following points confirm the actions taken to ensure transparency and trustworthiness.

1. Documentation was created and saved, such as interview guides and materials, analysis notes, codebooks and theme development tables.

2. Triangulation by roles and sites (HFA/RN/MD; NH1-NH3)
3. Use of a focus group, triangulating responses with the individual interviews.
4. Actions to deal with potential biases: given the researcher's role as a nurse in the sector several precautions were taken, such as the: use of neutral prompts, avoiding direct subordinates as participants, assuring participants that the study is not looking for medical accuracy (there is not right or wrong answer), and avoiding reactive comments or facial expressions.

Study's limitations: Only one physician (MD) participated in an interview, and one interview used transcription based on notes; another used mostly notes due to incomplete audio.

3.8 Ethics and GDPR

The study was carried out with regard to all the principles of research integrity, as well as all the relevant data protection regulations, specifically the GDPR. The study was carried out on a voluntary basis, with consent being given prior to data collection. The participant was given an information statement about the data being processed.

In accordance with the participant's consent, the audio recordings were transcribed using the AI tool WhisperAI. Although the participant was assured that no sensitive data would be shared or stored outside of the EEA, the researcher recognized the inherent risks of cloud data processing, which included the location of the server and the caching of data. In order to mitigate the risks, the following data protection strategies were employed:

- Pseudonymization: the data was given only an alphanumeric code, e.g., NH1_RN_P03.
- Data sovereignty: the data was under the control of the researcher, as it was processed locally using the WhisperAI tool.
- Redaction: any mention of the participant's name was immediately redacted, and the audio was deleted once the transcript was verified.

3.9 Use of AI tools

The use of AI tools was for the purposes of brainstorming, organization, and refinement of language. There was no processing of confidential participant information using generative AI tools. Audio transcription using WhisperAI is outlined in the Ethics and GDPR sections.

Chapter 4 - Findings

4.1 Data set and analytic output

This chapter reports findings from three nursing homes in Stavanger (NH1–NH3). The qualitative material includes 15 individual interviews and one focus group (4 participants), giving a total of n=19 participants: 11 healthcare assistants (HFAs/helsefagarbeidere), 7 registered nurses (RNs), and 1 physician (MD). Participants described how non-urgent symptom changes are recognised, assessed, documented and escalated in everyday nursing home work, and they discussed a simple prototype/storyboard walkthrough ("TryggVakt") used to make decision-support needs, workflow fit, and safety risks more concrete.

Participants also completed a short tick box background form (Annex I). The sample included 11 HFAs, 7 RNs and 1 physician (n=19). Most worked in dementia units (12/19), with the remainder in mixed somatic/dementia units (5/19) or somatic LTC (1/19). Experience in nursing homes ranged from 1–3 years (3 participants) to 16+ years (5 participants), with the largest group reporting 8–15 years (8 participants). Most participants described non-urgent symptom change situations as weekly (10/19) or daily (7/19), and digital confidence was mainly medium to high (8/19 and 9/19 respectively). In daily work, CosDoc (EPJ) and IKOS were the most frequently reported systems. Background form data were used to describe the sample and check variation across roles and sites; the thematic analysis is based on the interview/focus group transcripts.

Table 4.1 - Participant background characteristics (tick box form, n = 19)[6]

Characteristic	Category	n
Role	Healthcare assistant (HFA)	11
	Registered nurse (RN)	7
	Physician (MD)	1
Unit type	Dementia unit	12
	Mixed somatic + dementia	5
	Somatic long-term care	1
	Not specified	1
Years of NH/LTC experience	1–3 years	3
	4–7 years	3
	8–15 years	8
	16+ years	5
Frequency of non-urgent symptom change situations	Daily	7
	Weekly	10
	Other/unclear	2
Self-reported comfort with digital tools	High	9
	Medium	8
	Low	1
	Unclear	1

[6] Background characteristics were collected using a short tick box form (Annex I) and were used to describe the sample and check variation across roles and sites. The thematic analysis is based on the interview and focus group transcripts.

Most reported digital systems used in daily work	CosDoc (EPJ)	17
	IKOS	15
	Sensio	5
	TQM	4
Use of any support tools (e.g., checklists/scales/alerts)	Yes	16
	No	2
	Not specified	1

The analysis followed Attride-Stirling's thematic network logic (Attride-Stirling, 2001): meaning units were coded, codes were grouped into basic themes, basic themes were clustered into organising themes, and these were synthesised into one global theme. Findings are presented primarily by organising themes (the main insight clusters), with each theme linked to RQ2 (how escalation work is done in practice) and RQ3 (requirements/safeguards for safe and usable decision support).

4.2 Global theme

4.2.1 Turning scattered observations into a shared, actionable escalation story under time pressure and fragmented systems

Across roles and sites, participants described non-urgent symptom escalation as a process of noticing small changes, interpreting what they might mean for this resident, and communicating them in a way that supports safe action by the next person in the chain. HFAs were often described as the first to notice change (especially in residents with dementia or limited ability to explain symptoms), but nurses also frequently notice change directly, since in Norwegian nursing homes they often combine hands-on care with clinical responsibility. Staff, particularly HFAs, rely on knowing the resident's baseline to detect when something is 'off'—even if they cannot immediately put it into clinical terms" (NH1_P01, NH1_P03 and NH2_focus group).

Two pressures were repeated across interviews and the focus group: (1) time/workload, and (2) fragmented documentation and digital tools. In that setting, "decision support" was mainly understood as support for structured assessment, next steps, and better communication, rather than "AI giving a diagnosis."

4.3 Organising theme 1: From bedside noticing to "relevant signs" and next actions

4.3.1 What participants described (RQ2)

Participants highlighted that "noticing" a change is not merely about a clinical score; it is deeply rooted in knowing the resident and their baseline. One nurse explained that knowing daily habits is what makes a sign "alarming" (RN_P07, NH2_focus group; RN_P01, NH1). Participants repeatedly described the early stage as a "diffuse" concern that must be translated into clinically relevant information: What do I check now? What is important to say? When is it enough to escalate? One nurse summarised the

value of structuring escalation as being able to “describe the situation” through “de relevante tegn” (relevant signs).

HFAs frequently highlighted the importance of the support being concise, concrete, and communicable. The following quote from an HFA describes the value of the support in terms of “tre punkter,” i.e., three bullet points of what to do, combined with ISBAR for what to communicate to the nurse or physician: “…tre punkter – hva du skulle gjøre. Og noen punkter av ISBAR – hva du skal si til sykepleieren eller legen…” (NH1_HFA_P03). Furthermore, the HFAs frequently stated a need for support that would help them translate a diffusive concern into action steps. Participants in both NH2 and NH3 expressed that a tool could help corroborate their “gut feeling” with objective facts: “Det ville vært nyttig å se det i fakta før man går til sykepleieren for vurdering” (P11, P13).

Notably, the participants did not agree on the factors being most important in the moment. A registered nurse downplayed the importance of the baseline information in favour of information about action steps: “Baseline … er ikke viktig. Nyttig: hva skal gjøres.” (NH1_RN_P02). The participant emphasised that on busy shifts, knowing what to do immediately is more useful than reading a full baseline. Therefore, this should not be mistaken for a general downplaying of the importance of the baseline information. Rather, it highlights the diversity of the perceived importance of information depending on the situation. For instance, the situation might call for rapid information about action steps during a busy shift, whereas a broader context might be relevant when escalations occur.

4.3.2 What this implies for requirements/safeguards (RQ3)

This theme translates into three core requirements:

- Action-oriented guidance: short “what to do now” steps (often described as a few bullets, not long text). Instead of generic risk scores, the tool should provide a list of concrete actions—such as checking vital signs, performing a urine stix, or giving fluids—which participants identified as their immediate first steps. This supports the prioritization of 'what should be done' because, as one RN noted, “we don't always have time” to dig through extensive history (NH1_HFA_P03, NH2_HFA_P05, NH2_focus group, NH2_RN_P06 and NH3_HFA_P13).
- Communication support: help converting observations into a structured escalation message (e.g., a structure based on ISBAR).
- Speed and simplicity: usefulness was repeatedly linked to how fast the tool is to fill in and read under shift pressure.

4.4 Organising theme 2: Workflow friction and fragmented digital ecosystems shape what "usable" means

4.4.1 What participants described (RQ2)

Participants described a working reality with many parallel tools and documentation surfaces. In the focus group this was summarised clearly: "Vi har hundre verktøy, men vi har ikke tid til å bruke dem." (NH2_focus group). Efficiency was a primary concern; one participant was very clear that they shouldn't have to log in constantly, stating, "You shouldn't have to log in... it takes a long time" (NH1_HFA_P03 and NH3_HFA_P13). They described it as a time problem more than a technology problem — the main issue was having to manage everything inside a shift. Furthermore, there is an existing level of skepticism regarding new tech; some noted that current digital systems, like the Sensio alarms, do not always work "optimally" (NH2_RN_P06).

Another concern was related to time wasted with double documentation. The healthcare professionals expressed that they do not want to write the same thing in several places: "gidder ikke å skrive 2 ganger samme ting… [det] må … rett til CosDoc." (NH1_HFA_P04). The effort to document in multiple tools is therefore seen as an inconvenience and a barrier to consistent use and a risk for information loss.

4.4.2 What this implies for requirements/safeguards (RQ3)

This theme leads into strong implementation and design requirements:

- Integration is an essential safety requirement: to avoid duplication, delays, and lost information.
- Workflow fit: the tool must match timing and handover realities in nursing homes, including how information travels during and between shifts.
- Low "documentation effort": minimal fields, minimal clicks, and a dual and clear purpose (a) support for decision-making now and (b) documentation for the formal record.

This also helps explain why interviewees often spoke more about workflow and documentation than about AI itself: the main barrier is not whether AI is "impressive", but whether the tool can realistically be used inside a shift.

4.5 Organising theme 3: Role-dependent thresholds — what "support" means differs across HFAs, nurses, and the physician

4.5.1 What participants described (RQ2)

Participants shared the same overall aim (safe care and appropriate escalation) but described different thresholds and different support needs depending on role and responsibility.

HFAs often described needing support that turns "a feeling" or observation into something they can check, act on safely, and communicate clearly. They frequently framed support as reassurance plus concrete steps and clearer escalation triggers.

Nurses more often described support as a structuring and prioritisation aid — ensuring key observations are checked, helping summarise and communicate efficiently, and supporting decisions about whether to escalate further. This aligns with the RN emphasis on “relevant signs” and a small set of action points (NH1_RN_P02).

The physician (MD) argued that the main advantage of a support tool would be to improve the quality and clarity of the information from the ward, which is the foundation ('i bunn') of the escalation process. The doctor’s point of view is that decision support should be more focused on ensuring that information is detailed and structured, rather than on replacing clinical judgment or providing a definitive diagnosis. Improved documentation will ultimately improve communication efficiency, saving time.

Regardless, the physician was sceptical towards tools that remind staff of basic assessment and tasks, like checking vital signs, because they should be part of the professionals’ standard competence. A core concern was that the tool should not be perceived as an absolute authority. This is particularly relevant for complex residents where, as the MD noted, there is rarely a single 'correct answer' or 'fasit': 'det er jo sjeldent en fasitløsning på de her pasientene' (NH3_MD_P18). Consequently, the MD argued that support should prioritize improving 'upstream' information quality, allowing the doctor to receive a clearer clinical picture from the start.

4.5.2 What this implies for requirements/safeguards (RQ3)

This theme supports a key design principle: decision support should be tailored to role, because “one size fits all” may under support HFAs and not meeting nurses’ prioritisation needs.

Practical requirements suggested in the material:

Output for HFAs: what to check now + when to escalate + what to say.

Output for Nurses: prioritised signs, concise summary, and structured content ready for escalation.

Output for Physicians: a short clinically meaningful situation summary (trend + relevant baseline + what was already checked) to reduce back and forth.

4.6 Organising theme 4: Trust, safety, and “explainability” are practical and patient-specific

4.6.1 What participants described (RQ2)

The accounts of the participants related to RQ2 show that safety was often linked to the notion of what is normal for a specific resident. The participants stressed the point that general levels might not always be appropriate. They used the example of chronic obstructive pulmonary disease (COPD or KOLS) to emphasize the point that an alarm for a low level of oxygen might not always be appropriate

if the resident has a naturally lower level of oxygen: “...har litt lav O2 fordi han har KOLS” (NH1_HFA_P04). This was a general reality in the environment of the nursing home. Safe decision-making was about knowing what was normal for a specific resident, rather than what was normative for the general population (NH1_HFA_P04, NH2_RN_P06, NH1_RN_P1, NH1_RN_P02, NH1_HFA_P04).

The participants were of the view that a micro explanation of terms and fields would help in avoiding input errors. They suggested the need for a brief explanation of terms such as ‘debut’ and ‘baseline’ (NH1_HFA_P04). However, this was not related to the idea of an AI function. In addition, the usefulness of the system was related to the presence of a ‘baseline/context’ field. This would help in the efficient communication of observations during escalation (NH2_HFA_P05, NH2_RN_P06).

Lastly, the issue of trust was a crucial prerequisite for the usefulness of the system. The participants were of the view that the system would have to have the approval of the municipality in order to be useful. In other words, the system would have to have the official stamp of approval in order for them to use it (NH1_HFA_P03, NH1_HFA_P04, NH2_RN_P06).

4.6.2 What this implies for requirements/safeguards (RQ3)

In this dataset, "explainability" is best understood as actionable explainability, not technical interpretability. Staff did not mention the need for better technology architecture; they required:

Actionable justifications: Short clinical reasons for why specific steps are suggested ("Why these steps?").

Escalation triggers: Clear indications of what specific data point changed the decision or triggered the alert ("Why now?").

Field-level guidance: Brief definitions or examples (e.g., for "baseline" or "debut") within the tool to prevent incorrect data entry.

The "Human Filter": A critical safeguard is that staff must still use their "head" and professional filter rather than "blindly follow" the AI's suggestions (NH1_HFA_P04, NH3_HFA_P13).

Transparency of Uncertainty: The tool should not hide uncertainty; outputs should signal when information is missing, when a baseline is unknown, and when human clinical judgement is required.

Contextual Sensitivity: Decision support must be sensitive to individual clinical histories and patient-specific "norms" to prevent false alarms.

4.7 Thematic network summary (Attride-Stirling output)

To make the analytic structure explicit, the themes can be summarised as a thematic network:

4.7.1 Global theme

The global theme is: Turning scattered observations into a shared, actionable escalation story under time pressure and fragmented systems

4.7.2 Organising themes:

From bedside noticing to relevant signs and next actions (RQ2; informs RQ3 outputs).

Workflow friction and fragmented digital ecosystems shape usability (RQ2; informs RQ3 implementation conditions).

Thresholds vary by role and define what “support” means (RQ2; informs RQ3 role adapted requirements).

Trust, safety and explainability are practical and patient-specific (RQ2/RQ3; informs safeguards and explanation needs).

Examples of basic themes feeding into these organising themes include: “three concrete steps / short outputs”, communication structured around ISBAR. Other examples are prioritisation of “relevant signs”, double documentation fatigue and need for integration, “hundre verktøy” and lack of time, baselines specific to each patient and risk of generic thresholds.

Chapter 5 - Discussion

5.1 Purpose of the discussion

The goal of this research was to explore the potential for improving the management and escalation of non-urgent changes in symptoms through the use of AI-based decision support. This research was conducted in Stavanger, Norway. In the interview and focus group, the participants emphasised the feasibility of a decision support solution as long as access could be obtained quickly within a work shift and the solution could be integrated with existing documentation tools (e.g., EPJ). They emphasised the value of knowing what is normal for each resident and of incorporating perspectives from each role (nurse, HFA, MD) for understanding changes in symptoms.

The focus was on non-urgent symptoms because urgent deterioration follows different protocols and pathways (emergency routines/acute response), and the data collection (case-based discussions and storyboard walkthrough) was built around everyday escalation and handover.

Next, the findings regarding RQ1, RQ2, and RQ3 will be discussed.

5.2 Summary of key findings

The results were organised into four themes that explained how non-urgent symptom escalation is managed in daily work and what makes it challenging. First, the process of escalation starts with a diffuse bedside worry and progresses to relevant symptoms, a summary, and next steps. Second, fragmented digital systems and workflow friction have a significant impact on usability. Third, the physician, nurses, and HFAs had different support needs. Fourth, there was a practical, patient-specific understanding of explainability, safety, and trust.

5.3 Summary of the findings in relation to RQ2 and RQ3

5.3.1 Organising theme 1- From bedside noticing to “relevant signs” and next actions

The research shows that changes in symptoms and signs are first recognised by the healthcare professionals closest to the residents: HFAs and nurses. It is essential to know the resident to even recognise the sometimes subtle changes. Then comes the decision of how to manage it, including evaluating further, acting and/or escalating. The identification of what constitutes a symptom change is based on individual professional knowledge and their knowledge of the patient's baseline, whilst being diffuse and not structured. Subsequent actions do not follow a rigid structure, yet they consistently include a “baseline assessment” of common vital signs. Whether or not extra data/tests are collected is varied/differed across participants. Escalation is also subjective and occurs both before and after making a full evaluation of the measurements. It also differs in the goal: it can be to inform if they know no further action is to be taken at the moment, or to escalate. Gaps often occur in the full and summarised symptom change evolution, and sometimes in the current status of the management

of those recent non-urgent symptoms. Several participants described that Information is more likely to be lost between weekends and week shifts.

This theme also explains the limited translation of many AI/CDSS tools identified in the literature to non-urgent escalation of symptoms. In this scoping review, we found that many studies simply referred to various types of decision-support functions without clearly distinguishing the specific point of decision to be supported (monitor, assess, or escalate) or how to map results to specific next steps or checks to be taken by staff. The results of this study suggest that it is exactly this translation phase, from general change to specific signs, and to clearly communicate escalation, which causes problems for staff. While the literature treats "deterioration" as a single event, the participants interviewed view it as a process (noticing to validating to communicating to acting).

This finding can be compared with previous predictive studies in nursing homes. For example, Garcés-Jiménez et al. (2024) developed a machine-learning model for infectious-disease detection using vital signs and contextual information and introduced a baseline module for resident-specific comparison, reporting anticipation of infection 2–4 days before antibiotic treatment. The study was based on structured telemonitoring, heavy data cleaning and doctor-labelled infection outcomes. The present study addresses a different and earlier stage of practice. The starting point here is often a vague concern at the bedside. Then staff have to decide what's important, turn that into a brief summary another staff member can quickly understand, and decide whether to monitor, assess further or escalate. So, although baseline-sensitive prediction is technically possible, that does not solve the everyday problem. In real nursing-home workflow, staff also need support for interpretation, communication and knowing what the next step really is.

5.3.2 Organising theme 2 -Workflow friction and fragmented digital ecosystems shape usability (RQ2; informs RQ3 implementation conditions)

The results indicate that, given HCPs already have to use so many tools in their daily work and documentation, any new tool, even to help them access and act in a situation needs to be quick to access and faster than existing workarounds. This means that the tool needs to be integrated with the EPJ, where the HCPs already document clinical information.

Decision support therefore needs concise, action-oriented outputs and communication support (i.e. ISBAR ready summary) without adding extra documentation. If it takes longer to use the tool than to find a nurse, the tool is likely to be abandoned.

These findings are largely consistent with the work of Lukkien et al. (2024), who found that routinization of AI-DSS use, balanced proactivity, and customization were important prerequisites for responsible AI-assisted decision-making in long-term care. The thesis confirms those general findings, but makes them more specific to the Norwegian nursing-home context, showing that workflow fit is greatly affected by fragmented digital systems, documentation burden and the need for access within

seconds during a shift. In this context, usability is not just a question of the quality of the interface, but a question of whether the tool is sufficiently embedded in the existing documentation and handover routines so it saves time rather than creating extra work.

5.3.3 Organising theme 3: Role-dependent thresholds define what “support” means (RQ2; informs RQ3 role-adapted requirements)

Several interpretations of the potential offerings of technological clinical decision support were identified. The health care professionals described a need for what to verify immediately, when to escalate, and what to say. The nurses need a priority sign, a concise synopsis, and a structural composition ready to initiate escalation. The physicians need a concise and meaningful situational summary, which includes trends, relevant baselines, and previous verifications to improve communication efficiency. The summary should particularly focus on improving the quality of the information from the ward. The need for a concise and meaningful summary from the physician implies a preference for improved information quality rather than the automated ability to diagnose. The physician’s recognition that there is rarely a “correct answer” (*fasitløsning*) for the patients (P18) establishes a definitive boundary for the application of artificial intelligence: it should be used to improve the quality of the information provided to the clinician so that he or she can make a better decision, rather than to provide the definitive “answer.”

In addition, the findings also nuance the literature by showing that, in Norwegian nursing homes, both HFAs and nurses may notice a non-urgent change in symptoms first, which influences what “role-appropriate” decision support should look like.

Another important aspect is that the tool should not just list the steps. Given the diffuse nature of symptom change detection in this context, the HCPs might feel the increased burden of responsibility. They use the tool to reduce this burden by having "fakta" (facts) to back up their intuition when talking to a nurse. They want to get a confirmation that their judgment is correct. Escalation in this context could be a collaborative process of sense making where staff move from a “gut feeling” to objective clinical signs.

Consequently, “safe and usable” decision support in this context should provide outputs tailored to the roles (HFA vs RN vs MD). Without understanding the roles, it may not sufficiently support users, may overload users, or may cause unsafe ambiguity about which person to take action and escalate.

The present findings also make earlier long-term care literature more concrete. Lukkien et al. (2024) argued that AI-DSSs should be adapted to different user groups, including caregivers with different roles and responsibilities. The current study is consistent with that view, but shows more clearly what this means in practice during non-urgent symptom escalation in nursing homes. HFAs were mostly looking for help with what to check, when to escalate and how to communicate the concern. Nurses wanted support for prioritising, a brief overview of the situation and support with

escalation. The physician, in turn, emphasized the need to keep a brief but clinically meaningful summary, especially baseline, trend, and what had already been checked. In that sense, this study does not just support the idea of customization. It shows how role-specific support needs to be built into the design of decision support if it is to match the actual work of escalation in nursing homes.

5.3.4 Organising theme 4: Trust, safety and explainability are practical and patient-specific (RQ2/RQ3; informs safeguards and explanation needs)

Currently, the escalation process functions as a translation of the concern to clinically significant symptoms, then to communication with other healthcare professionals, and finally to action. The decision to escalate a concern is often graded (monitor, assess, or escalate). Explainability is best understood by staff as actionable explainability rather than technical interpretability.

This means that for a decision support tool to be safe and usable, it must provide actionable justifications and escalation triggers that act as a 'second opinion' for the human to consider. This finding serves to reinforce the notion that clinical judgment is the prerogative of the user, especially when the tool is able to indicate a level of uncertainty through the presence of missing information. The move toward the notion of contextual sensitivity, whereby the tool is able to adapt to the clinical histories and resident norms through micro explanations (such as the need to define “debut” or “baseline” in order to input data correctly), serves to legitimise the notion of the safety requirement of personalised AI tools.

These findings contrast with the literature identified in RQ1, which abounds with studies that describe explainability in terms of technical interpretability, e.g., in relation to the mechanisms of the model or the identification of the important variables. However, the participants in the present study defined “explainability” as practical and focused on action, ensuring that the clinical judgment resides in the hands of the user and that the subsequent actions are performed in a safe manner.

The participants’ focus on resident specific “norms” (e.g., a COPD/KOLS patient having a lower “normal” O2 saturation) demonstrates that a tool using generic clinical thresholds could lead to false escalations or missed deteriorations. Therefore, personalised AI—which accounts for these individual baselines—is an essential clinical safety safeguard rather than a secondary feature.

This finding also clarifies explainability in the context of nursing homes. The literature so far has often discussed transparency and responsible AI at a fairly high level, while the evidence reviewed from the studies on fall prediction has identified typical gaps in transparent reporting and fairness assessment. The present study supports the importance of transparency, but for frontline staff, explainability is experienced mainly as practical rather than technical. The staff were more concerned about whether the recommendation made clinical sense for this resident, what information was

missing and what safe next step should be taken, than about the internal logic of the model. In this sense, the current study adds to the existing literature by demonstrating that explainability in the context of non-urgent symptom management must be actionable, role-sensitive and anchored in patient-specific baselines.

5.3.5 Link back to the scoping review (RQ1): what the findings explain or extend

When combined, the qualitative results provide an explanation for why many AI/CDSS studies are difficult to apply in real-world settings. The decision point, intended user, or ownership of the next step are frequently not explicitly stated when workflow touchpoints are identified in the literature. The thesis demonstrates why those omissions are important in nursing homes: staff members require useful advice that can be translated into safe action, responsibility varies across roles, and escalation occurs under time constraints. In this context, actionable explainability, role clarity, and workflow fit seem to be more important for practical use than just predictive performance.

5.4 Implications for design and implementation

In summary, the following are the practical implications to design, functionality and implementation of AI-supported non-urgent symptom management in Norwegian nursing homes:

- Integration with EPJ systems, like CosDoc/IKOS, to prevent fragmented data and redundant documentation.
- Role tailored results for doctors, nurses, and HFAs.
- Event-driven use with assistance for handover and evaluation as soon as a change in symptoms is seen.
- Summary prepared for ISBAR to enhance communication.
- Access to pertinent baseline and historical data to avoid generic thresholds
- Quick, user-friendly design that takes into consideration different levels of digital confidence.
- Governance that upholds the clinical responsibilities of RNs and MDs and views the tool as a supplement rather than a substitute.

5.5 Strengths, limitations and reflexivity

This study's main strengths are the number of nursing homes where the participants work: 3; the role mix (from HFAs to nurses and a physician; the feedback from one focus group; the use of real cases and of a prototype/walkthrough, which makes the perspectives more concrete.

Limitations include only one medical doctor (MD) was interviewed, one interview transcript was largely based on notes (RN_P14), no direct observation of escalation in practice; no real AI tool was tested; the Stavanger based sample and the potential researcher role bias. On the other hand, the

researcher, being a nurse, also promotes a perspective closer to the context and reality. This allowed for building trust with participants.

Recommendations and next steps

For future work, it is recommended to conduct interviews in other Norwegian municipalities and in other European countries. Furthermore, a pilot study should assess a prototype integrated into the workflow, with attention to usability, safety, and impact on the quality of communication and the speed of escalation.

5.6 Conclusion of the Discussion

A main point from this discussion is that AI support in this setting has to work with practice, not against it. Staff need something they can access quickly, use during a shift, and connect to the documentation and handover systems they already have. The support also has to differ by role, because HFAs, nurses, and physicians are not doing the same job in the escalation process. Resident baseline and recent change have to stay central. So does communication. What participants described was not a system that takes over judgement, but one that helps them think, check, and explain.

Conclusion

The thesis undertakes to explore the possibility of AI-based CDS to improve non-urgent symptom escalation management in nursing homes, particularly with respect to the Stavanger region of Norway and the day to day CDS needs of healthcare assistants and nurses. The thesis also recognizes the specific design, workflow, and governance needs to ensure that AI-based CDS can be considered safe, usable, and trustworthy. The research combines a scoping review of existing AI/CDSS approaches to non-urgent symptom escalation management in nursing homes/long-term care with qualitative research conducted in three nursing homes in the Stavanger region. Through case-based discussions and walkthroughs of a storyboard/prototype, the research makes non-urgent escalation management and CDS needs more concrete.

Direct answers to each research question

RQ1: What does the existing evidence show, and why is there a clinical translation gap?

The thesis shows that while there is increasing literature in the field of AI/CDSS that is relevant to nursing homes and residential aged care, there is still a challenge in translating solutions into the practical management of non-urgent changes in symptoms. This is clearly shown in the qualitative findings, whereby there is a lack of specificity in what exact decision point the tool supports (e.g., monitor vs assess now vs escalate), as well as a lack of information regarding exactly what the roles are that will be involved in the next steps in the process. In nursing homes, non-urgent symptom escalation is not a singular event. It involves a sequence of actions that begin with recognition, proceed to validation using a limited list of relevant signs, communicate a brief situation summary, and end in action. When the exact decision point is not clearly specified, as well as the roles that will be involved in the next steps in the process, translating model output into safe action in a time pressured environment is difficult.

Moreover, the thesis shows that there is a lack of understanding of exactly what is meant by “explainability” in the literature, in that while there is a focus in the literature on technical explanations of model output, in nursing home practice, “explainability” is a practical concept, involving whether or not the next action suggested is reasonable in terms of the particular resident, as well as whether or not the tool supports safe action in uncertain situations.

RQ2: How are non-urgent symptom changes currently managed and escalated in nursing homes, and what makes it difficult?

The thesis shows that non-urgent symptom changes are typically first recognised by the staff closest to the residents, healthcare assistants (HFAs) and nurses, especially for residents with dementia or limited ability to describe symptoms. Early concerns are frequently experienced as diffuse, and then gradually clarified through baseline knowledge and checking a small set of signs and measurements.

The escalation process depends, therefore, on knowing the resident: knowing what is “normal” for the resident is essential for recognising subtle change and judging whether the situation is developing or stabilising.

The findings also show that subsequent actions are not performed through one rigid structure, yet they consistently involve a form of baseline assessment (often including vital signs), followed by decisions to monitor, evaluate further, act, and/or escalate. Thresholds for escalation vary across situations and participants. Escalation may occur before completing a full set of measurements (to ensure safety), or after additional information has been gathered to strengthen the clinical message. A central practical challenge is not only making the assessment, but translating the concern into a short, clinically relevant communication to the next person—particularly across shifts and roles. Participants described that information is at risk of being lost across shift changes, with specific vulnerability between weekends and weekdays, where the “red thread” of symptom change and the specific plan for follow-up may be lost.

The results suggest that workflow friction and fragmented digital ecosystems affect escalation quality and usability conditions. A clinician must deal with several tools and documentation needs in a constrained time schedule. Therefore, if a clinician must deal with a new tool that takes a long time to access, or if it does not reduce the workload, it will be perceived as not being useful. Thus, in this time pressured environment, usability is effectively measured in seconds; any tool that adds to the documentation burden is unlikely to be adopted. This context explains how non-urgent symptom management can be difficult even for skilled clinicians because of the fragmented nature of the documentation environment.

RQ3: What requirements must AI decision support meet to be safe, usable, and trustworthy in this context?

The thesis shows that “safe and usable” AI decision support for non-urgent symptom management in nursing homes must be designed around workflow, roles, and resident complexity, rather than functioning as a generic guideline tool or a system that attempts to deliver a single “correct answer.” The findings indicate that support needs differ across roles. HFAs need guidance on what to check, what to do now, and what to communicate (often seeking confirmation that their judgment is reasonable and sufficiently “factual” to communicate). Nurses require prioritized indicators, a summary that is concise and structured, and next steps that are clear and appropriate to their role in assessment and escalation. The physician perspective emphasises the value of a short, clinically meaningful situation summary (trend, relevant baseline, and what has already been checked) to improve communication efficiency and information quality, acknowledging that in complex residents there is rarely a single “one fits all” solution (“fasitløsning”).

In terms of the different roles, the research identifies the essential requirements for implementation and design. The first is that the tool should be integrated into the existing documentation ecosystem (EPJ), which would help to eliminate the possibility of "double documentation" and "information fragmentation." The use of a "silo" approach is considered a safety risk. The second is that the tool should be able to support use on a "event-driven" basis at the time a change is noticed during a shift, as well as supporting the handover and periodic review processes. Third, outputs must be concise and action oriented, including communication support such as an ISBAR ready summary, because escalation and safe follow-up depend heavily on the quality and structure of information passed between staff.

The present study proves that trust and safety are based on the particular clinical realities of individual residents rather than on generic guidelines. Some participants mentioned that, for example, residents with certain medical conditions have different thresholds/different “normal range” clinical values (like COPD/”KOLS”). Here, personalization emerges as a critical safety prerequisite rather than a technical nicety. In a similar vein, staff perceptions of explainability are based on the practical bedside understanding of why a particular check might be suggested and what action might be taken when the data is incomplete. By signalling uncertainty, the tool reinforces the “human filter” identified in the interviews. It serves as a second professional opinion.

The study also proves that technology cannot bypass existing hierarchies. Governance has to explicitly ensure the RN and physician roles as clinical leads, with the AI in a supportive role. The role of the AI is to improve the quality of the upstream information and make it easier for safe escalation. In the particular context of the Norwegian nursing home, feasibility and safety are not based on the performance metrics of the model but rather on the alignment with the workflow, the respect for professional roles, and the accommodation of each resident’s unique baseline.

Contributions of the thesis

The thesis contributes to the digital transformation of nursing home care by defining the requirements for AI to safely support staff when a resident’s condition changes (non urgently). The findings suggest that the "translation gap" in AI is not just about how well a model performs, but about whether its output can be turned into a safe clinical action within a busy, fragmented work environment.

Practical Contributions: Requirements for the Frontline

A major contribution of this work is a set of implementation requirements built directly from the perspectives of those working at the bedside. These go beyond technical specs and focus on clinical reality:

- Seamless Integration: Multiple systems often lead to duplicate documentation. It is necessary to move past "silos" to ensure tools work within the existing EPJ.

- Role Awareness: Recognising that an HFA needs different support than a physician.
- Communication Support: Using ISBAR-ready summaries to bridge the gap between shifts and roles.
- Safety Safeguards: Using resident-specific baselines (like the COPD/KOLS example) to prevent the dangers of generic clinical thresholds.
- Governance: Establishing that AI is a subordinate tool while the RN and physician remain the clinical leads.

Theoretical Contributions: Reframing "Explainability"

This research reframes how explainability is understood in the reviewed literature, by emphasising practical, action oriented explanations for frontline use. While the literature often focuses on the accuracy and characteristics of an algorithm (technical interpretability), the findings show that for frontline staff, explainability is related to actions. It is about knowing why a certain assessment is suggested, what to do if information is missing, and when it is necessary to pause and escalate. The participants revealed that it is less important 'how the model works' and more important “how the model helps me act safely” at the bedside.

Methodological Contributions: Bridging Theory and Practice

By combining the scoping review with case based discussions and prototype walkthroughs, the thesis operationalises AI supported decision making in the nursing home context. This approach allowed the researcher to uncover workflow and governance needs that are usually missed in purely technical studies. Furthermore, by gathering data across three different nursing homes and multiple professional roles, this study provides a more nuanced, context sensitive look at the thresholds and needs of the Norwegian nursing home sector.

Practical recommendations

Building on the study findings and within the scope of non-urgent symptom management, the following recommendations are proposed for municipalities, nursing home leadership, and technology providers implementing AI/CDSS in Norwegian nursing homes.

1. Prioritise EPJ integration and semantic interoperability from the start.

   Avoid stand alone solutions “in a silo,” as they increase fragmented clinical pictures and the risk of information loss across shifts and roles. A safe solution must “speak” to existing documentation systems and minimise duplicate documentation.

2. Align the support provided with professional roles and clinical responsibility

   Decision support should be tailored to the unique needs and role of the user. HFAs need verification checks and support on how to express concerns. Nurses need prioritised indicators

and clear next steps to manage the situation. Physicians need brief clinical summary of trends, baseline data, and past interventions to enable rapid decision-making.

3. Treat communication support as a fundamental safety feature of the decision support tool
   The weakest link in the escalation chain in many systems is effective communication. A decision support tool should assist in structuring the handover and escalation. ISBAR statements can be a solution to the gaps in information transfer that occur during and between staff.
4. Embed resident baseline data as a safety measure in the decision support tool
   Universal thresholds, which do not take into account the complexities of the patients in a nursing home can lead to unnecessary escalation or missed deteriorating patient. Therefore, the tool should be able to take into account the individual's knowledge of the patient's baseline clinical signs, symptoms and clinical history.
5. Support event-driven use with continuity across time.
   The tool should be usable at the moment a symptom change is noticed during a shift, while also feeding into shift change routines and periodic evaluations, so that "what is happening now and what to do next" remains visible over time.
6. Design for usability under time pressure and cognitive load.
   Access must be fast, navigation simple, and the tool must reduce—not add—workload. If using the tool adds time to current steps, staff will not adopt it. Interface design should assume variable digital confidence and prioritise low cognitive load.
7. Implement safeguards that reinforce the "human filter."
   It should be able to provide a form of secondary opinion that enhances rather than replaces human judgment. It should make uncertainty visible, such as missing data, and provide clear boundaries on when to stop and seek more information or when to proceed.
8. Governance, accountability, and credible endorsement should be established before implementation.
   Well defined and unambiguous protocols should be established to clearly define the clinical responsibility of the nurse (RN) and physician leaders, with the tool being viewed as support for information gathering and communication. The need for credible organizational endorsement and attribution of accountability was highlighted as a key factor.
9. Pilot and evaluate under real-world conditions before implementation.
   Evaluate workflow integration to assess usability, safety, quality of communication, and speed of escalation rather than predictive metrics alone.
10. Extend evidence across contexts to strengthen transferability.

To support broader Norwegian and European guidance, replicate qualitative work in other municipalities and countries and compare findings with Stavanger, as a foundation for context sensitive implementation guidelines for nursing home AI decision support.

Limitations

This study has limitations that should be considered when interpreting the findings. Only one medical doctor (MD) was interviewed, one interview transcript was largely based on notes (RN_P14), and there was no direct observation of real escalation episodes in practice. In addition, no real AI tool was tested, the empirical context was limited to Stavanger, and there is a risk of researcher role bias given the researcher's nursing background. To mitigate this, neutral prompts were used to avoid leading participants, no direct subordinates were interviewed, participants were reminded that the study was not testing medical accuracy, and reactive comments or facial expressions were avoided during data collection. On the other hand, the fact that the researcher has a nursing background served as an "interpretative lens" that allowed the researcher to understand the "nursing language" and subtle cues (like the "gut feeling" mentioned by HFAs) which a non clinical researcher might have missed.

Future scope

Future research should strengthen transferability by replicating the qualitative work in other Norwegian municipalities and in other European countries. This would enable comparison with the Stavanger context and support the development of broader implementation guidance that is sensitive to local context for AI clinical decision support in nursing homes. In addition, the next logical step is a workflow-integrated prototype pilot evaluated in real practice with outcomes focused on usability, safety, impact on the quality of communication, and the speed/quality of escalation, rather than predictive performance alone.

Closing statement

The thesis indicates that in Norwegian nursing homes, the success and safety of AI based CDSS for non-urgent symptom changes depend less on marginal gains in predictive accuracy and more on workflow fit, fast access, role clarity, resident baseline awareness, and clear actionable next steps. A safe system must strengthen communication and continuity across shifts while remaining governed as a "second opinion"—with clinical responsibility and final judgement remaining human, under clear RN/physician leadership and accountability structures.

Bibliographic References and Sources

Abdellatif, A., Bouaud, J., Nghiem, D., Lafuente-Lafuente, C., Belmin, J., & Séroussi, B. (2020). Clinical decision support systems in nursing homes: A scoping review. *Studies in Health Technology and Informatics, 270*, 542–546. https://doi.org/10.3233/SHTI200219

Abdellatif, A., Bouaud, J., Lafuente-Lafuente, C., Belmin, J., & Séroussi, B. (2021). Computerized decision support systems for nursing homes: A scoping review. *Journal of the American Medical Directors Association, 22*(5), 984–994. https://doi.org/10.1016/j.jamda.2021.01.080

Alruwaili, M. M., Al-Otaibi, A. S., & Al-Zahrani, S. M. (2024). Exploring nurses' awareness and attitudes toward artificial intelligence: Implications for nursing practice. *Digital Health, 10*, 1–10. https://doi.org/10.1177/20552076241271803

Andargoli, A. E., Ulapane, N., Nguyen, T. A., Shuakat, N., Zelcer, J., & Wickramasinghe, N. (2024). Intelligent decision support systems for dementia care: A scoping review. *Artificial Intelligence in Medicine, 150*, 102815. https://doi.org/10.1016/j.artmed.2024.102815

Atalla, A. D. G., El-Sayed, S. H., & El-Gazar, H. E. (2025). Embracing artificial intelligence in nursing: Exploring the relationship between artificial intelligence-related attitudes, creative self-efficacy, and clinical reasoning competency among nurses. *BMC Nursing, 24*, 661. https://doi.org/10.1186/s12912-025-03306-3

Atee, M., Hoti, K., Parsons, R., & Hughes, J. D. (2018). A novel pain assessment tool incorporating automated facial analysis: Interrater reliability in advanced dementia. *Clinical Interventions in Aging, 13*, 1245–1258. https://doi.org/10.2147/CIA.S168024

Attride-Stirling, J. (2001). Thematic networks: An analytic tool for qualitative research. *Qualitative Research, 1*(3), 385–405. https://doi.org/10.1177/146879410100100307

Australian Institute of Health and Welfare. (2024, May 22). *People's care needs in aged care*. GEN Aged Care Data. https://www.gen-agedcaredata.gov.au/topics/people-s-care-needs-in-aged-care

Badawi, B., et al. (2025). Multimodal detection of agitation in people with dementia in clinical settings: Observational pilot study. *JMIR Aging, 8*(1), e68156. https://doi.org/10.2196/68156

Badawy, W., & Shaban, M. (2025). Exploring geriatric nurses' perspectives on the adoption of artificial intelligence in elderly care: A qualitative study. *Geriatric Nursing, 61*, 41–49. https://doi.org/10.1016/j.gerinurse.2024.10.078

Bagchus, C., Zee, M. S., van der Steen, J. T., Klapwijk, M. S., Lemos Dekker, N., Onwuteaka-Philipsen, B. D., & Pasman, H. R. W. (2024). Challenges in recognizing and discussing changes in a resident's condition in the palliative phase: Focus group discussions with nursing staff working in nursing homes about their experiences. *BMC Palliative Care, 23*, 144. https://doi.org/10.1186/s12904-024-01479-3

Barker, R. O., Hanratty, B., Kingston, A., Ramsay, S. E., & Matthews, F. E. (2021). Changes in health and functioning of care home residents over two decades: What can we learn from population-based studies? *Age and Ageing, 50*(3), 921–927. https://doi.org/10.1093/ageing/afaa227

Barstad, A., & Hammersland, R. (2025, August 12). *Eldrebefolkningen, nå og i framtiden: Humanitære behov og omsorgsressurser* (Rapporter 2025/22). Statistisk sentralbyrå. https://www.ssb.no/helse/helsetjenester/artikler/eldrebefolkningen-na-og-i-framtiden

Berkhout, M., Smit, K., Sent, D., Kusters, R., Versendaal, J., & van Houwelingen, T. (2025). Understanding the role of clinical decision support systems among hospital nurses using the FITT framework: Qualitative study. *Journal of Medical Internet Research, 27*, e76025. https://doi.org/10.2196/76025

Boyce, R. D., Kravchenko, O. V., Perera, S., Karp, J. F., Kane-Gill, S. L., Reynolds, C. F., Albert, S. M., & Handler, S. M. (2022). Falls prediction using the nursing home minimum dataset. *Journal of the American Medical Informatics Association, 29*(9), 1497–1507. https://doi.org/10.1093/jamia/ocac111

Capodici, A., et al. (2025). A scoping review of machine learning models to predict risk of falls in elders, without using sensor data. *Diagnostic and Prognostic Research, 9*(1), 11. https://doi.org/10.1186/s41512-025-00190-y

Charon, C., et al. (2024). One month prediction of pressure ulcers in nursing home residents with Bayesian networks. *Journal of the American Medical Directors Association, 25*(6), 104945. https://doi.org/10.1016/j.jamda.2024.01.014

Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. M. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *British Journal of Surgery, 102*(3), 148–158. https://doi.org/10.1002/bjs.9736

Cunningham, A., et al. (2025). Prevalence of agitation and depression amongst people with dementia in Australian residential aged-care facilities: A cross-sectional study using LLMs. *Studies in Health Technology and Informatics, 316*, 1255–1264. https://doi.org/10.3233/SHTI251255

Delespierre, T., Denormandie, P., Bar-Hen, A., & Josseran, L. (2017). Empirical advances with text mining of electronic health records. *BMC Medical Informatics and Decision Making, 17*, 84. https://doi.org/10.1186/s12911-017-0519-0

Elhaddad, M., & Hamam, S. (2024). AI-driven clinical decision support systems: An ongoing pursuit of potential. *Cureus, 16*(4), e57728. https://doi.org/10.7759/cureus.57728

Fossum, M., Ehnfors, M., Fruhling, A., & Ehrenberg, A. (2011). An evaluation of the usability of a computerized decision support system for nursing homes. *Applied clinical informatics, 2*(4), 420–436. https://doi.org/10.4338/ACI-2011-07-RA-0043

Garcés-Jiménez, A., Polo-Luque, M.-L., Gómez-Pulido, J. A., Rodríguez-Puyol, D., & Gómez-Pulido, J. M. (2024). Predictive health monitoring: Leveraging artificial intelligence for early detection of infectious diseases in nursing home residents through discontinuous vital signs analysis. *Computers in Biology and Medicine, 174*, 108469. https://doi.org/10.1016/j.compbiomed.2024.108469

Garg, A. X., et al. (2005). Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: A systematic review. *JAMA, 293*(10), 1223–1238. https://doi.org/10.1001/jama.293.10.1223

Hacking, C., Verbeek, H., Hamers, J. P. H., Sion, K., & Aarts, S. (2022). Text mining in long-term care: Exploring the usefulness of artificial intelligence in a nursing home setting. *PLOS ONE, 17*(8), e0268281. https://doi.org/10.1371/journal.pone.0268281

Helsedirektoratet. (2021). *Status, muligheter og behov relatert til kunstig intelligens i kommunal helse- og omsorgstjeneste*. https://www.helsedirektoratet.no/rapporter/status-muligheter-og-behov-relatert-til-kunstig-intelligens-i-kommunal-helse-og-omsorgstjeneste

Helsepersonellkommisjonen. (2023). *Tid for handling: Personellet i en bærekraftig helse- og omsorgstjeneste (NOU 2023: 4)*. Helse- og omsorgsdepartementet. https://www.regjeringen.no/contentassets/337fef958f2148bebd326f0749a1213d/no/pdfs/nou202320230004000dddpdfs.pdf

Hendriks, A., Hacking, C., Verbeek, H., & Aarts, S. (2024). Data science techniques to gain novel insights into quality of care: A scoping review of long-term care for older adults. *Exploration of Digital Health Technologies, 2*, 67–85. https://doi.org/10.37349/edht.2024.00012

Hoogendijk, E. O., et al. (2023). Optimising the care for older persons with complex chronic conditions in home care and nursing homes: Design and protocol of I-CARE4OLD, an observational study using real-world data. *BMJ Open, 13*(6), e072399. https://doi.org/10.1136/bmjopen-2023-072399

Kandel, B., et al. (2025). Development of a predictive hospitalisation model for skilled nursing facility patients. *Journal of the American Medical Directors Association, 26*(1), 105288. https://doi.org/10.1016/j.jamda.2024.105288

Kang, M.-J., et al. (2024). Designing and testing clinical simulations of an early warning system for implementation in acute care settings. *JAMIA Open, 7*(4), ooae092. https://doi.org/10.1093/jamiaopen/ooae092

Kawamoto, K., Houlihan, C. A., Balas, E. A., & Lobach, D. F. (2005). Improving clinical practice using clinical decision support systems: A systematic review of trials to identify features critical to success. *BMJ, 330*(7494), 765. https://doi.org/10.1136/bmj.38398.500764.8F

Kruse, C. S., et al. (2017). Impact of electronic health records on long-term care facilities: Systematic review. *JMIR Medical Informatics, 5*(3), e35. https://doi.org/10.2196/medinform.7958

Lapp, L., Egan, K. J., McCann, L., Mackenzie, M., Wales, A., & Maguire, R. (2022). Decision support tools in adult long-term care facilities: Scoping review. *Journal of Medical Internet Research, 24*(9), e39681. https://doi.org/10.2196/39681

Lee, S. K., Shin, J. H., Ahn, J., Lee, J. Y., & Jang, D. E. (2020). Application of machine learning methods in nursing home research. *International Journal of Environmental Research and Public Health, 17*(17), 6234. https://doi.org/10.3390/ijerph17176234

Lee, S. K., Shin, J. H., Ahn, J., Lee, J. Y., & Jang, D. E. (2021). Identifying the risk factors associated with nursing home residents' pressure ulcers using machine learning methods. *International Journal of Environmental Research and Public Health, 18*(6), 2954. https://doi.org/10.3390/ijerph18062954

Liberati, E. G., et al. (2017). What hinders the uptake of computerized decision support systems in hospitals? A qualitative study and framework for implementation. *Implementation Science, 12*(1), 113. https://doi.org/10.1186/s13012-017-0644-2

Linqvist Leonardsen, A.-C., Högstedt, Å., Lundgren, P., & Hardeland, C. (2025, April 4). Norge og Sverige bruker ulike beslutningsstøtteverktøy i kommunehelsetjenesten. *Sykepleien, 113*, e-98564. https://doi.org/10.4220/Sykepleiens.2025.98564

Lukkien, D. R. M., et al. (2024). AI-assisted decision-making in long-term care: Qualitative study on prerequisites for responsible innovation. *JMIR Nursing, 7*, e55962. https://doi.org/10.2196/55962

Masot, O., et al. (2022). Decision support-tools for early detection of infection in older people (aged >65 years): A scoping review. *BMC Geriatrics, 22*, 552. https://doi.org/10.1186/s12877-022-03218-w

McDonald, E. G., et al. (2025). Electronic decision support for deprescribing in older adults living in long-term care: A stepped-wedge cluster randomized trial. *JAMA Network Open, 8*(5), e2512931. https://doi.org/10.1001/jamanetworkopen.2025.12931

Meunier, P.-Y., Raynaud, C., Guimarães, E., Gueyffier, F., & Letrilliart, L. (2023). Barriers and facilitators to the use of clinical decision support systems in primary care: A mixed-methods systematic review. *Annals of Family Medicine, 21*(1), 57–69. https://doi.org/10.1370/afm.2908

Moon, K. J., Son, C.-S., Lee, J.-H., & Park, M. (2022). The development of a web-based app employing machine learning for delirium prevention in long-term care facilities in South Korea. *BMC Medical Informatics and Decision Making, 22*, 220. https://doi.org/10.1186/s12911-022-01966-8

Nanevski, I., et al. (2025). The potential of AI in nursing care: Multicenter evaluation in fall risk assessment. *Journal of Medical Internet Research, 27*, e71034. https://doi.org/10.2196/71034

National Academies of Sciences, Engineering, and Medicine. (2022). *The national imperative to improve nursing home quality: Honoring our commitment to residents, families, and staff*. National Academies Press. https://doi.org/10.17226/26526

Olakotan, O., Samuriwo, R., Ismaila, H., & Atiku, S. (2025). Usability challenges in electronic health records: Impact on documentation burden and clinical workflow: A scoping review. *Journal of Evaluation in Clinical Practice, 31*(4), e70189. https://doi.org/10.1111/jep.70189

Porcellato, E., et al. (2025). Exploring applications of artificial intelligence in critical care nursing: A systematic review. *Nursing Reports, 15*(2), 55. https://doi.org/10.3390/nursrep15020055

Rony, M. K. K., et al. (2025). The role of artificial intelligence in nursing care: An umbrella review. *Nursing Inquiry, 32*(2), e70023. https://doi.org/10.1111/nin.70023

Shahidi, F., et al. (2023). Machine learning risk estimation and prediction of death in continuing care facilities using administrative data. *Scientific Reports, 13*, 17708. https://doi.org/10.1038/s41598-023-43943-9

Shao, L., et al. (2024). Development and external validation of a machine learning based fall prediction model for nursing home residents: A prospective cohort study. *Journal of the American Medical Directors Association, 25*(9), 105169. https://doi.org/10.1016/j.jamda.2024.105169

Skuban-Eiseler, T., Orzechowski, M., Denkinger, M., Kocar, T. D., Leinert, C., & Steger, F. (2023). Artificial intelligence–based clinical decision support systems in geriatrics: An ethical analysis. *Journal of the American Medical Directors Association, 24*(11), 1271–1276. https://doi.org/10.1016/j.jamda.2023.06.008

Solaiman, B. (2024). Legal and ethical considerations of artificial intelligence for residents in post-acute and long-term care. *Journal of the American Medical Directors Association, 25*, 105105. https://doi.org/10.1016/j.jamda.2024.105105

Stancu, A., Roșca, C.-M., & Iovanovici, E. M. (2025). Applications of machine learning algorithms in geriatrics. *Applied Sciences, 15*(15), 8699. https://doi.org/10.3390/app15158699

Standing, M. (2008). Clinical judgement and decision-making in nursing: Nine modes of practice in a revised cognitive continuum. *Journal of Advanced Nursing, 62*(1), 124–134. https://doi.org/10.1111/j.1365-2648.2007.04583.x

Sun, C. L. F., et al. (2020). Predicting coronavirus disease 2019 infection risk and related risk drivers in nursing homes: A machine learning approach. *Journal of the American Medical Directors Association, 21*(11), 1533–1538.e6. https://doi.org/10.1016/j.jamda.2020.08.030

Susnjak, T., Ganeshan, K., & Barczak, A. L. (2025). Towards clinical prediction with transparency: An explainable AI approach to survival modelling in residential aged care. *Computer Methods and Programs in Biomedicine, 263*, 108653. https://doi.org/10.1016/j.cmpb.2025.108653

Suzuki, M., et al. (2020). Deep-learning prediction of falls among nursing home residents with Alzheimer's disease. *Geriatrics & Gerontology International, 20*(9), 838–844. https://doi.org/10.1111/ggi.13920

Takeshita, Y., et al. (2025). Machine learning prediction for postdischarge falls in older adults: A multicenter prospective study. *Journal of the American Medical Directors Association, 26*(2), 105414. https://doi.org/10.1016/j.jamda.2024.105414

Thapa, R., et al. (2022). Predicting falls in long term care facilities: Machine learning study. *JMIR Aging, 5*(2), e35373. https://doi.org/10.2196/35373

Thompson, C., et al. (2021). Effects of computerised clinical decision support systems on nursing and allied health professional performance and patient outcomes: A systematic review of experimental and observational studies. *BMJ Open, 11*(11), e053886. https://doi.org/10.1136/bmjopen-2021-053886

Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care, 19*(6), 349–357. https://doi.org/10.1093/intqhc/mzm042

Ventura Silva, J., et al. (2024). Artificial intelligence in the organization of nursing care: A scoping review. *Nursing Reports, 14*, 2733–2745. https://doi.org/10.3390/nursrep14040202

Vithanage, D., Yu, P., Xie, Q., Xu, H., Wang, L., & Deng, C. (2025). A comprehensive evaluation of large language models for information extraction from unstructured electronic health records in

residential aged care. *Computers in biology and medicine, 197*(Pt A), 111013. https://doi.org/10.1016/j.compbiomed.2025.111013

Woo, K., et al. (2021). Identifying urinary tract infection–related information in home care nursing notes. *Journal of the American Medical Directors Association, 22*(5), 1015–1021.e2. https://doi.org/10.1016/j.jamda.2020.12.010

Yesankar, P., Puri, C., Bhagat, M., & Wankhade, G. (2025, March). AI-powered clinical decision support systems (CDSS): Challenges, benefits, applications and future directions [Paper presentation]. *2025 International Conference on Machine Learning and Autonomous Systems (ICMLAS)*. https://doi.org/10.1109/ICMLAS64557.2025.10969014

Zhang, M., et al. (2025). Leveraging retrieval-augmented generation-driven large language models to extract dementia agitation symptoms and triggers from free-text nursing notes. *Studies in Health Technology and Informatics, 316*, 950–959. https://doi.org/10.3233/SHTI250950

Zhu, Y., Song, T., Zhang, Z., Deng, C., Alkhalaf, M., Li, W., … Yu, P. (2022). Agitation prevalence in people with dementia in Australian residential aged care facilities: Findings from machine learning of electronic health records. *Journal of Gerontological Nursing, 48*(4), 57–64. https://doi.org/10.3928/00989134-20220309-01

Annexes

Annex A – Study Sourcing Log

Title	Database / Source	Platform/Interface	Search Category	Search Strategy / String	Snowballing / Hand-search Notes	Filters / Limits Applied	Date Searched (YYYY-MM-DD)[7]	Results Returned (n)	File / Export Name	Why adapted	Comments / Adjustments
A comprehensive evaluation of large language models for information extraction from unstructured electronic health records in residential aged care	PubMed	Web browser (Google → ResearchGate → PubMed → ScienceDirect → Elsevier PDF)	Targeted / known-item search	"Vithanage S. et al. Large language model–based info extraction on residential aged-care EHR. Computers in Biology & Medicine, 2025. ResearchGate" then exact title "A comprehensive evaluation of large language models for information extraction from unstructured electronic health records in residential aged care"	Could not access full text via ResearchGate, so repeated search with exact title, opened PubMed, then Elsevier full article on ScienceDirect and viewed/downloaded PDF	None (no filters, general web search)	2025-09 to 2025-12	1	20_Vithanage2025_LLM_IE_RAC_EHR.pdf	Direct match from AI-suggested reference list; not captured in database export	Added deliberately because it evaluates LLMs on RAC EHR, relevant as backend data-extraction layer for symptom CDS
A novel pain assessment tool incorporating automated facial analysis interrater reliability in advanced dementia	PubMed Central	Web browser Google to PMC	Targeted / known-item search	Automated facial analysis ePAT PainChek for pain in aged care residents with dementia Clinical Interventions in Aging 2018 PMC	Clicked first Google result PMC article then used PDF link	None	2025-09 to 2025-12	1	26_Atee2018_PainChek_AutomatedFacialAnalysis.pdf	From AI generated recommendation list not from original database search	Primarily psychometric reliability study may be borderline for inclusion but retained as example of AI assisted pain assessment
A scoping review of machine learning models to predict risk of falls in elders, without using sensor data	PubMed	Web browser: Google → PubMed → Free full text PMC → journal PDF	Targeted / known-item search	Scoping review: ML models to predict falls from EHR (no sensors) npj Digital Medicine 2025; then refined to full title string from PubMed	Clicked first PubMed hit, followed Free full text PMC link, then opened journal site to download full PDF.	None	2025-09 to 2025-12	1	39_Capodici2025_ScopingReview_ML_Falls_NoSensors.pdf	Included as cross-setting background on ML falls-risk modelling and reporting standards.	Used mainly for feature ideas and methodological critique rather than LTC-specific data.
Agitation prevalence in dementia in residential aged-care facilities using AI/NLP on nursing notes	PubMed	Web browser (Google → PubMed → publisher full-text page)	Targeted / known-item search	"Zhu Y. et al. Agitation prevalence in dementia in RACFs using AI/NLP on nursing notes. Journal of Gerontological Nursing, 2022. PubMed"	From Google was opened the PubMed record, then the publisher's full-text page, but there was no access via the affiliated universities, so only the abstract could be used	None (no filters, general web search)	2025-09 to 2025-12	1	25_Zhu2022_JGN_Agitation_Prevalence_NLP_abstractOnly.pdf	Added as AI/NLP-based agitation prevalence study to contextualise symptom burden, even though only abstract was accessible	Also marked as abstract-only; limited methodological detail means it will have lower weight in the synthesis
AI Powered Clinical Decision Support Systems CDSS Challenges Benefits Applications and Future Directions	IEEE Xplore	ISCTE library search leading to Google Scholar then IEEE Xplore	Targeted / known-item search	gayot s et al AI CDSS 2025 then AI-Powered Clinical Decision Support Systems CDSS Challenges Benefits Applications and Future Directions	Followed chain ISCTE library to Google Scholar to IEEE Xplore and downloaded PDF	None	2025-09 to 2025-12	1	33_Yesankar2025_AIPowered_CDSS_Review.pdf	Added from LTC ethics readiness and organisation list as general CDSS background	Broad CDSS overview across healthcare not LTC specific but useful for generic design and governance principles
AI-assisted decision-making in long-term care: Qualitative study on prerequisites for responsible innovation	Google Scholar	Google	Bibliographic search (Google Scholar)	"ai for decision making in nursing homes norway"	Direct title match from GS	from 2018	2025-10-08	18400	Yesankar_2025_AIPowered_Clinical_Decision_Support_Systems_CDSS_Challenges_Benefits_Applications_and_Future_Directions_ICMLAS.PDF		Target: first 100; logging incrementally

[7] Where exact search dates or export filenames were not available, the log records the best available retrieval window (Sep–Dec 2025) and the current stored filename to ensure traceability.

Title	Database / Source	Platform/Interface	Search Category	Search Strategy / String	Snowballing / Hand-search Notes	Filters / Limits Applied	Date Searched (YYYY-MM-DD)	Results Returned (n)	File / Export Name	Why adapted	Comments / Adjustments
AI-Driven Clinical Decision Support Systems: An Ongoing Pursuit of Potential (Elhaddad & Hamam, 2024)	PubMed	Web	Database search	Elhaddad, M., et al. (2024). AI Driven Clinical Decision Support Systems: A Review. Journal of Personalized Medicine, 14(1), 29 typed on google and the chose the first result	Followed publisher link (Cureus) from PubMed; verified metadata; downloaded PDF	Publication date: 2023–2025; Humans; English; Article / PubMed current	2025-11-18	1	Elhaddad_2024_Cureus_AI-Driven-CDSS.pdf	Started from full citation via Google → PubMed to confirm OA; focused terms on title/author to avoid false positives	Used this as foundational review for governance/XAI/workflow principles
An Evaluation of the Usability of a Computerized Decision Support System for Nursing Homes	Google Scholar	NR	Bibliographic search (Google Scholar)	"decision making nursing homes norway"	NR	NR (not recorded)	2025-10-09		An Evaluation of the usability of a computerized decision support system for nursing homes	NR	
Application of machine learning methods in nursing home research	Publisher website (MDPI)	Web browser Google to MDPI	Targeted / known-item search	Application of ML methods in nursing home research IJERPH 2020 PMC	Clicked first Google result leading directly to IJERPH article and downloaded PDF	None	2025-09 to 2025-12	1	27_Lee2020_IJERPH_ML_in_NH_research.pdf	Chosen based on AI recommendation to capture facility level ML falls model	Broad methods paper some parts not directly about clinical decision support but useful for organisational risk context
Applications of Machine Learning Algorithms in Geriatrics	Google Scholar	Web browser	Bibliographic search (Google Scholar)	ML in geriatrics applications and directions	Google Scholar search; second result (MDPI Applied Sciences review) chosen as it matched description	None	2026-01-04		Applications of Machine Learning Algorithms in Geriatrics.pdf	Broad geriatrics ML background review	Gives general landscape of ML in older adults and supports background section on AI in geriatrics beyond LTC
Artificial Intelligence in the Organization of Nursing Care A Scoping Review	PubMed	Web browser Google to PubMed to PMC to journal site	Targeted / known-item search	How nurses use work with AI scoping review Nursing Reports 2024 Portugal authorship PubMed ventura silva	Clicked first PubMed result then Free full text PMC then journal site PDF	None	2025-09 to 2025-12	1	30_VenturaSilva2024_AI_Organization_NursingCare.pdf	Added from Norway and Portugal context list as background on AI in nursing care organisation	Broad scoping review across settings not specific to nursing homes but highly relevant for conceptual frame and design requirements
Artificial Intelligence–Based Clinical Decision Support Systems in Geriatrics: An Ethical Analysis	PubMed	Google	Citation chasing / hand-search	Skuban-Eiseler, T., et al. (2023). Artificial Intelligence–Based Clinical Decision Support Systems: Ethical Implications in Long-Term Care. JAMDA. (led to PubMed record and open-access PDF)	Reviewed methods (five-step process) and conclusions on autonomy/beneficence/nonmaleficence/justice	none	2025-11-09	1	google_2025-11-09_skuban-eiseler2023_jamda_ethics.bib		Ethical analysis (not metrics study); suitable to cite in thesis “safety/ethics” section for LTC AI-CDSS
Data science techniques to gain novel insights into quality of care: a scoping review of long-term care for older adults	Google Scholar	Web browser	Bibliographic search (Google Scholar)	Hendriks A. et al. Digital health in LTC for older adults	Targeted Google Scholar search; first Exploration of Digital Health Technologies result used	None	2026-01-04		Data science techniques to gain novel insights into quality of care_a scoping review of long term care for older adults.pdf	Scoping review on data science and AI in LTC	Selected to map existing analytics uses in LTC and identify gaps relevant to my thesis
Decision support-tools for early detection of infection in older people (aged >65 years): a scoping review	Google Scholar	Web	Database search	Masot et al 2022, Decision support-tools for early detection of infection in older people	Followed citations within the review to identify NH-relevant tools	Since 2018; Review articles	2025-11-09	10	Masot_2022_BMCGeriatrics_infection_DST_scoping_review.pdf		Use for infection cue standardisation and action ladder in LTC

Title	Database / Source	Platform/Interface	Search Category	Search Strategy / String	Snowballing / Hand-search Notes	Filters / Limits Applied	Date Searched (YYYY-MM-DD)	Results Returned (n)	File / Export Name	Why adapted	Comments / Adjustments
Deep-learning prediction of falls among nursing home residents with Alzheimer's disease	Wiley Online Library	Web browser (Google → Wiley abstract → ISCTE library search)	Targeted / known-item search	"Suzuki M. et al. Deep-learning prediction of falls among nursing home residents with Alzheimer's disease. Geriatrics & Gerontology International, 2020. PubMed"	From Google was opened the Wiley abstract; then searched the same title in the ISCTE library but still no access to the full PDF	None (no filters; Google and local library search)	2025-09 to 2025-12	1	24_Suzuki2020_GGI_DL_Falls_AD_abstractOnly.pdf	Included despite abstract-only access because it is one of few DL falls models in NH residents with Alzheimer's	Marked as abstract-only; methods and effect sizes must be interpreted with caution
Designing and testing clinical simulations of an early warning system for implementation in acute care settings	PubMed	Web browser: Google → CONCERN project site → copied citation → Google → PubMed → JAMIA Open full text	Targeted / known-item search	ARTICLE CONCERN clinical decision support using nurse-entered narratives early warning kang 2023; then pasted full citation from CONCERN site	Clicked CONCERN study website, copied Kang et al. citation, searched in Google/PubMed, then opened JAMIA Open and downloaded PDF.	None	2025-09 to 2025-12	1	38_Kang2024_JAMIAOpen_CONCERN_simulation.pdf	Included as methodological example of simulation-based testing of nurse-facing early-warning CDS.	Not LTC but highly relevant for CDS testing methods and human-in-the-loop design.
Development and external validation of a machine learning-based fall prediction model for nursing home residents: A prospective cohort study	PubMed	Web browser – Google → PubMed record (no open PDF)	Citation chasing / hand-search	"Shao L. et al. 'Development & external validation of an ML fall-risk model for nursing home residents.' JAMDA, 2024."	From the AI-generated list; I pasted the string into Google, opened the first PubMed result, but could not access the full PDF (paywalled).	None (no filters; direct Google search)	2025-09 to 2025-12	1	Shao_2024_JAMDA_Falls_Prospective_abstract Only	Kept because it is a prospective, externally validated nursing-home falls model; data extracted from abstract/methods only.	Recorded as abstract-only inclusion, obtained via targeted Google → PubMed lookup, not via database export.
Development of a predictive hospitalization model for skilled nursing facility patients	PubMed	Web browser – Google → PubMed record → Elsevier Open Access → Download PDF	Citation chasing / hand-search	"Kandel A. et al. 'Predicting 7-day hospitalization or death in skilled nursing facilities.' JAMDA, 2025."	From the AI-generated list; I pasted the citation string into Google, opened the first PubMed result and then the Elsevier Open Access full-text PDF.	None (no filters; direct Google search)	2025-09 to 2025-12	1	Kandel_2025_JAMDA_Hosp7d_SNF.pdf	Included as a major deterioration/hospitalisation risk model for skilled nursing facility patients.	Outside the formal database search; added through targeted Google → PubMed → OA retrieval pathway.
Electronic decision support tool for deprescribing in older adults in long-term care	PubMed	Web	Database search	McDonald EG. et al, 2025, Impact of an electronic decision support tool on deprescribing in older adults in long-term care	Checked trial registry and related deprescribing CDSS papers; none with LTC stepped-wedge design matched better	2023–2025; English; Humans; Clinical Trial	2025-11-18	20	McDonald_2025_JAMA_Netw_Open_deprescribing_LTC_stepped-wedge.pdf	Initial results mixed settings → added "long-term care" and Clinical Trial filter to isolate LTC RCT/SW-CRT evidence	Use to justify medication-review card + measurable deprescribing outcomes in prototype
Embracing artificial intelligence in nursing: Exploring the relationship between artificial intelligence-related attitudes, creative self-efficacy, and clinical reasoning competency among nurses	SpringerLink / BMC Nursing	Web browser (Google Scholar → BMC Nursing full text)	Bibliographic search (Google Scholar)	Atalla ADG et al. AI attitudes & clinical reasoning competency among nurses BMC Nursing 2025	Clicked first Google Scholar result then journal full-text page and downloaded PDF.	None	2025-09 to 2025-12	1	36_Atalla2025_BMCNursing_AI_attitudes_clinical_reasoning.pdf	Added from nurse experience / capability building list.	Correlational study on attitudes and competencies in ICU nurses, not LTC but highly relevant for training design.
Empirical advances with text mining of electronic health records	Google Scholar	Web browser	Bibliographic search (Google Scholar)	Delespierre T. et al. Empirical advances with text mining of EHRs (includes NH-related use). BMC Med Inform Decis Mak 2017	Added from targeted Google Scholar search after AI recommendations on LTC text mining	None	2026-01-04		Empirical advances with text mining of.pdf	Methods paper on LTC EHR text mining to inform free text handling	Used for background and design requirements, not as effectiveness evidence
Exploring applications of artificial intelligence in critical care nursing: A systematic review	Publisher website (MDPI)	Web browser (Google Scholar → MDPI → Download PDF)	Bibliographic search (Google Scholar)	Applications of AI in critical care nursing systematic review	No extra snowballing recorded beyond references in the review.	None	2025-09 to 2025-12	1	37_Porcellato2025_NursingReports_AI_critical_care_nursing_SR.pdf	Added as methods/background paper on AI & nursing from nurse capability/experience theme.	ICU-focused; used mainly for evaluation metrics and design lessons rather than LTC-specific content.

Title	Database / Source	Platform/Interface	Search Category	Search Strategy / String	Snowballing / Hand-search Notes	Filters / Limits Applied	Date Searched (YYYY-MM-DD)	Results Returned (n)	File / Export Name	Why adapted	Comments / Adjustments
Exploring geriatric nurses' perspectives on the adoption of artificial intelligence in elderly care: A qualitative study	ScienceDirect	Web browser (Google Scholar → ScienceDirect)	Bibliographic search (Google Scholar)	"Geriatric nurses' perspectives on AI adoption in elderly care"	No additional snowballing yet; full text to be retrieved later via library if possible.	None	2025-09 to 2025-12	1	35_Badawy2025_GeriatricNursing_AI_adoption_elderlycare_abstractOnly.pdf	Included from nurse experience category despite only abstract access to capture adoption themes.	Flagged as abstract-only; methods and results to be updated once full text is obtained.
Exploring nurses' awareness and attitudes toward artificial intelligence: Implications for nursing practice	PubMed	Web browser (Google → PubMed → PMC → journal PDF)	Targeted / known-item search	Nurses' awareness & attitudes toward AI tools Digital Health 2024 PMC	Clicked first Google/PubMed result then used Free full text (PMC) and PDF link.	None	2025-09 to 2025-12	1	34_Alruwaili2024_DigitalHealth_NurseAttitudesAI.pdf	Added from nurse experience / adoption list (category B).	Survey on general AI attitudes, not LTC-specific but conceptually relevant.
Falls prediction using the nursing home minimum dataset.	PubMed	Web browser (Google search → PMC article page → "PDF").	Citation chasing / hand-search	Boyce RD. et al. "Falls prediction using the nursing home Minimum Data Set (MDS)." JAMIA, 2022. PMC	Article was part of the "direct matches" Word document list generated by ChatGPT; I copied the suggested citation string into Google, clicked the first result (PMC article "Falls prediction using the nursing home minimum dataset"), then clicked "PDF" to download the full article.	None (general Google search; no filters).	2025-11-23		Boyce_2022_JAMIA_FallsPrediction_MDS.pdf	Not found via structured database search but deliberately added through targeted search because it is an NH-specific AI/ML risk prediction model highly relevant to my thesis.	Included as a key example of a nursing-home fall risk prediction model using routinely collected assessment and medication data (MDS + drugs), informing design of risk modules in my prototype.
Identifying the risk factors associated with nursing home residents pressure ulcers using machine learning methods	Publisher website (MDPI)	Web browser Google to MDPI	Targeted / known-item search	Lee S K et al ML to identify risk factors for pressure ulcers in nursing home residents IJERPH 2021 MDPI	Clicked first Google result to MDPI article then downloaded PDF	None	2025-09 to 2025-12	1	28_Lee2021_IJERPH_ML_PressureUlcers.pdf	Selected from AI recommendation to represent facility level ML pressure ulcer risk model	Pairs with the 2020 Lee falls paper to show organisational risk modelling for key geriatric complications
Identifying urinary tract infection–related information in home care nursing notes	PubMed	Google	Citation chasing / hand-search	Woo, K., et al. (2021). Identifying Urinary Tract Infection-Related Information in Nursing Notes Using NLP. JAMDA	Opened PubMed, then full text/PMC PDF	none	2025-11-07	1	google_2025-11-07_woo2021_jamda.bib		Targeted retrieval; home-care setting (not NH).
Impact of electronic health records on long-term care facilities: Systematic review	Google Scholar	Web browser	Bibliographic search (Google Scholar)	Impact of EHRs on Long term care facilities	Direct Google Scholar search then first JMIR Med Inform hit opened	None	2026-01-04		Impact of Electronic Health Records on Long Term Care Facilities.pdf	Context paper on EHR adoption in LTC	Provides organisational and workflow context for any AI or CDS implementation in Norwegian nursing homes
Intelligent decision support systems for dementia care: A scoping review	PubMed	Google	Citation chasing / hand-search	Andargoli, A. E., et al. (2024). Intelligent decision support systems for dementia care: A scoping review. International Journal of Medical Informatics (led to PubMed/Elsevier; actual journal is Artificial Intelligence in Medicine)	Verified bibliographic details on PubMed; Elsevier link confirmed article/DOI	none	2025-11-08	1	google_2025-11-08_andargoli2024_artmed.bib		User's initial journal guess (IJMI) corrected to Artificial Intelligence in Medicine
Legal and Ethical Considerations of Artificial Intelligence for Residents in Post Acute and Long Term Care	PubMed	Web browser Google to PubMed to Elsevier open access	Targeted / known-item search	Legal & ethical considerations of AI for residents in post acute & LTC JAMDA 2024 Pepper et al	Clicked first PubMed result then Elsevier open access then downloaded PDF	None	2025-09 to 2025-12	1	31_Solaiman2024_LegalEthical_AI_PALTC.pdf	Added from LTC ethics readiness and organisation list	Conceptual legal ethics paper no empirical data but directly addresses PA LTC AI issues

Title	Database / Source	Platform/Interface	Search Category	Search Strategy / String	Snowballing / Hand-search Notes	Filters / Limits Applied	Date Searched (YYYY-MM-DD)	Results Returned (n)	File / Export Name	Why adapted	Comments / Adjustments
Leveraging retrieval-augmented generation-driven large language models to extract dementia agitation symptoms and triggers from free-text nursing notes	PubMed	Web browser (Google → PubMed → IOS Press → Download PDF)	Targeted / known-item search	"Zhang M. et al. RAG-based detection of agitation triggers from nursing notes in RACFs. (peer-reviewed 2025). PubMed"	From Google opened the first PubMed result, then clicked "IOS Press full text" to reach the ebook page and download the PDF	None (no filters, general web search)	2025-09 to 2025-12	1	21_Zhang2025_RAG_LLM_AgitationTriggers.pdf	Direct match from AI-suggested list; outside original database strategy	Included because it illustrates RAG + LLM pipeline to extract agitation symptoms and triggers from RACF nursing notes
Machine Learning Prediction for Postdischarge Falls in Older Adults: A Multicenter Prospective Study	PubMed	Web browser: Google → PubMed → Elsevier JAMDA full text → Download PDF	Targeted / known-item search	ML prediction of early post-discharge falls older adults JAMDA 2025; then used full article title from PubMed	Clicked first PubMed result, followed Elsevier open access link to JAMDA, then downloaded PDF.	None	2025-09 to 2025-12	1	40_Takeshita2025_JAMDA_PostdischargeFalls_ML.pdf	Included as hospital-to-community transitions paper relevant to falls-risk modelling.	Prospective but not LTC; used primarily for predictor selection and performance benchmarking.
Machine learning risk estimation and prediction of death in continuing care facilities using administrative data	PubMed	Google	Citation chasing / hand-search	Shahidi 2023 Machine learning risk estimation and prediction of death in continuing care residents Scientific Reports	Verified model table (Table 4) and aORs; noted lack of calibration.	none	2025-11-08	1	google_2025-11-08_shahidi2023_sci-rep.bib		Article number 17708 confirmed in PDF; DOI matches.
Multimodal detection of agitation in people with dementia in clinical settings: Observational pilot study	Publisher website (JMIR)	Web browser (Google → JMIR Aging → Download PDF)	Targeted / known-item search	"Badawi B. et al. Multimodal (wearable + behavioral) AI detection of agitation in dementia. JMIR Aging, 2025. aging.jmir.org"	Clicked first JMIR Aging result from Google, then used "Download PDF" to get the full article	None (no filters, general web search)	2025-09 to 2025-12	1	23_Badawi2025_JMIR_Multimodal_Agitation.pdf	Directly added from AI-suggested list to capture multimodal (wearable + behavioural) agitation detection	Relevant for understanding sensor-based symptom monitoring and its feasibility in dementia care
One-month prediction of pressure ulcers in nursing home residents with Bayesian networks	PubMed	Web browser – Google → PubMed record → Elsevier Open Access → Download PDF	Citation chasing / hand-search	"Charon C. et al. 'One-month prediction of pressure ulcers in nursing homes: ML study.' 2024. PubMed"	Article from AI-generated list; I pasted the string into Google, opened the first PubMed result, then used the Elsevier Open Access link to get the full-text PDF.	None (no filters; direct Google search)	2025-09 to 2025-12	1	Charon_2024_JAMDA_PressureUlcers_Bayesian.pdf	Included as a key pressure-ulcer risk prediction study in nursing homes using explainable Bayesian networks.	Not found in structured database runs; deliberately added via targeted Google → PubMed → OA workflow.
Predicting Coronavirus Disease 2019 Infection Risk and Related Risk Drivers in Nursing Homes: A Machine Learning Approach	Google Scholar	Google	Citation chasing / hand-search	Predicting COVID 19 Infection Risk and Outbreaks in Nursing Homes Using Machine Learning. JAMDA.	Opened PubMed record from GS; clicked "See full article" to PMC PDF	none	2025-11-04	1270	Sun_2020_Predicting coronavirus disease 2019 infection risk and related risk drivers in nursing homes_ Journal of the American Medical Directors Association.PDF		Targeted retrieval via GS → PubMed → PMC full text
Predicting falls in long-term care facilities: Machine learning study	Publisher website (JMIR)	Web browser – Google → JMIR Aging article page → Download PDF	Citation chasing / hand-search	"Thapa R. et al. 'Predicting falls in long-term care facilities: ML using EHR.' JMIR Aging, 2022. PMC"	Article was part of the AI-generated 'direct matches' list; I pasted the suggested string into Google, clicked the first JMIR Aging result and downloaded the PDF.	None (no filters; direct Google search)	2025-09 to 2025-12	1	Thapa_2022_JMIR_Falls_LTC.pdf	Included as a key long-term care falls prediction model using routine EHR data, obtained outside the formal database search.	Not identified through structured database searching; deliberately added via targeted lookup from the AI-generated list.

Title	Database / Source	Platform/Interface	Search Category	Search Strategy / String	Snowballing / Hand-search Notes	Filters / Limits Applied	Date Searched (YYYY-MM-DD)	Results Returned (n)	File / Export Name	Why adapted	Comments / Adjustments
Predictive health monitoring: Leveraging artificial intelligence for early detection of infectious diseases in nursing home residents through discontinuous vital signs analysis	NR	NR	Citation chasing / hand-search	NR	NR	NR	2025-09 to 2025-12		Garces_2024_Predictive health monitoring_Leveraging artificial intelligence for early detection of infectious diseases in nursing home residents through discontinuous vital signs analysis_Computers in Biology and Medicine.pdf	NR	NR
Prevalence of agitation and depression amongst people with dementia in Australian residential aged-care facilities: A cross-sectional study using LLMs	PubMed	Web browser (Google → PubMed → IOS Press → Download PDF)	Targeted / known-item search	First: "Cunningham A. et al. Prevalence of agitation & depression in RACFs via LLM-NLP. 2025. PubMed" (0 results), then: "Prevalence of agitation & depression in RACFs via LLM-NLP 2025 article"	Initial PubMed search with full suggested citation failed; using only the title found the correct PubMed record, then IOS Press full text and PDF	None (no filters, general web search)	2025-09 to 2025-12	1	22_Cunningham2025_LLM_Agitation_Depression_Prevalence.pdf	Added because it uses LLMs on nursing notes to estimate agitation and depression prevalence in RACFs	Complements decision-support papers with population-level symptom burden data from AI/NLP
Text mining in long-term care Exploring the usefulness of artificial intelligence in a nursing home setting	PLOS ONE journal website	Web browser Google to PLOS ONE	Targeted / known-item search	Hacking C et al Text mining nursing-home interviews to support quality-of-care insights AI used on LTC data PLOS ONE 2022	Clicked first Google result leading to PLOS ONE article page and accessed full text PDF	None	2025-09 to 2025-12	1	29_Hacking2022_TextMining_LongTermCare.pdf	Added from combined AI recommendation list not from core database searches	Included as an example of AI text mining on nursing home narrative data to support quality-of-care insight rather than symptom-level decision support
The development of a web-based app employing machine learning for delirium prevention in long-term care facilities in South Korea.	BMC Medical Informatics and Decision Making	Web browser (Google search → BMC article page → "Download PDF").	Citation chasing / hand-search	Moon G. et al. "Development of a web-based app using ML to predict delirium in long-term care facilities." BMC Medical Informatics & Decision Making, 2022. jamda.com	Article was part of the "direct matches" list produced by ChatGPT; I copied the suggested citation string into Google, clicked the first result (BMC Medical Informatics and Decision Making article page) and then clicked "Download PDF".	None (general Google search; no database filters). / Not applicable (Google / journal site).	2025-11-23		Moon_2022_BMC_Delirium_LTCF.pdf	Not from a structured database search but deliberately added via targeted search based on the AI-generated list of LTC nurse decision-support studies.	Included as a key example of a nurse-facing delirium CDSS in long-term care, combining ML-based prediction, assessment and interventions, directly relevant to my thesis scope.
The Potential of AI in Nursing Care: Multicenter Evaluation in Fall Risk Assessment	Google Scholar	Web browser	Bibliographic search (Google Scholar)	Nanevski I. et al. Multicenter evaluation potential of AI in nursing care falls case JMIR 2025	Directly from Google Scholar; only JMIR result for this title	None	2026-01-04		The Potential of AI in Nursing Care_Multicenter Evaluation in Fall Risk Assessment.pdf	Multicentre AI fall risk model and fairness analysis	Used to derive requirements around external validation, transportability and fairness for fall risk AI
Towards clinical prediction with transparency: An explainable AI approach to survival modelling in residential aged care	PubMed	Web browser – Google → PubMed record → View PDF	Citation chasing / hand-search	"Susnjak T., Griffin E. 'Explainable-AI survival modelling in residential aged care.' Computer Methods & Programs in Biomedicine, 2025."	From the AI-generated list; I pasted the suggested string into Google, opened the first PubMed result, then used the 'View PDF' link to access the full article.	None (no filters; direct Google search)	2025-09 to 2025-12	1	Susnjak_2025_CMPB_Survival_XAI_RAC.pdf	Included as a core reference on explainable survival modelling in residential aged care.	Added via targeted Google → PubMed → open-PDF path rather than database search export.
Unobtrusive nighttime movement monitoring to support nursing home continence care: algorithm development & validation	Google Scholar	Google	Other	Unobtrusive nighttime movement monitoring continence care JMIR Nursing 2024	Direct GS → JMIR full-text; noted related bed-sensor papers (Gong; T'Jonck).	None	2025-11-05	1	gscholar_<2025-11-05>_strauven2024_jmir nursing.bib		Targeted retrieval; supplementary source (GS)

Annex B – Included studies table (n=44)

Study (author, year)	Country	Study design	Key focus/aim	Key findings relevant to thesis
Alruwaili et al. (2024). Exploring nurses' awareness and attitudes toward artificial intelligence: Implications for nursing practice.	Saudi Arabia	Cross-sectional descriptive survey.	To assess nurses' awareness of and attitudes toward AI applications in clinical practice and their association with demographic factors.	Considering AI-literacy and influencing attitude of HCP is important in AI tools' implementation framework so HFAs and nurses see AI as a support, not a threat.
Andargoli et al. (2024) Intelligent decision support systems for dementia care: A scoping review	N/A / Multiple	Scoping review	Map AI-powered CDSS for dementia care and identify gaps/challenges vs WHO Global Action Plan domains	Scoping review: dementia AI-CDSS research is diagnosis-heavy; adoption hampered by data/cost/performance—calls for digital twins and trust-building; little LTC evidence.
Atalla et al. (2025). Embracing artificial intelligence in nursing: Exploring the relationship between artificial intelligence-related attitudes, creative self-efficacy, and clinical reasoning competency among nurses	Egypt	Cross-sectional descriptive-correlational study.	To examine relationships between nurses' AI-related attitudes, creative self-efficacy and clinical reasoning competency, and identify predictors of these competencies.	Evidence that shaping attitudes and self-efficacy is essential for safe AI-supported symptom management and escalation decisions in nursing homes.
Atee et al. (2018). A novel pain assessment tool incorporating automated facial analysis interrater reliability in advanced dementia.	Australia	Prospective observational reliability study	To examine inter rater reliability of the electronic Pain Assessment Tool for pain assessment in residents with advanced dementia.	Example of AI augmented symptom recognition for hidden pain in dementia supporting need for structured observation based tools.
Badawi et al. (2025). Multimodal detection of agitation in people with dementia in clinical settings: Observational pilot study.	Canada	Observational pilot model-development and feasibility study.	To evaluate a multimodal AI system combining wearable and behavioural data for detecting agitation in people with dementia in clinical settings.	Supports multimodal approach but reinforces that technical performance alone is not enough; workflow and ethical issues for continuous monitoring are central.
Badawy & Shaban (2025). Exploring geriatric nurses' perspectives on the adoption of artificial intelligence in elderly care: A qualitative study.	Saudi Arabia	Qualitative phenomenological interview study.	To explore how geriatric nurses perceive the adoption of AI technologies in elderly care, including benefits, risks and prerequisites.	Useful as qualitative evidence that nurse buy-in depends on seeing AI as a support for, not a replacement of, the caring relationship and professional judgement.
Boyce et al. (2022). Falls prediction using the nursing home minimum dataset	United States	Retrospective prognostic model-development and internal validation; no prospective clinical trial.	To develop and validate a fall prediction model for nursing home residents using only MDS 3.0 and drug data to predict falls within 90 days after an assessment.	Development and internal validation of an interpretable falls risk model using routine NH data, showing realistic performance and highlighting the need for external validation and implementation research.
Capodici et al. (2025). A scoping review of machine learning models to predict risk of falls in elders, without using sensor data	Multiple countries (US, Europe, Asia - high-income focus)	Scoping review (evidence mapping).	To identify and describe ML models predicting falls in older adults without sensor data, and summarise their performance and reporting.	Supports the argument that ML models must include validation, calibration and equity analyses, and that implementation in LTC requires more than just a promising AUC.
Charon et al. (2024). One month prediction of pressure ulcers in nursing home residents with Bayesian networks.	France	Retrospective case control model development and internal validation	To predict incident pressure ulcers one month ahead using nursing home EHR data	Shows benefits and risks of using highly sensitive risk prediction where many residents are flagged but few develop ulcers

Study (author, year)	Country	Study design	Key focus/aim	Key findings relevant to thesis
Cunningham et al. (2025). Prevalence of agitation and depression amongst people with dementia in Australian residential aged-care facilities: A cross-sectional study using LLMs.	Australia	Retrospective cross-sectional NLP study.	To estimate the prevalence of agitation and depression among residents with dementia in Australian RACFs using LLM-based analysis of free-text notes.	Shifts focus from single decisions to population-level insights; emphasizes documentation quality and the risk of under-recognised depression and agitation.
Delespierre et al. (2017)	France	Retrospective cross-sectional data-mining study.	To show that qualitative and quantitative text mining of physiotherapy clinical narratives in LTC EHRs can create a corpus and generate new, useful variables about residents' care and health.	Useful as early example of LTC EHR text mining; supports inclusion of free-text–derived indicators (e.g., mobility, feeding issues) in my conceptual framework.
Elhaddad et al. (2024). AI-Driven Clinical Decision Support Systems: An Ongoing Pursuit of Potential	International (authors Canada/UK)	Review	Review AI technologies in CDSS, applications, human-centred considerations, challenges (interpretability, bias, regulation), and adoption strategies	Useful to justify: "Why this" panel, data quality warnings, role-based views, feedback loop
Fossum et al. (2011). An Evaluation of the Usability of a Computerized Decision Support System for Nursing Homes.	Norway	Usability/implementation evaluation of a clinical decision support system (CDSS) in nursing homes. Qualitative usability study (multi-site); cognitive walkthrough and questionnaire.	Identify facilitators and barriers affecting nursing personnel's effective use of a CDSS for planning/treating pressure ulcers and malnutrition and following risk-assessment guidelines.	Early CDSSs succeed when task-fit (data quality/locatability), training, and integration are solved; "advice not automation" aligns with the thesis' escalation model.
Garcés-Jiménez et al. (2024). Predictive health monitoring: Leveraging artificial intelligence for early detection of infectious diseases in nursing home residents through discontinuous vital signs analysis.	Spain and Dominican Republic	Model development/validation. Retrospective ML development with internal repeats/bootstrapping; no external validation.	Early detection of acute respiratory, UTI, and other infections in nursing-home residents using ML on discontinuous daily vital signs and contextual info.	Targets respiratory & UTI. NH ML on daily vitals and context predicted infections 2–4 days early with basal baseline; promising for proactive escalation, but lacks external validation, calibration, thresholds and user/clinical impact.
Hacking et al. (2022). Text mining in long-term care Exploring the usefulness of artificial intelligence in a nursing home setting.	Netherlands	Exploratory cross-sectional methodological study	Explore usefulness of text-mining approaches for narrative data in a nursing home to evaluate experienced quality of care	Adjacent LTC AI exemplar (NLP for quality/experience data); useful background contrast with the thesis' focus on symptom decision points and escalation support.
Hendriks et al. (2024). Data science techniques to gain novel insights into quality of care: a scoping review of long-term care for older adults.	Primarily Europe and North America	Scoping review (evidence mapping).	To map and describe how data-science techniques (including AI and text mining) are used to gain insights into quality of care for older adults in long-term care settings.	Useful as major background source for LTC data-science landscape, justifying focus on explainable, nurse-centred CDS rather than purely technical models.
Kandel et al. (2025). Development of a predictive hospitalisation model for skilled nursing facility patients.	United States	Retrospective model development and internal validation	To develop a model that predicts hospitalisation or death within seven days for SNF (skill nursing facility) patients and compare its performance with clinician judgement	Shows how daily prognostic lists from routine EHR data can work while highlighting trade offs between sensitivity workload and actionable follow up.

Study (author, year)	Country	Study design	Key focus/aim	Key findings relevant to thesis
Kang et al. (2024). Designing and testing clinical simulations of an early warning system for implementation in acute care settings.	United States	Simulation-based formative study (implementation science / usability).	To use clinical simulations to evaluate a nurse-facing early-warning CDS and identify usability, workflow and training needs before implementation.	Useful methodological template for running simulations with Norwegian nursing-home staff when testing prototype symptom-support or escalation tools.
Kruse et al. (2017). Impact of electronic health records on long-term care facilities: Systematic review.	Primarily United States (plus some international LTC studies)	Systematic review (narrative).	To identify potential benefits, common conditions and EHR features associated with positive or negative impacts in LTC facilities.	Background for organisational/implementation chapter, especially barriers/facilitators when layering AI on top of existing EHRs.
Lee et al. (2020). Application of machine learning methods in nursing home research.	South Korea	Retrospective cross sectional model development with internal cross validation	To compare several machine learning methods in predicting facility level falls and to identify falls related to organisational and resident mix factors.	Useful methodological example showing that organisational factors can be combined with ML for facility risk insight alongside resident level approaches.
Lee et al. (2021). Identifying the Risk Factors Associated with Nursing Home Residents' Pressure Ulcers Using Machine Learning Methods.	South Korea	Retrospective cross sectional model development with internal validation	Use and compare ML models on nursing-home data to identify key risk factors for pressure ulcers (and predict PU risk).	Together with the paper above about fall prediction, illustrates ML use on facility data for safety outcomes offering organisational context for future resident level clinical decision support.
Lukkien et al. (2024). AI Assisted Decision Making in Long Term Care Qualitative Study on Prerequisites for Responsible Innovation.	Netherlands	Qualitative study (exploratory/descriptive; thematic analysis of interviews with LTC stakeholders)	Explore opportunities, risks, and prerequisites for responsible AI-assisted decision-making in long-term care from nurses' and other stakeholders' perspectives.	Identifies prerequisites for responsible AI in LTC: clear role responsibility and human oversight, transparency/explainability to support trust, staff training/competence needs, data quality and interoperability, GDPR-aware governance, and organisational support—highly relevant to this research's governance and workflow-integration findings.
Masot et al. (2022). Decision support-tools for early detection of infection in older people (aged >65 years): a scoping review	N/A / Multiple	Scoping review	Map DSTs for detecting infection in adults >65 and assess settings, stages of development, and gaps	Scoping review shows many infection DSTs but few LTC-ready and little implementation evidence; calls for general, practical tools for NHs.
McDonald et al. (2025). Electronic Decision Support for Deprescribing in Older Adults Living in Long-Term Care: A Stepped-Wedge Cluster Randomized Trial.	Canada	Stepped-wedge cluster randomized trial (prospective RCT).	To determine whether evidence-based electronic decision support for deprescribing, integrated into quarterly medication reviews in LTC homes, increases deprescribing of potentially inappropriate medications compared with usual care.	High-quality, recent LTC stepped-wedge RCT illustrating rule-based LTC CDS, workflow integration, and evaluation; focused on medication deprescribing rather than general symptom triage, but very informative for LTC AI/CDS design in the thesis.
Moon et al. (2022). The development of a web-based app employing machine learning for delirium prevention in long-term care facilities in South Korea.	South Korea	Multi-phase development / validation study (algorithm development, pilot validation, usability and feasibility evaluation); not an RCT.	To develop a web-based delirium prevention app for long-term care facilities, using a machine-learning–derived prediction algorithm plus evidence-based interventions, and to evaluate its validity, usability and feasibility among nurses.	Shows that symptom-specific, nurse-facing mobile CDSS can fit daily LTC workflow (risk levels + screening + checklist-based next steps) and is acceptable to nurses in a pilot; mainly informs design (actionable next steps/alerts/checklists) rather than effectiveness (small sample, no RCT or outcome evidence, limited alert-burden evaluation).

Study (author, year)	Country	Study design	Key focus/aim	Key findings relevant to thesis
Nanevski et al. (2025). The Potential of AI in Nursing Care: Multicenter Evaluation in Fall Risk Assessment.	Germany	Multicenter retrospective evaluation	Evaluate AI fall-risk prediction across sites vs rule-based tools	AI outperformed rule-based tools; cross-site training/federated learning gave limited added benefit; fairness checks suggested no sex differences but age-related performance variation—supports the thesis's governance/validation points.
Porcellato et al. (2025). Exploring applications of artificial intelligence in critical care nursing: A systematic review.	N/A / Multiple	Systematic review of quantitative and mixed-methods studies.	To systematically review how AI has been applied in critical care nursing and what outcomes and implementation challenges have been reported.	Provides a method and outcome benchmark showing current evidence gaps (prospective evaluation, usability, alert burden, ethics), which the thesis' LTC-focused work aims to address conceptually.
Shahidi et al. (2023). Machine learning risk estimation and prediction of death in continuing care facilities using administrative data.	Canada	Algorithm development & comparative modeling on retrospective cohort.	Identify factors associated with 60-day all-cause mortality in continuing-care residents during COVID-19 and compare LR vs multiple ML models; assess preprocessing and optimal cut-points.	COVID-era LTC/DSL (designated supportive living) admin-data model predicts 60-day mortality; good for risk alerts but needs local validation.
Shao et al. (2024). Development and external validation of a machine learning based fall prediction model for nursing home residents: A prospective cohort study.	China	Prospective cohort with external validation between homes	To develop and externally validate machine learning models that predict falls within six months among ambulatory nursing home residents	Shows how to structure baseline assessments and external validation while underlining need for local adaptation and usability testing
Skuban-Eiseler et al. (2023). Artificial Intelligence–Based Clinical Decision Support Systems in Geriatrics: An Ethical Analysis	Germany (lead); international	Normative ethical analysis (principle-based) with scoping of arguments	Provide a systematic ethical analysis of AI-CDSS use in geriatrics using Beauchamp & Childress's four-principles framework	Useful for the governance subsection. Ethical analysis flags autonomy, safety, privacy, and justice risks for AI-CDSS in geriatrics/LTC and recommends transparent, human-in-the-loop deployment
Solaiman (2024). Legal and Ethical Considerations of Artificial Intelligence for Residents in Post Acute and Long Term Care.	Qatar based author focus on United States law.	Narrative legal ethics analysis labelled original study but without empirical data	To propose a framework for analysing ethical and legal concerns when using AI for residents in post acute and long term care.	High level explainability requirement emphasised but no technical XAI methods described
Stancu et al. (2025). Applications of Machine Learning Algorithms in Geriatrics.	N/A / multiple	Systematic review	Synthesise how ML is used in geriatrics for prevention/diagnosis/prediction/monitoring/ treatment and identify commonly used models.	Highlights need for explainability (e.g., SHAP/LIME), better data standardisation/quality, and more real-world/prospective validation before routine clinical use.
Strauven et al. (2024). Unobtrusive nighttime movement monitoring to support nursing home continence care: algorithm development & validation.	Belgium	Lab algorithm development with LOSOCV.	Explore an unobtrusive bed-integrated sensor system to monitor nighttime movement (in bed / agitation / turn / out of bed) to support continence care in nursing homes.	Under-mattress sensors + XGBoost classified in-bed/turn/agitation/out-of-bed with good accuracy; promising for night-time continence care prompts, but needs real NH validation, thresholds, and integration.
Sun et al. (2020). Predicting Coronavirus Disease 2019 Infection Risk and Related Risk Drivers in Nursing Homes: A Machine Learning Approach.	USA	Model development with 10-fold CV + prospective external validation in a different state.	Build and validate an ML model to predict which NHs will have ≥1 resident COVID-19 case and identify key risk drivers to inform prevention policies.	Facility-level ML risk index predicted NH COVID-19 presence with AUC≈0.72 and strong correlation to LTCF outcomes; main drivers were community incidence and facility size/density; useful for targeting testing/PPE, not a bedside DSS.
Susnjak (2025). Towards clinical prediction with transparency: An explainable AI approach to survival modelling in residential aged care	New Zealand and Australia	Retrospective survival model development and internal validation	To develop an interpretable survival model for residents in aged care facilities and demonstrate transparent prediction of mortality at several horizons	Key reference for XAI in residential aged care: calibrated survival prediction with patient-level explanations to support transparent, ethically informed palliative/end of life decision-making.

Study (author, year)	Country	Study design	Key focus/aim	Key findings relevant to thesis
Suzuki et al. (2020). Deep learning prediction of falls among nursing home residents with Alzheimer's disease	Japan	Retrospective model-development study (inferred).	To develop a deep-learning model to predict falls in nursing-home residents with Alzheimer's disease.	Treated as background evidence that deep learning is being tried; weighted lower in synthesis due to abstract-only access.
Takeshita et al. (2025). Machine Learning Prediction for Postdischarge Falls in Older Adults: A Multicenter Prospective Study.	Japan	Multicenter prospective cohort with internal cross-validation.	To develop and internally validate ML models predicting falls within 3 months after discharge in older adults using easily collected hospital data.	Provides concrete example of prospective ML risk prediction using routine data and highlights need for external validation before use in LTC; good benchmark for AUC expectations.
Thapa et al. (2022). Predicting falls in long term care facilities: Machine learning study.	United States	Retrospective ML model development and internal validation	To develop and validate machine learning models to predict which LTC residents will fall within a future time window	Supports use of ML with routine LTC data to safely rule out low risk residents while acknowledging low PPV and need to avoid unnecessary alerts
Ventura Silva et al. (2024). Artificial intelligence in the organization of nursing care: A scoping review.	NA / multiple	Scoping review evidence synthesis	To synthesise the evidence on the use of artificial intelligence in the organisation of nursing care.	Provides a contemporary Portuguese perspective on AI in nursing care organisation summarising global tools and highlighting themes of monitoring prediction decision support and interaction as well as training ethics and governance which can support the thesis' background and discussion chapters.
Vithanage et al. (2025). A comprehensive evaluation of large language models for information extraction from unstructured electronic health records in residential aged care.	Australia	Retrospective technical evaluation of NLP/LLM models.	To evaluate the performance of several large language models for information extraction from unstructured EHR narratives in residential aged care.	Positions LLM IE as an enabling layer; reinforces that CDS quality depends on accurate extraction from messy nursing documentation rather than only on the prediction model.
Woo et al. (2021). Identifying urinary tract infection–related information in home care nursing notes.	USA	Algorithm development and validation on gold standard; large-scale descriptive application	Create and validate an NLP algorithm to automatically identify UTI-related information in home-care nursing notes and examine its occurrence around UTI-related ED visits/hospitalizations.	Rule-based NLP on nursing notes detected UTI-related language with significant accuracy and showed pre-event spikes before UTI ED/hospitalization—useful concept for proactive escalation.
Yesankar et al. (2025). AI-Powered Clinical Decision Support Systems (CDSS): Challenges, Benefits, Applications, and Future Directions. (conference paper)	N/A / multiple	Narrative review and conceptual analysis not systematic	To review challenges benefits applications and future directions of AI powered clinical decision support systems across healthcare.	High level CDSS overview that can be used to justify design principles such as human oversight XAI safety monitoring and workflow integration
Zhang et al. (2025). Leveraging retrieval-augmented generation-driven large language models to extract dementia agitation symptoms and triggers from free-text nursing notes.	Australia	Retrospective NLP model-development and evaluation	To evaluate retrieval-augmented LLMs for extracting agitation symptoms and triggers from free-text nursing notes in residential aged care.	Supports using LLM pipelines to move beyond simple risk scores towards actionable factors staff can address (noise, interactions, routines).
Zhu et al. (2022). Agitation Prevalence in People With Dementia in Australian Residential Aged Care Facilities: Findings From Machine Learning of Electronic Health Records.	Australia	Retrospective observational study using NLP	Estimate prevalence and types of agitation behaviours in people with dementia in Australian residential aged care using ML/NLP on EHR notes.	Shows free-text nursing notes can be mined to identify agitation patterns; agitation was common and often co-occurred (multiple behaviours in many residents), indicating high care burden and a need for better support/workflows.

Annex C – Prisma

Annex C1 — PRISMA flow

Reports identified via targeted methods + supplementary searching	52
Duplicate records removed	2
Reports assessed for eligibility (full text)	50
Reports excluded (full text)	6 *(reasons in Annex C2)*
Studies included in scoping review	44

Table C1. PRISMA-style table. The evidence set was identified through a targeted approach (seed papers recommended and iterative supplementary searching), therefore the flow focuses on full-text eligibility rather than large-volume database screening.

Annex C2 — Full-text exclusions table and duplicates

Reason codes

- **DUP** = duplicate record (retained once)
- **SET** = wrong setting (not NH/LTC / not residential aged care decision support)
- **AIM** = wrong aim (not clinical decision support / symptom decision-making)
- **TYPE** = protocol/conceptual (no evaluation/results)
- **GREY** = project webpage/grey literature (not a study report)

Table C1. Full-text excluded (n = 6)

Excluded full-text paper (APA)	***Code***	***Why excluded (1 line)***
Vitorino, L. M., Yoshinari Júnior, G. H., & Lopes-Júnior, L. C. (2025). Artificial intelligence in nursing: advancing clinical judgment and decision-making. ***Revista Brasileira de Enfermagem, 78****(4), e780401.*	*SET/AIM*	*General nursing perspective; not NH/LTC decision support focus. (SciELO)*
Torrado, J. C., Husebo, B. S., Allore, H. G., Erdal, A., Fæø, S. E., Reithe, H., et al. (2022). Digital phenotyping by wearable-driven AI… Protocol of the cyclic ActiveAgeing study. ***PLOS ONE, 17****(10), e0275747.*	*TYPE/SET*	*Protocol + digital phenotyping; not NH/LTC symptom decision support evaluation. (PLOS)*
Hoogendijk, E. O., Onder, G., Smalbil, L., Vetrano, D. L., et al. (2023). Optimising the care… design and protocol of I-CARE4OLD. ***BMJ Open, 13****(6), e072399.*	*TYPE*	*Protocol/design paper (no evaluated CDSS outcomes yet). (PubMed)*

Boyle, L. D., Giriteka, L., Marty, B., et al. (2025). Activity and behavioral recognition using sensing technology…: Umbrella review. ***Sensors, 25****(3), 668.*	*AIM*	*Sensing/behavior recognition review; not decision-support for symptom escalation/workflow. (MDPI)*
Hacking, C., Verbeek, H., Hamers, J. P. H., & Aarts, S. (2023). Comparing text mining and manual coding methods…. ***PLOS ONE, 18****(11), e0292578.*	*AIM*	*Method paper on analysing interview data; not clinical decision support. (ResearchGate)*
University of Bergen. (n.d.). The 5-D Project.	*GREY/TYPE*	*Project webpage (ongoing study), not a peer-reviewed evaluation paper. (www4.uib.no)*

Duplicates removed (before eligibility assessment) (n = 2)

- **Hacking et al. (2022)** – **DUP** (duplicate record; study retained once)
- **Lukkien et al. (2024)** – **DUP** (duplicate record; study retained once)

Annex D - Quality Rubric

Citation (short)	NH/LTC setting	Decision point	Effect size	External validation	XAI present	Workflow	Transferability to	HFA escalation	Score
Alruwaili 2024 nurse AI awareness attitudes	1	1	0	0	1	1	1	0	5
Andargoli 2024 — Artificial Intelligence in Medicine	0	1	0	0	0	1	1	1	4
Atalla 2025 AI attitudes clinical reasoning	1	1	0	0	1	1	1	0	5
Atee 2018 PainChek dementia pain assessment	1	1	0	0	0	1	1	1	5
Badawi 2025 JMIR multimodal agitation	1	1	0	0	0	1	1	1	5
Badawy 2025 geriatric nurses AI adoption	1	0	0	0	1	1	1	0	4
Boyce 2022 – hybrid CART-logit falls model using MDS	1	1	1	0	0	1	1	1	6
Capodici 2025 scoping review falls ML	1	1	0	0	1	1	1	0	5
Charon 2024 JAMDA pressure ulcer Bayesian network	1	1	1	0	0	1	1	1	6
Cunningham 2025 LLM agitation depression prevalence	1	0	0	0	0	1	1	0	3
Delespierre 2017 EHR text mining LTC	1	0	0	0	0	0	1	0	2
Elhaddad 2024 – AI-Driven CDSS (review)	0	0	0	0	1	1	1	0	3
Fossum 2011 - Applied Clinical Informatics	1	1	0	0	1	1	1	1	6
Garcés-Jiménez 2024 — CBM	1	1	1	0	1	1	1	1	7
Hacking 2022 text mining LTC quality	1	0	0	0	1	1	1	0	4
Hendriks 2024 data science in LTC	1	0	0	0	0	0	1	0	2
Kandel 2025 JAMDA SNF hospitalisation	1	1	1	0	0	1	1	1	6
Kang 2024 CONCERN simulation	1	1	0	0	1	1	1	1	6
Kruse 2017 EHRs in LTC systematic review	1	0	0	0	0	0	1	0	2
Lee 2020 ML methods in nursing homes	1	1	0	0	0	1	1	0	4
Lee 2021 ML pressure ulcer risk	1	1	0	0	0	1	1	0	4
Lukkien 2024 AI-assisted decision making LTC	1	1	0	0	1	1	1	1	6
Masot 2022 – Infection DSTs (scoping review)	1	0	0	0	0	0	1	1	3
McDonald 2025 – Deprescribing in LTC (SW-CRT)	1	1	1	1	0	1	1	0	6
Moon 2022 – Web_DeliPREVENT_4LCF delirium app (BMC)	1	1	1	0	0	1	1	1	6
Nanevski 2025 AI fall risk multicentre	0	1	1	0	0	1	1	1	5
Porcellato 2025 AI in critical care nursing SR	1	1	0	0	1	1	1	0	5
Shahidi 2023 — Scientific Reports	1	1	1	0	0	1	1	1	6
Shao 2024 JAMDA prospective falls	1	1	1	0	0	1	1	0	5
Skuban-Eiseler 2023 — JAMDA	1	0	0	0	0	1	1	1	4
Solaiman 2024 legal ethical PA LTC AI	1	0	0	0	0	1	1	1	4
Stancu 2025 ML in geriatrics review	0	0	0	0	0	0	1	0	1
Strauven 2024 — JMIR Nursing	1	1	1	0	1	1	1	1	7
Sun 2020 — JAMDA	1	1	1	1	1	1	1	0	7
Susnjak 2025 CMPB survival XAI RAC	1	1	1	0	1	1	1	1	7
Citation (short)	**NH/LTC setting**	**Decision point**	**Effect size**	**External validation**	**XAI present**	**Workflow**	**Transferability to**	**HFA escalation**	**Score**

Suzuki 2020 DL falls in AD (abstract only)	1	1	0	0	0	0	1	1	4
Takeshita 2025 postdischarge falls ML	1	1	0	0	1	1	1	0	5
Thapa 2022 JMIR falls prediction LTC	1	1	1	0	0	1	1	1	6
Ventura-Silva 2024 AI organisation of nursing care	1	1	0	0	1	1	1	0	5
Vithanage 2025 LLM IE RAC EHR	1	0	0	0	0	0	1	0	2
Woo 2021 — JAMDA	0	1	1	0	1	1	1	1	6
Yesankar 2025 AI-powered CDSS review	0	0	0	0	1	1	1	1	4
Zhang 2025 RAG LLM agitation triggers	1	1	0	0	1	1	1	1	6
Zhu 2022 NLP agitation prevalence	1	0	0	0	0	0	1	0	2

Annex E - Recruitment texts

Used in Norwegian and translated here to English.

Email to Participants

Subject: Invitation to participate in a master's thesis study (interview/focus group)

Hello,

My name is **Tania Braga**, and I am a master's student in **ManagiDiTH – Managing Digital Transformation in the Health Sector**. I am conducting a study about how **nurses and healthcare assistants (HFAs/helsefagarbeidere)** in nursing homes recognise, assess, document, and escalate **non-urgent symptom changes**, and what would be required for **AI-based decision support** to be useful and safe in everyday practice.

You are invited to take part in **a 35-45 minute individual interview / a 60 minute focus group**. Participation is voluntary, and you may withdraw at any time without giving a reason. We will not collect patient data, and we will ask you not to share identifiable resident information.

If you are interested, please contact [email] **/** phone number] or reply to this message. I will then contact you with further details.

Thank you for considering to participate.

Best regards,

Tania Braga

Master's student, ManagiDiTH

Supervisor: **Mitha Jose /** [email]

Email to Nursing Home Managers

Hi [NAME],

Thank you for last time.

I'm in the final stage of my master's thesis and would like to carry out a small pilot/user test and interviews related to everyday work in nursing homes (decision support/structured escalation when new symptoms occur).

I'm requesting permission to:

conduct 4–6 short interviews or user tests (approx. 30–35 minutes) with staff (nurses/healthcare assistants)

test a simple prototype (not a patient system—only a sketch tool/flow) and collect feedback on content and workflow

collect anonymous notes (no patient information, no personally identifiable data)

Plan:

When: week 4 (exact date can be adapted)

Where: at the nursing home / in a suitable room

Participation: voluntary, and I will adjust timing to minimize impact on daily work.

I can send a short information sheet to participants (purpose, voluntariness, data processing) and a simple interview guide. If you prefer, we can also have a short call (10–15 minutes) to clarify practical arrangements.

Please let me know whether this can be approved, and whether there are any internal procedures I need to follow before I begin.

Thank you.

Tania Braga

Sms to Participants

Sms A

Hi [NAME]. I'm Tania, a nurse at [NURSING HOME]. I spoke with [NAME] and she said I could contact you. I'd like to ask if you could take part in a short interview for my master's thesis. Would you be able to participate (35–45 minutes) tomorrow from xx:xx to xx:xx?

You don't need to prepare. This is not a test of knowledge—it's about your experience and how everyday work functions in practice. The topic is how we identify and follow up non-acute symptom changes in nursing homes, and what could make a simple AI support tool safe and useful.

Best regards,

Tania

Sms B

Hi [NAME]. Can you help me with my master's thesis? I'm researching what healthcare professionals think about using artificial intelligence in nursing homes. I need to speak with healthcare staff and ask a few questions. It will take about 45 minutes.

Best regards,

Tania

Annex F - Participant Information Documents

Used in Norwegian and translated here to English.

Individual interview

Study title: AI Decision Support for Non-Urgent Symptom Management in Nursing Homes

Researcher: Tania Braga, Laurea University of Applied Sciences/ISCTE, [email]

Supervisor: Mitha Jose, Laurea University of Applied Sciences, [email]

Version/date: v1/19/01/2026

1) Invitation to participate

You are invited to take part in a research study as part of my master's thesis. This sheet explains why the research is being done and what participation involves. Please read it carefully and ask any questions you may have before deciding.

2) Purpose of the study

The study explores:

- how staff notice and interpret **non-urgent symptom changes**,
- how decisions are made about **monitoring, assessment, documentation, and escalation**, and
- what requirements and safeguards an **AI decision-support tool** would need to be **usable, reliable, and trustworthy** in nursing homes.

3) Why you have been invited

You are invited because you work as a **[nurse / healthcare assistant]** in a nursing home and have experience with symptom observation and escalation in daily care.

4) Participation content

If you agree to participate, you will take part in one interview that will last 35 to 45 minutes. Topics that can be discussed include:

· Examples of non-urgent symptom changes and how they are handled,

· Challenges and characteristics of clinical decision in nursing homes

· Requirements for a tool to be safe and useful in supporting decision making in non acute situations.

An example of a prototype will be shown and discussed.

Audio recording will be conducted if you permit it and only to help accurate transcription. Please inform me if you do not consent it.

Important: Patient data is not being collected. Avoid sharing personal and sensitive data about yourself or others. If such details appear, they will be removed from transcripts.

5) Voluntary participation and withdrawal

Participation is voluntary and you may withdraw at any time, without giving a reason, and without any consequences. If you withdraw, you might be asked if the researcher may keep and use data collected up to that point, or if you want it deleted.

6) Possible benefits and disadvantages

You may find it useful to think about and analyse topics such as clinical decision-making and workflow challenges. Your perspective and contributions may help improve the design of future digital tools for clinical decision support in nursing homes.

Minimal risk is expected, but discussing difficult situations may feel uncomfortable. You may skip questions or stop the interview at any time. Please consider this is not a test of ability or competency and there is only a focus on understanding several perspectives, experiences and opinions.

7) Confidentiality

All the data collected will be treated confidentially. The information given in the interview can be quoted in the thesis, but no personal names will be revealed (i.e. Nurse 3). The place of employment will also not be revealed.

8) Data protection / GDPR (summary)

Personal data such as role, years of experience or voice recording, will be processed under GDPR principles. There is a Privacy Statement and a Data Management Plan which describes how data are handled, stored, and disposed.

9) Results and dissemination

Findings will be reported in a master's thesis and may be presented in academic contexts. A short summary of findings can be shared with your workplace if requested.

10) Contact details

Researcher: Tania Braga, [email]

Supervisor: Mitha Jose, [email]

Version/date: v1/19/01/2026

Focus group

Study title: AI Decision Support for Non-Urgent Symptom Management in Nursing Homes

Researcher: Tania Braga, Laurea University of Applied Sciences/ISCTE, [email]

Supervisor: Mitha Jose, Laurea University of Applied Sciences, [email]

Version/date: v1/19/01/2026

1) Invitation to participate

You are invited to take part in a research study as part of my master's thesis. This sheet explains why the research is being done and what participation involves. Please read it carefully and ask any questions you may have before deciding.

2) Purpose of the study

The study explores:

- how staff notice and interpret **non-urgent symptom changes**,
- how decisions are made about **monitoring, assessment, documentation, and escalation**, and
- what requirements and safeguards an **AI decision-support tool** would need to be **usable, reliable, and trustworthy** in nursing homes.

3) Why you have been invited

You are invited because you work as a **[nurse / healthcare assistant]** in a nursing home and have experience with symptom observation and escalation in daily care.

4) Participation content

If you agree to participate, you will take part in one interview that will last 35 to 45 minutes. Topics that can be discussed include:

· Examples of non-urgent symptom changes and how they are handled,

· Challenges and characteristics of clinical decision in nursing homes

· Requirements for a tool to be safe and useful in supporting decision making in non acute situations.

An example of a prototype will be shown and discussed.

Audio recording will be conducted if you permit it and only to help accurate transcription. Please inform me if you do not consent it.

Important: Patient data is not being collected. Avoid sharing personal and sensitive data about yourself or others. If such details appear, they will be removed from transcripts.

5) Voluntary participation and withdrawal

Participation is voluntary and you may withdraw at any time, without giving a reason, and without any consequences. If you withdraw, you might be asked if the researcher may keep and use data collected up to that point, or if you want it deleted.

6) Possible benefits and disadvantages

You may find it useful to think about and analyse topics such as clinical decision-making and workflow challenges. Your perspective and contributions may help improve the design of future digital tools for clinical decision support in nursing homes.

Minimal risk is expected, but discussing difficult situations may feel uncomfortable. You may skip questions or stop the interview at any time. Please consider this is not a test of ability or competency and there is only a focus on understanding several perspectives, experiences and opinions.

7) Confidentiality

All the data collected will be treated confidentially. The information given in the interview can be quoted in the thesis, but no personal names will be revealed (i.e. Nurse 3). The place of employment will also not be revealed.

The researcher will anonymise data, but because this is a group setting, it cannot be fully guaranteed what other participants will share outside the session. All participants will be asked to respect group confidentiality and not repeat personal statements outside the group.

8) Data protection / GDPR (summary)

Personal data such as role, years of experience or voice recording, will be processed under GDPR principles. There is a Privacy Statement and a Data Management Plan which describes how data are handled, stored, and disposed.

9) Results and dissemination

Findings will be reported in a master's thesis and may be presented in academic contexts. A short summary of findings can be shared with your workplace if requested.

10) Contact details

Researcher: Tania Braga, Laurea University of Applied Sciences/ISCTE, [email]
Supervisor: Mitha Jose, Laurea University of Applied Sciences, [email]

Version/date: v1/19/01/2026

Doctor Interview

Study title: AI Decision Support for Managing Non-Urgent Symptom Changes in Nursing Homes

Researcher: Tania Braga, Laurea University of Applied Sciences/ISCTE, [email]

Supervisor: Mitha Jose, Laurea University of Applied Sciences, [email]

Version/date: v1 / 19.01.2026

1) Invitation to participate

You are invited to take part in a research project as part of my master's thesis. This information sheet explains why the study is being conducted and what participation involves. Please read it carefully and feel free to ask questions before deciding.

2) Purpose of the study

This study examines how nursing home staff recognize and interpret non-urgent changes in residents' symptoms; how decisions are made about observation, assessment, documentation, and escalation; and what does a future AI tool to support decision making in non urgent symptom changes are essencial for its design and safe use.

Your interview focuses on the physician perspective, including how you receive escalations, gaps do you identify and , how you decide next steps. Finally, it will be discussed your perspective of requirements to trust AI-supported decision recommendations in this context.

3) Why you are invited

You are invited because you work as a physician and are involved in healthcare for nursing home / long-term care residents and have experience with clinical decisions and escalation pathways from nursing homes.

4) What participation involves

If you consent to participate, you will take part in one individual interview (35–45 minutes).

We will talk about:

- examples of non-urgent symptom changes and how escalation is handled from a physician perspective
- what information is necessary to make safe decisions (and what is often missing)
- thresholds for "wait and see" versus further assessment versus escalation (e.g., rapid medical assessment/visit, emergency clinic, hospital)
- a short review of a simple storyboard/prototype (low-fidelity) to discuss requirements for an AI tool

Audio recording will be conducted if you permit it and only to help accurate transcription. Please inform me if you do not consent it.

Important: Patient data is not being collected. Avoid sharing personal and sensitive data about yourself or others. If such details appear, they will be removed from transcripts.

5) Voluntary participation and withdrawal

Participation is voluntary and you may withdraw at any time, without giving a reason, and without any consequences. If you withdraw, you might be asked if the researcher may keep and use data collected up to that point, or if you want it deleted.

6) Possible benefits and disadvantages

You may find it useful to think about and analyse topics such as clinical decision-making and workflow challenges. Your perspective and contributions may help improve the design of future digital tools for clinical decision support in nursing homes.

Minimal risk is expected, but discussing difficult situations may feel uncomfortable. You may skip questions or stop the interview at any time. Please consider this is not a test of ability or competency and there is only a focus on understanding several perspectives, experiences and opinions.

7) Confidentiality

All the data collected will be treated confidentially. The information given in the interview can be quoted in the thesis, but no personal names will be revealed (i.e. Nurse 3). The place of employment will also not be revealed.

8) Data protection / GDPR (summary)

Personal data such as role, years of experience or voice recording, will be processed under GDPR principles. There is a Privacy Statement and a Data Management Plan which describes how data are handled, stored, and disposed.

9) Results and dissemination

Findings will be reported in a master's thesis and may be presented in academic contexts. A short summary of findings can be shared with your workplace if requested.

10) Contact information

Researcher: Tania Braga, Laurea University of Applied Sciences/ISCTE, [email]

Supervisor: Mitha Jose, Laurea University of Applied Sciences, [email]

Version/date: v1 / 22.01.2026

Annex G – Interview Consent forms

Used in Norwegian and translated here to English.

Individual Interview

Study title: AI Decision Support for Non-Urgent Symptom Management in Nursing Homes

Participant ID:

Date:

Please initial each bullet point:

- I confirm that I have read and understood the Participant Information Sheet (version/date _____).
- I confirm that I have received and understood the privacy policy.
- I understand participation is voluntary and I can withdraw at any time without consequences.
- I understand the interview will focus on work practices and that I should not share resident-identifying information.
- I consent to the interview being audio-recorded.

- I understand recordings will be transcribed and anonymised; identifiable information will be removed where possible.
- I consent to anonymised quotes being used in the thesis/publications/presentations.
- I know who to contact with questions or complaints.

Optional

- I agree to be contacted for clarification/follow-up (email/phone: ___).

Participant name: _____ Signature: _____ Date: _____

Researcher name: __Tania Braga_____ Signature: _____ Date: _____

Focus Group

Study title: AI Decision Support for Non-Urgent Symptom Management in Nursing Homes

Participant ID: [assigned code]

Date: [YYYY-MM-DD]

Please initial each bullet point:

- I confirm that I have read and understood the Participant Information Sheet (version/date _____).
- I understand participation is voluntary and I can withdraw at any time without consequences.
- I understand the interview will focus on work practices and that I should not share resident-identifying information.
- **I consent to the interview being audio-recorded.**
- I understand recordings will be transcribed and anonymised; identifiable information will be removed where possible.
- I consent to anonymised quotes being used in the thesis/publications/presentations.
- I know who to contact with questions or complaints.
- I understand that confidentiality in a focus group cannot be fully guaranteed, and I agree to respect the confidentiality of other participants by not sharing what is discussed outside the session.

Optional

- I agree to be contacted for clarification/follow-up (email/phone: ___).

Participant name: _____ Signature: _____ Date: _____

Researcher name: __Tania Braga_____ Signature: _____ Date: _____

Annex H — Privacy statement (Personvernerklæring)

Used in Norwegian and translated here to English.

Title: Privacy statement for research participants

Study: AI Decision Support for Non-Urgent Symptom Management in Nursing Homes

Data controller: Laurea University of Applied Sciences/ISCTE

Contact: Tania Braga [email]

Supervisor: Mitha Jose [email]

What data we collect

- Role and background variables (e.g., years of experience, unit type, shift type)
- Audio recording of interview/focus group (if consented)
- Transcript notes from the session (anonymised)

Purpose of processing

To answer the thesis research questions about workflow decision points and requirements for safe AI decision support in nursing homes.

Legal basis (GDPR)

Task in the public interest/research + explicit consent for recording.

How data are stored and protected

- Audio files will be stored on the researcher's mobile phone or in WhisperAI, without data sharing. Only the researcher will have access.
- Transcripts pseudonymised (Participant IDs)
- Access restricted to research team
- Data retained until 23/01/2027 and then deleted/anonymised according to the DMP

Participant rights

- Withdraw consent; access; rectification; restriction; complaint pathway
- Limits: in focus groups, complete deletion of all contributions may be difficult if already merged into anonymised thematic results.

Data transfer

No data will be transferred outside EEA.

No participant or resident confidential data will be entered into generative AI tools.

Annex I — Demographic/context form

Used in Norwegian and translated here to English.

Nurse and HFA

Participant ID: ___ Date: _____

1. Role: ☐ Nurse ☐ HFA/helsefagarbeider ☐ Other: _____
2. Years working in nursing home/LTC: ☐ <1 ☐ 1–3 ☐ 4–7 ☐ 8–15 ☐ 16+
3. Current shift pattern (typical): ☐ Day ☐ Evening ☐ Night ☐ Mixed
4. Main unit type: ☐ Somatic LTC ☐ Dementia unit ☐ Mixed ☐ Other: _____
5. Frequency of non-urgent symptom change situations: ☐ Daily ☐ Weekly ☐ Monthly ☐ Rare
6. Digital tools used in daily work (tick all): ☐ EPJ (Cosdoc, osv) ☐ IKOS ☐ TQM (quality system) ☐ Other: ___
7. Comfort with digital tools: ☐ Low ☐ Medium ☐ High
8. Have you used any help tools at work (hjelpeverktøy), like a checklist (sjekkliste), a scoring form, or a warning/alert tool (varslingsverktøy) that helps decide what to check or who to call? ☐ Yes ☐ No ☐ Not sure

Doctor

Participant ID:	_____	Date:	_____

Current role (tick one):

☐ Nursing home physician (attending/supervising physician)
☐ GP responsible for nursing home residents
☐ Out-of-hours / emergency clinic physician
☐ Other: _____

Years working with nursing home / LTC patients:

☐ <1 ☐ 1-3 ☐ 4-7 ☐ 8-15 ☐ 16+

Typical contact format with nursing homes (tick all that apply):

☐ Scheduled rounds/visits/supervision
☐ Phone consultations during daytime
☐ Contact outside working hours / via emergency clinic
☐ Digital consultations (video/secure messaging)
☐ Other: _____

How often do you receive escalations about non-urgent symptom changes?

☐ Daily ☐ Weekly ☐ Monthly ☐ Rarely

Which information sources do you usually use? (tick all that apply):

☐ Verbal report from nurse (phone)

☐ Written nursing notes (free text)
☐ Vital signs / observations
☐ Medication list
☐ Lab tests (if available)
☐ Previous medical history/diagnoses
☐ Structured tools (e.g., checklists/scoring tools)
☐ Other: _____

How comfortable are you with digital tools used in municipal health services?

☐ Low ☐ Medium ☐ High

Have you used a decision-support tool related to symptom assessment/escalation?

☐ Yes ☐ No ☐ Unsure
If yes (optional): _____

Annex J – Interview Scripts

Used in Norwegian and translated here to English.

Individual interview script – HFA and Nurse

≤45 minutes total

- **0–8 min**: greeting + info sheet check + consent + record permission
- **8–43 min**: interview questions
- **43–45 min**: close + withdrawal reminder

Remind main rules: avoid information that could identify residents; focus on non-urgent symptom changes and decision processes; the participant may skip questions or stop at any time.

PART 1

1 (2 minutes): Can you briefly describe your role and what a typical shift looks like?

Q2 (2 minutes): Think of a recent situation where a resident had a *non-urgent* change. What did you notice first? What made you notice it (who noticed, how, when)?

Q3 (4 minutes): What made you think it could matter? What were you unsure about?

Q4 (4 minutes): What did you do next, step by step?

Q5 (4 minutes): What information did you check (for example observations, vital signs, notes, medication list), and where did you find it?

Q6 (4 minutes): What was missing, hard to find, or took too long?

Q7 (5 minutes): How did you decide between: **keep monitoring**, **assess more now**, or **escalate** (call a nurse/doctor)?

Q8 (4 minutes): What did you document, and what tends to get lost between shifts or staff?

Q9 (5 minutes): In that kind of situation, what would be most helpful from a digital tool to support your decision?

Q10 (5 minutes): What would it need to show so you could trust it *in the moment*? And what rules/safeguards would be needed for safe use?"

Q11 (2 minutes): If you choose only **three requirements** for a future tool, what are they?

PART 2 – Scenario card

a) instructions

1. This is a just a hypothetical scenario. We are not evaluating accuracy.
2. Present scenario card

b) Scenario cards - printed

Scenario A — Early deterioration/infection suspicion (non-urgent)

Resident: 84, dementia mild/moderate, baseline: eats well, walks with aid.

Change: “More tired today, ate half lunch, slightly confused.”

Available info: last temp 37.6 (yesterday), RR unknown, no BP measured yet.

Question: What do you do now? What would make you escalate?

Scenario B — Delirium risk / acute confusion (non-urgent but concerning)

Resident: 79, usually calm and oriented.

Change: sudden confusion in evening, restless, says “people are in my room”, poor sleep.

Available info: new medication started 3 days ago; mild dehydration suspected.

Question: What information do you need and what is your escalation threshold?

Scenario C — Behaviour change / agitation trigger (non-urgent escalation)

Resident: 76, dementia severe, usually calm after breakfast.

Change: agitation every afternoon for 3 days; staff notes mention noise in hallway and toileting resistance.

Available info: nursing notes (free text), no continuous monitoring.

Question: What would help you identify triggers and decide actions/escalation?

PART 3 – prototype/storyboard walkthrough

a) instructions

1. Show the prototype/storyboard walkthrough (on the pc)
2. “This is a concept storyboard/prototype. It is not in actual use, nor is it a final tool. We are not evaluating accuracy.”
3. Click on some of the fields to demonstrate basic functionality – few seconds
4. Ask participant to “use” the screens (think aloud).
5. Capture: decision point, missing info, trust requirements, governance needs.

b) questions

- Which choice does this help you make: observe, assess further, or escalate?
- What is helpful here? What is missing?
- Would you trust this? Why/why not?
- What explanation would you need *in that moment*?
- Who should see this output (HFA, nurse, physician)?
- What could go wrong if staff follow it? How to reduce that risk? What needs to be safeguarded?

Individual interview script – Doctor

≤45 minutes total

- **0–8 min**: greeting + info sheet check + consent + record permission
- **8–43 min**: interview questions
- **43–45 min**: close + withdrawal reminder

Remind main rules: avoid information that could identify residents; focus on non-urgent symptom changes and decision processes; the participant may skip questions or stop at any time.

Q1. (3 minutes): Could you briefly describe your role and how you usually work with the other healthcare professionals and the residents at the nursing home (e.g., scheduled visits, phone contact)?

Ask about: frequency, typical communication channel, who contacts you, decision authority, constraints/framework conditions.

(10 minutes)

Q2. Think of a recent situation where a resident had a non-urgent symptom change (not an immediate emergency). What was the first information you received?

Ask about: who called, what they said, what they were concerned about.

Q3. What made you think this could be significant (or not)? What were you uncertain about?

Ask about: unclear baseline, dementia/communication limitations, complex comorbidity, medication effects.

Q4. What did you do next, step by step?

Ask about: what questions you asked, what you asked staff to check, plan for observation/monitoring, timing for follow-up.

Q5. What information did you need most, and where should it ideally come from?

Ask about: vital signs, trend/change from baseline, fluids/food/diuresis, altered mental status, pain, medication changes, notes, objective findings/measurements.

(10 minutes)

Q6. What is often missing, unclear, or takes too long when a nursing home contacts you?

Ask about: missing vital signs, unclear baseline, unclear timeline, lack of structure in information, fragmented documentation.

Q7. How do you choose between if the resident should be observed more, or whether more info should be collected/further evaluation now or whether to escalate to a hospital or emergency department?

Ask about: typical thresholds; decisive red flags; when you require a nurse assessment before deciding.

Q8. Which types of escalation errors can occur?

Ask about: delay in treatment or assessment, unnecessary treatment or hospitalization, medication safety, missed delirium, dehydration, delayed sepsis/infection.

(3 minutes)

Q9. What information is often lost between shifts and between staff and physician?

Ask about: handover challenges, free-text notes, lack of structured summary.

Q10. Could you point one improvement that would help the decision making process be better in these type of scenarios (non-urgent symptom changes)? Do you think the staff communicates non-urgent changes to you always effectively and efficiently? What can be improved?

Scenario cards (10 minutes)

Scenario D - Telephone triage: possible infection vs dehydration (non-urgent onset, uncertain course)

Resident: 86 years old, moderate dementia.

Baseline: mobile with assistive device, usually eats well.

Change (reported by staff): “More tired since yesterday, eating less, a bit more confused.”

Information you have by phone now: Temp 37.8 today (37.4 yesterday). RR not measured. BP not measured. SpO2 unknown.

Other context: less urine today; one episode of loose stool; no clear cough.

Doctor decision task:

- What are your first five questions to staff?
- What is the minimum dataset you need before you can give safe advice?
- What do you recommend now (observation, further assessment, orders/actions, possibly treatment)?
- What are your escalation thresholds (emergency clinic / admission)?

Scenario E - Medication-related delirium risk (non-urgent but concerning)

Resident: 78 years old, usually calm and oriented.

Change: acute confusion and agitation in the evenings for 2 days. Sleeping poorly. Says things that may resemble hallucinations (“people are in the room”).

Medication timeline: new medication started 4 days ago (or dose increase). PRN sedative/sleeping medication used sometimes.

Available information: reduced food/fluid intake; suspected mild dehydration; pain status unclear; latest BP/pulse unknown.

Questions:

- What do you suspect, and what must be clarified/excluded first?
- Which medication actions do you consider (pause/adjust/avoid)?
- Which observations/measurements or tests do you need from the nursing home?
- How and when should staff escalate?

Storyboard / prototype walkthrough (8 minutes)

What information is absolutely necessary for you to make a safe decision here? What is missing in the storyboard output?
Would you trust an AI-supported recommendation here? Why/why not?
What explanation would you need in the moment (baseline/trend, data quality, uncertainty, "why this recommendation")?
Who should see AI output (HCA, nurse, physician)? What can go wrong if staff follow it? What safeguards are needed (workflow rules, training, escalation rules, logs/audit trail, requirements for updates/validation)?

Prioritization and closing (2 minutes)

If you could choose only three requirements for safe AI decision support in nursing homes, what are they?

What are the three most important risks you want to prevent?

- Closing. Reminder of option to withdraw.

Focus group script

0) Interview set-up (5min)

- Thank participant, confirm time.
- Confirm they read the information sheet and signed consent.
- Confirm permission to record.
  Ground rules:
- Remind: *no resident-identifying details*.
- Respect each other; one person at a time.
- Keep what others say confidential outside the group.

1) Shared workflow mapping (10 minutes)

Prompt: "Let's map what usually happens when a resident has a non-urgent change."

- Who notices first?
- What information is checked?

- When is a nurse called? When is a doctor contacted?
- Where does documentation happen?

2) Two contrasting cases (10 minutes)

Discuss two typical scenarios (e.g., early infection suspicion vs behaviour change).

- How would each be handled?
- What information gaps occur?

3) Scenario cards (10 minutes)

- "This is a just a hypothetical scenario. We are not evaluating accuracy.
- Present scenario cards

Scenario A — Early deterioration/infection suspicion (non-urgent)

Resident: 84, dementia mild/moderate, baseline: eats well, walks with aid.

Change: "More tired today, ate half lunch, slightly confused."

Available info: last temp 37.6 (yesterday), RR unknown, no BP measured yet.

Question: What do you do now? What would make you escalate?

Scenario B — Delirium risk / acute confusion (non-urgent but concerning)

Resident: 79, usually calm and oriented.

Change: sudden confusion in evening, restless, says "people are in my room", poor sleep.

Available info: new medication started 3 days ago; mild dehydration suspected.

Question: What information do you need and what is your escalation threshold?

Scenario C — Behaviour change / agitation trigger (non-urgent escalation)

Resident: 76, dementia severe, usually calm after breakfast.

Change: agitation every afternoon for 3 days; staff notes mention noise in hallway and toileting resistance.

Available info: nursing notes (free text), no continuous monitoring.

Question: What would help you identify triggers and decide actions/escalation?

4) Prototype/storyboard walkthrough (15 minutes)

a) instructions

6. Show the prototype/storyboard walkthrough (on the pc)
7. "This is a concept storyboard/prototype. It is not in actual use, nor is it a final tool. We are not evaluating accuracy."
8. Click on some of the fields to demonstrate basic functionality – few seconds
9. Ask participants to "use" the screens (think aloud).
10. Capture: decision point, missing info, trust requirements, governance needs.

b) questions

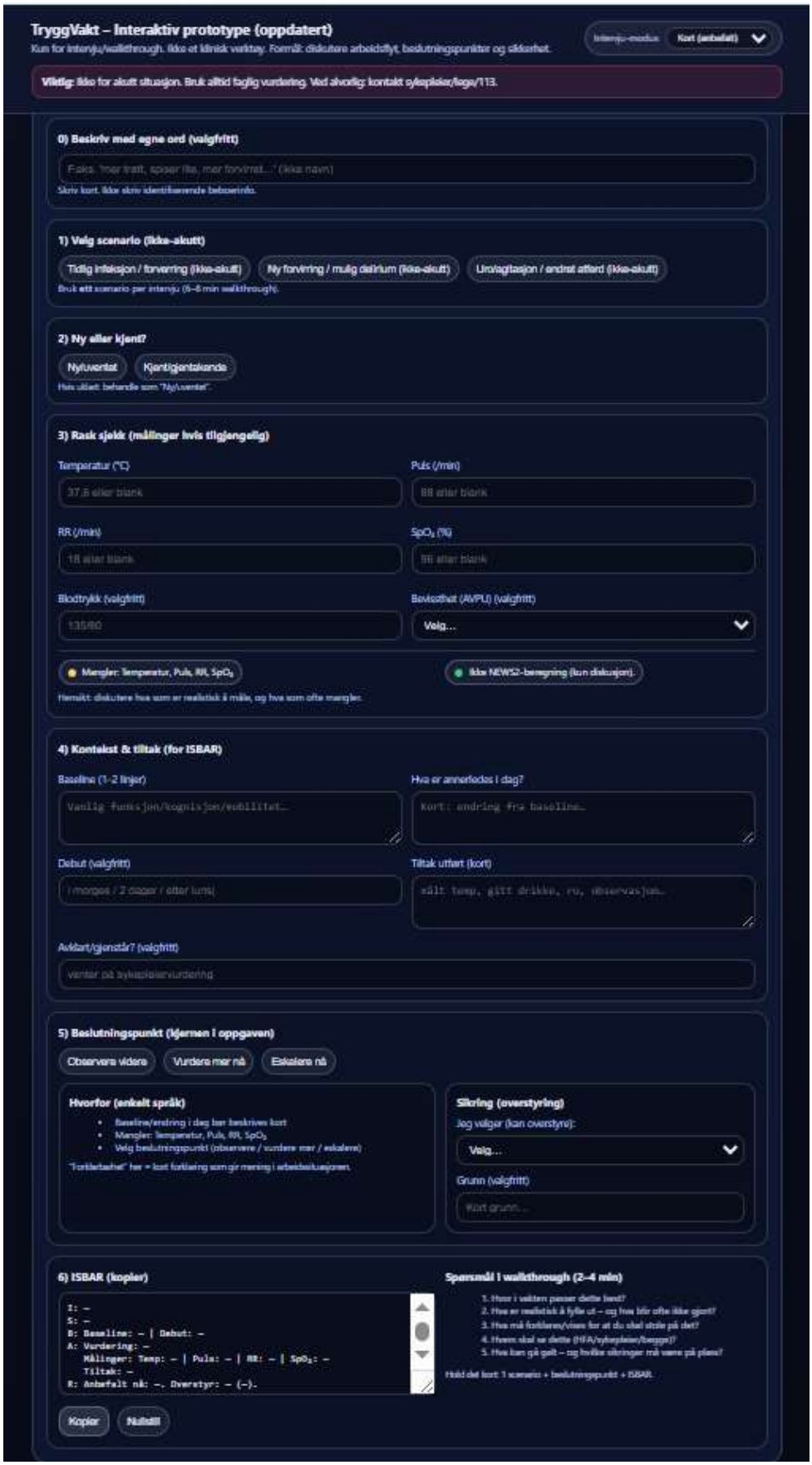
- Which choice does this help you make: observe, assess further, or escalate?
- What is helpful here? What is missing?
- Would you trust this? Why/why not?
- What explanation would you need *in that moment*?

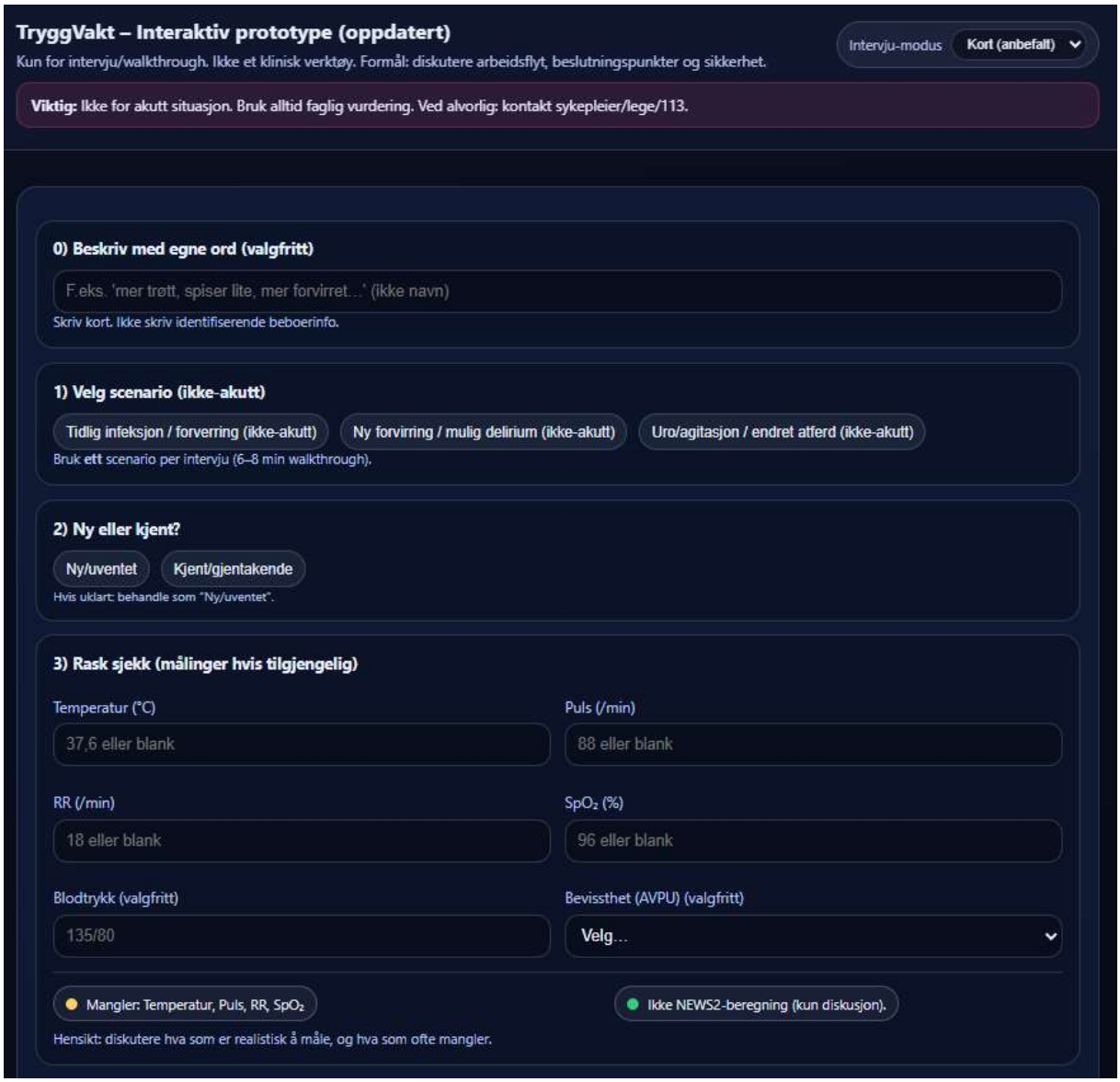
- Who should see this output (HFA, nurse, physician)?
- What could go wrong if staff follow it? How to reduce that risk? What needs to be safeguarded?

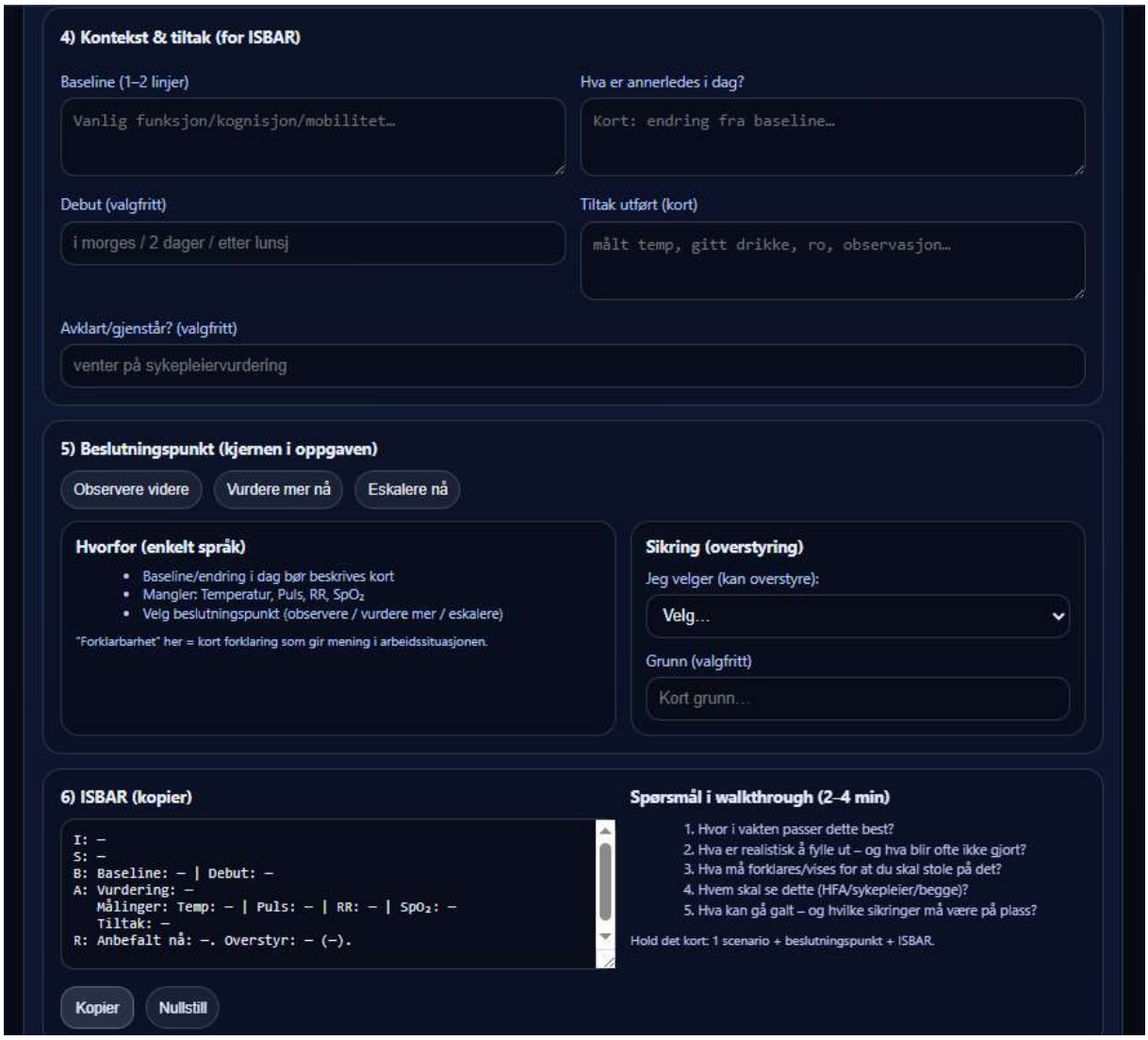
5) Prioritisation (10 minutes)

Group ranks top 5 requirements and top 5 risks/concerns.

Annex K - Prototype/storyboard walkthrough screenshots




TryggVakt – Interaktiv prototype (oppdatert)
Kun for intervju/walkthrough. Ikke et klinisk verktøy. Formål: diskutere arbeidsflyt, beslutningspunkter og sikkerhet.
Intervju-modus
Kort (anbefalt)
Viktig: Ikke for akutt situasjon. Bruk alltid faglig vurdering. Ved alvorlig: kontakt sykepleier/lege/113.
0) Beskriv med egne ord (valgfritt)
F.eks. 'mer trøtt, spiser lite, mer forvirret…' (ikke navn)
Skriv kort. Ikke skriv identifiserende beboerinfo.
1) Velg scenario (ikke-akutt)
Tidlig infeksjon / forverring (ikke-akutt)
Ny forvirring / mulig delirium (ikke-akutt)
Uro/agitasjon / endret atferd (ikke-akutt)
Bruk ett scenario per intervju (6–8 min walkthrough).
2) Ny eller kjent?
Ny/uventet
Kjent/gjentakende
Hvis uklart: behandle som "Ny/uventet".
3) Rask sjekk (målinger hvis tilgjengelig)
Temperatur (°C)
37,6 eller blank
Puls (/min)
88 eller blank
RR (/min)
18 eller blank
SpO₂ (%)
96 eller blank
Blodtrykk (valgfritt)
135/80
Bevissthet (AVPU) (valgfritt)
Velg…
Mangler: Temperatur, Puls, RR, SpO₂
Ikke NEWS2-beregning (kun diskusjon).
Hensikt: diskutere hva som er realistisk å måle, og hva som ofte mangler.


4) Kontekst & tiltak (for ISBAR)
Baseline (1–2 linjer)
Vanlig funksjon/kognisjon/mobilitet…
Hva er annerledes i dag?
Kort: endring fra baseline…
Debut (valgfritt)
i morges / 2 dager / etter lunsj
Tiltak utført (kort)
målt temp, gitt drikke, ro, observasjon…
Avklart/gjenstår? (valgfritt)
venter på sykepleiervurdering
5) Beslutningspunkt (kjernen i oppgaven)
Observere videre
Vurdere mer nå
Eskalere nå
Hvorfor (enkelt språk)
Baseline/endring i dag bør beskrives kort
Mangler: Temperatur, Puls, RR, SpO2
Velg beslutningspunkt (observere / vurdere mer / eskalere)
"Forklarbarhet" her = kort forklaring som gir mening i arbeidssituasjonen.
Sikring (overstyring)
Jeg velger (kan overstyre):
Velg…
Grunn (valgfritt)
Kort grunn…
6) ISBAR (kopier)
I: –
S: –
B: Baseline: – | Debut: –
A: Vurdering: –
Målinger: Temp: – | Puls: – | RR: – | SpO2: –
Tiltak: –
R: Anbefalt nå: –. Overstyr: – (–).
Spørsmål i walkthrough (2–4 min)
1. Hvor i vakten passer dette best?
2. Hva er realistisk å fylle ut – og hva blir ofte ikke gjort?
3. Hva må forklares/vises for at du skal stole på det?
4. Hvem skal se dette (HFA/sykepleier/begge)?
5. Hva kan gå galt – og hvilke sikringer må være på plass?
Hold det kort: 1 scenario + beslutningspunkt + ISBAR.
Kopier
Nullstill

Annex L - Key gaps identified in the literature (NH/LTC symptom-related decision support)

Key gaps identified in the literature (NH/LTC symptom-related decision support)

Key gaps in the literature on symptom-related decision support systems in the context of Nursing Homes and Long-Term Care (NH/LTC) settings include:

- **Ambiguity in decision points**: the description of the decision support system does not always clearly articulate the exact clinical decision point being addressed.
- **Poor role alignment**: the role alignment does not consistently clarify which user the tool is designed for —whether the frontline HFA, the registered nurse with clinical accountability, or the physician— despite their different levels of authority.
- **Overlook of Norwegian practice**: the Norwegian role and workflow is rarely addressed; that is, the tendency of the nurse to combine direct care with clinical responsibility is not often addressed.
- **Lack of real-world evaluation**: external validation and RCT-level evidence is often missing. Most studies remain in the conceptual or pilot phase.
- **Technical vs clinical focu**s: Evaluation often prioritises technical model performance over practical impacts on workflow, workload, and the timeliness of escalation.
- **Abstract Explainability**: Many studies report explainability, but it is often presented as technical interpretability (such as feature importance) rather than an actionable explanation that supports bedside decision-making and clinical accountability.
- **Disconnected Implementation Barriers**: The issue of implementation barriers is also not consistently articulated in the context of workflow, documentation, and training in relation to the exact touchpoints and activities in NH settings.