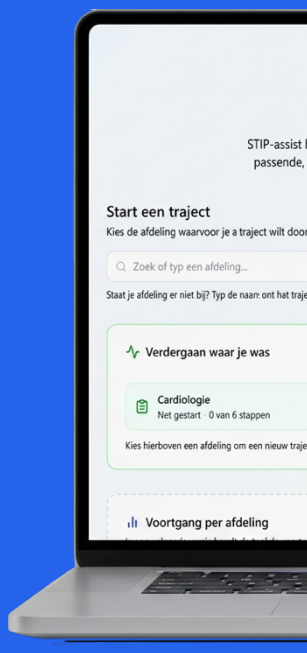
Closing the gap between what we know and what we do.

Designing an implementation tool for infection prevention.



TU Delft

Amsterdam UMC
Universitair Medische Centra

Thibauld
Demaret

Designing an implementation tool for infection prevention.

Introducing **STIP-assist**: a digital, AI-supported tool that helps infection prevention professionals turn behavioural research into tailored interventions for physicians, moving beyond generic campaigns towards context-specific implementation.

Thibauld Demaret

Amsterdam UMC
Universitair Medische Centra

TU Delft

MSc. Graduation Thesis
Strategic Product Design
June 2026

Thibauld Demaret
Strategic Product Design
Faculty of Industrial Design Engineering
Delft University of Technology

In collaboration with
Amsterdam UMC

Chair
Dr. V.P. Pannunzio
Design Value in Ecosystems
Design Organisation and Strategy (DOS)

Mentor
Dr. T. Wang
Human Factors
Human-Centered Design (HCD)

Company mentors
Dr. Rosa van Mansfeld
Miriam Schutte
Amsterdam UMC

"Now you know. And knowing is half the battle!"

~ G.I. Joe: A Real American Hero

Foreword

Dear reader,

When I started this project, I thought infection prevention was mostly a matter of knowing the rules. Wash your hands, wear your gloves, follow the guideline. If compliance is low, surely the answer is to remind people better, to teach them more clearly, to run another campaign.

The deeper I went, the more I realised how wrong that assumption was. The strategies to improve infection prevention already exist, and the people working in hospitals already know them. What stands in the way is something far more human: time pressure, habit, hierarchy, the culture of a particular ward, the quiet question of whether a colleague will do the same. Infection prevention turned out to be far less about knowledge and far more about behaviour, context, and the messy reality of clinical work.

That shift in understanding is what this project is really about. As a strategic designer, I am drawn to complex, human-centred problems that can make a real difference in practice. My Master's in Strategic Product Design, together with the Medisign specialisation, taught me to approach challenges like this from both a strategic and a design perspective. This graduation project gave me the opportunity to apply those skills in a context that truly matters, where better design can ultimately contribute to patient safety.

I am grateful to everyone at Amsterdam UMC who shared their time, expertise, and honest opinions with me, and to my supervisors at TU Delft for their guidance and support throughout the process.

I hope this report invites you to see infection prevention as I came to see it: not as a matter of telling people what to do, but as a design challenge worth solving.

Thibauld Demaret

Summary

Healthcare-associated infections and antimicrobial resistance are a major threat to patient safety. The everyday measures that prevent them, such as hand hygiene and the correct use of personal protective equipment, are applied inconsistently in hospitals. Physicians in particular are often identified as poorer compliers. The cause is rarely a lack of knowledge. Compliance is shaped by time pressure, habit, hierarchy, and the culture of a specific ward. It also varies a lot between departments within the same hospital. As a result, hospital-wide campaigns and clinical lessons rarely produce lasting change. What works instead is tailoring: matching interventions to the behavioural determinants at play in a given context.

Within the STIP (*Strategieën op maat tegen verspreiding van AMR door toepassing van infectiepreventie door artsen)* programme at Amsterdam UMC, the strategies needed to improve physician compliance have already been developed. They were co-developed with end users and grounded in established frameworks such as the Behaviour Change Wheel. What has been missing is a way to bring these strategies into practice. Infection prevention professionals currently have no structured means of deciding which strategy fits a particular ward, specialty, or barrier. This project therefore frames the problem differently. The challenge is not to develop new strategies, but to translate existing content from theory into a practical tool. That tool should support tailored strategy selection and fit the realities of clinical work.

The project follows an iterative, human-centred design approach. It began with a context analysis that combined desk research and interviews with infection prevention professionals. This made the user needs and contextual barriers visible that a supporting tool would have to address. The literature and interviews pointed to the same direction of effective change. Feedback works best when it is individualised and specialty-specific, rather than aggregated. Local physician champions can play an important role in successful implementation. Effective implementation involves identifying barriers, deciding which are most important, selecting suitable strategies, and creating an action plan. These insights were used to develop design requirements and explore different tool concepts.

The result is STIP-assist. It is a digital, AI-supported decision aid that guides infection prevention professionals through this sequence themselves. It moves from a contextual analysis of barriers, to a set of matched and explained strategies, to a concrete action plan. It does this without depending on external experts. In this way, it puts ideas from the literature into practice in a more consistent way. This report describes what as happened during the 20 weeks of this project. It covers the problem framing, the research and design insights behind the tool, the current prototype, and the steps planned towards user testing and final delivery.

Terminology & abbreviations

AMR (Antimicrobial Resistance) – The ability of micro-organisms to withstand antimicrobial treatment, reducing the effectiveness of antibiotics and increasing the risk of untreatable hospital-acquired infections.

STIP – *Strategieën op maat tegen verspreiding van AMR door toepassing van infectiepreventie door artsen* (tailored strategies against the spread of AMR through the application of infection prevention by physicians). The broader research programme at Amsterdam UMC within which this graduation project is positioned.

DI (Deskundige Infectiepreventie) – Infection prevention professional; the specialist responsible for advising departments on, and implementing, infection prevention practices. The primary intended user of the tool developed in this project.

PPE (Personal Protective Equipment) – Protective items such as gloves, gowns and masks used by healthcare workers to prevent transmission of micro-organisms.

Compliance / adherence – The extent to which healthcare workers actually follow infection prevention measures (e.g. hand hygiene) in daily practice. Distinct from *implementation*, which refers to the organisational process of introducing and embedding a measure or strategy.

Implementation strategy – A planned, theory-informed approach intended to improve the uptake of an infection prevention measure (for example, audit and feedback, role modelling, or environmental restructuring).

Tailored vs. personalised – In this report, *tailored* refers to strategies matched to the determinants of a specific context or group (e.g. a ward or specialty), whereas *personalised* would refer to the individual level. This project operates primarily at the tailored (group/context) level.

COM-B – A behaviour model stating that behaviour (B) results from the interaction of Capability, Opportunity and Motivation (Michie et al., 2011). The core of the Behaviour Change Wheel.

BCW (Behaviour Change Wheel) – A framework for designing behaviour change interventions, built around the COM-B model (Michie et al., 2011).

TDF (Theoretical Domains Framework) – A framework of behavioural determinants used to categorise the factors influencing a target behaviour, frequently applied in infection prevention research.

DCE (Discrete Choice Experiment) – A quantitative method used to obtain stakeholders' preferences by asking them to choose between alternatives with varying attributes; used within STIP to study physicians' preferences for strategy attributes.

Determinant – A factor (e.g. social norms, time pressure, risk perception) that influences whether a target behaviour occurs; the basis for matching tailored strategies.

ItFits-toolkit – An existing implementation toolkit (structured around the modules identify, match, design, and apply & review) used as a reference for the design of the STIP tool.

Table of Contents

Foreword *6*

Summary *7*

Terminology & abbreviations *8*

1 Project background & framing *12*

1.1 Introduction 12

1.2 Context 12

Figure 1. *Key stakeholder map* 13

1.3 Scope 13

1.4 Project aim and research objectives 14

1.5 Approach 14

2 Literature review *15*

2.1 Infection prevention and compliance problem 15

2.2 Determinants of physician (non-)compliance 16

2.3 Different IPC behaviours require different approaches 17

2.4 Theoretical frameworks for behaviour change 18

2.5 Tailored implementation and self-guided toolkits 19

2.6 Synthesis and implications for design 21

3 Interviews & design requirements *22*

3.1 Method in brief 22

.3.2 Findings 23

3.3 Stakeholders 24

3.4 Design requirements 25

3.5 Scope and delivery constraints 27

4 Brainstorm, sketches & design process *29*

4.1 From requirements to ideas 29

4.2 Clustering the ideas 30

4.3 From clusters to two concepts 31

5 Two concepts: comparison and development *32*

5.1 Concept 1: STIP-assist **32**

5.2 Concept 2: Implementation Co-Creation Toolkit **33**

5.3 Comparing the two **35**

5.4 Developing concept 1 ..... **36**

6 Prototype, feedback and feasibility *37*

6.1 The first prototype **37**

6.2 Feedback from stakeholders **41**

6.3 Feasibility **43**

6.4 The last iteration **44**

7 Final evaluation *55*

7.1 Evaluation approach **55**

7.2 Evaluation against the requirements **55**

7.3 Practicality and relevance **57**

7.4 A usability test ready to run **57**

8 Roadmap *59*

8.1 The journey, horizon by horizon **60**

Horizon 0 — Laying the foundation (current stage) 60

Horizon 1 — Pilots (starting now) 62

Horizon 2 — Scaling for wider adoption (mid-term, indicative 2028) 64

Horizon 3 — National adoption and long-term sustainability (long-term, indicative 2030 and onwards) 66

8.2 The five tracks of the roadmap **69**

8.3 Strategic value of the roadmap **72**

9 Reflection, recommendations and contribution to STIP *73*

9.1 Reflection **73**

9.2 Limitations **73**

9.3 Recommendations **74**

9.4 My contribution to STIP **75**

References *76*

List of figures *82*

Appendix *84*

List of appendices **84**

Appendix A. Project Brief Amsterdam UMC **85**

Appendix B Interviews **86**

Interview with infection prevention professional 86

Interview with observator infection prevention 88

Interview with an infection prevention professional 93

Interview with creator of ItFits toolkit 95

Interview PhD researcher AI in Healthcare 97

Meeting with data protection officer 99

Meeting with an AI architect 100

Appendix C **102**

Pilot particpating hospital 102

Appendix D **104**

VHIG Congress 104

Appendix E **106**

Mentimeter VHIG congres 106

Appendix F List of design requirements **108**

Appendix G Test protocol **110**

Appendix H Strategic roadmap **119**

..... **120**

1 Project background & framing

1.1 Introduction

Infection prevention in hospitals is essential to protect patients from harm. Insufficient infection prevention contributes to healthcare-associated infections and to the spread of antimicrobial resistance (AMR), which together form a great threat to patient safety (Huis et al., 2012). In the WHO European Region alone, an estimated 674,000 deaths in 2019 were associated with AMR, and around 8% of hospitalised patients acquire at least one healthcare-associated infection (Albers et al., 2025).

Despite this, infection prevention measures such as hand hygiene and the use of personal protective equipment (PPE) are often difficult to apply consistently in daily practice. Compliance among healthcare workers is suboptimal, and physicians in particular are repeatedly identified as poorer compliers, with an average compliance of around 45% compared with 52% among nurses (Bredin et al., 2022; Squires et al., 2014). This is driven less by a lack of knowledge than by contextual and behavioural factors: compliance varies significantly between wards within the same hospital, reflecting the strong influence of local culture, workflows, and professional norms (Cantrell et al., 2008; Schutte et al., 2024). Traditional interventions such as clinical lessons and campaigns rarely achieve substantial and lasting effects on compliance (Huis et al., 2012).

These findings highlight the need for implementation strategies that are tailored to the local context rather than applied through a one-size-fits-all approach (Albers et al., 2023). Within the STIP programme, the behaviour change strategies needed to improve physician compliance have already been co-developed with end users and grounded in theoretical frameworks such as the Behaviour Change Wheel (Michie et al., 2011; Schutte et al., 2025). However, infection prevention professionals currently lack a practical way to select the most appropriate strategy for a specific ward or behavioural barrier. The design challenge of this project is therefore not to develop new strategies, but to translate existing knowledge into a practical implementation tool that supports tailored strategy selection and helps infection prevention professionals improve physician compliance.

This report describes the development of that implementation tool. It begins with the project context, followed by the research and design process, the resulting concept, and its evaluation.

1.2 Context

This graduation project was conducted at the Department of Medical Microbiology and Infection Prevention at Amsterdam UMC as part of the STIP programme (Strategieën op maat tegen verspreiding van AMR door toepassing van infectiepreventie door artsen). The programme aims to

improve physician compliance with infection prevention measures through tailored, theory-based implementation strategies.

The project involves three main stakeholder groups. Infection prevention professionals are the primary users of the implementation tool and support clinical departments in improving infection prevention practice. Physicians are the target group whose behaviour the selected strategies aim to influence, while clinical microbiologists provide clinical expertise and support implementation. Figure 1 presents an overview of the key stakeholders involved in the project.

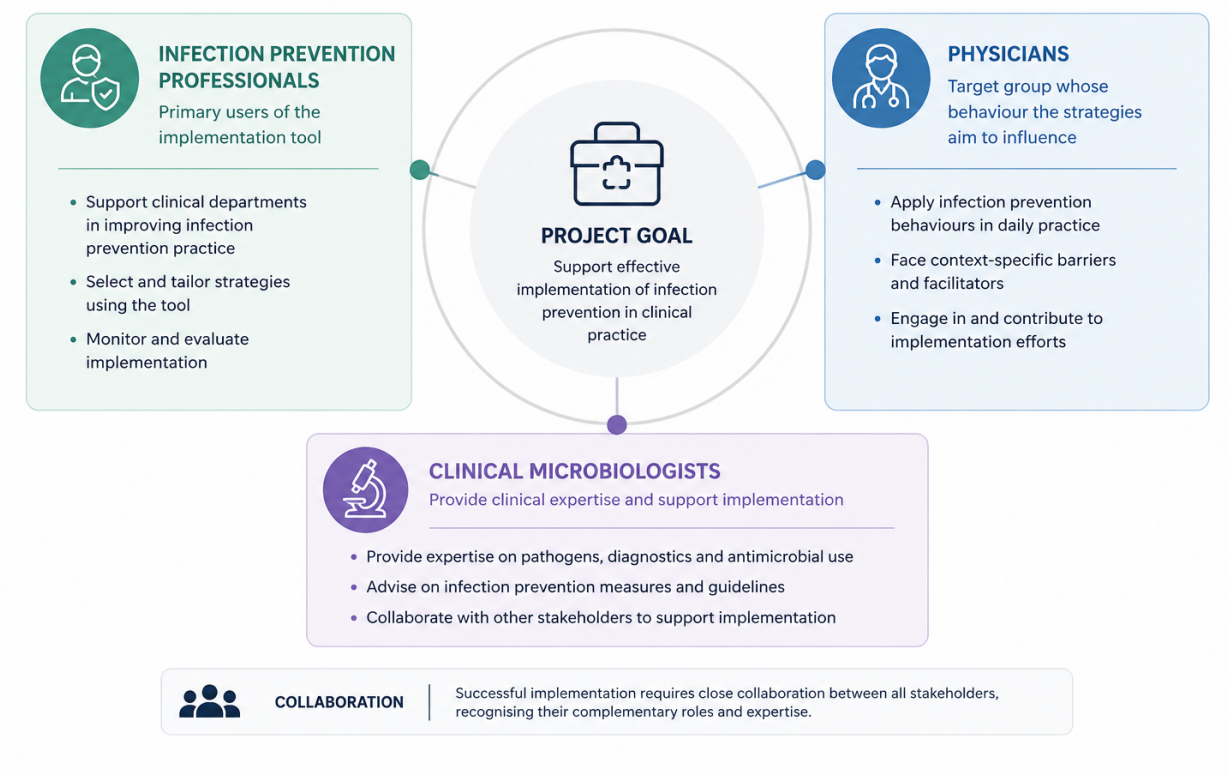


Figure 1. *Key stakeholder map*

1.3 Scope

This project focuses on the design of an implementation tool that supports the selection of tailored infection prevention strategies. It does not address the development of new behaviour change strategies, as these have already been created within the STIP programme. The project is conducted within the context of Dutch academic hospitals, with infection prevention professionals as the primary users of the tool. Within the available 100 working days, the project focuses on user research, concept development, and iterative prototyping and testing. Full clinical implementation and large-scale validation are outside the scope of this project and are addressed in the strategic roadmap.

1.4 Project aim and research objectives

Project aim

To support infection prevention professionals in selecting tailored implementation strategies that improve physician compliance and ultimately strengthen patient safety.

Design statement

Design a practical decision-support tool that translates existing, theory-based infection prevention strategies into a usable resource, helping infection prevention professionals select the right strategy for a specific ward, specialty, or barrier

.

Research objectives

- To understand the needs, workflows, and decision-making of infection prevention professionals when selecting and applying strategies.
- To identify the contextual and behavioural factors that influence physician compliance across different wards and specialties.
- To translate the existing STIP strategies into a structure that supports tailored strategy selection in practice.
- To develop and iteratively test tool concepts with intended users, evaluating them against the criteria of practicality and relevance.

1.5 Approach

This project followed an iterative, human-centred design approach. Figure 2 provides an overview of the project phases. The process started with a context analysis based on desk research, document analysis, and interviews with infection prevention professionals, toolkit and software developers, clinical microbiologists and implementation researchers. The resulting insights were translated into design requirements and concept directions, which were iteratively developed into prototypes and refined through user testing and feedback sessions.

Throughout the project, methods from the Delft Design Guide were selected to support different phases of the design process, including stakeholder mapping, journey mapping, and value proposition framing. The final concept was evaluated with infection prevention professionals to assess its usability, relevance, and applicability within the hospital context. The evaluation criteria are described in Chapter 7.



Figure 2. *Overview of the design process followed during the project*

2 Literature review

This chapter reviews the literature that forms the foundation of the present project. It first examines the infection prevention compliance problem and the behavioural and contextual factors that influence physician behaviour. It then discusses why different infection prevention behaviours require tailored approaches and introduces the theoretical frameworks used to understand behaviour and select appropriate implementation strategies. Building on these foundations, the chapter explores how tailored implementation is applied in practice through self-guided implementation toolkits. Finally, the literature is synthesised into a set of design implications that guide the development of STIP-assist.

2.1 Infection prevention and compliance problem

To understand why an implementation tool is needed, it is first necessary to examine the nature of the infection prevention problem and the challenges of achieving consistent compliance in clinical practice.

Infection prevention and control (IPC) practices, such as hand hygiene and the use of personal protective equipment (PPE), are among the most effective measures to prevent the transmission of micro-organisms and reduce healthcare-associated infections (Huis et al., 2012). Despite this, compliance with these measures in daily hospital practice is consistently low. A systematic review of 96 studies found an overall median hand hygiene compliance rate of 40%, with compliance significantly lower among physicians (32%) than among nurses (48%), and lowest of all before patient contact (21%) (Erasmus et al., 2010). A more recent meta-analysis confirmed this pattern,

reporting an overall compliance rate of 45% for physicians compared with 52% for nurses (Bredin et al., 2022). Physicians are therefore repeatedly identified as poorer compliers than other healthcare workers (Squires et al., 2014).

Compliance is not only low on average but also highly variable. The same systematic review found that compliance was lowest during periods of high clinical activity and in situations involving physicians (Erasmus et al., 2010). Compliance also differs considerably between hospital wards. Cantrell et al. (2008) observed physician hand hygiene compliance ranging from 4% in one ward to 96% in another within the same hospital, attributing this variation primarily to differences in local ward culture rather than individual or structural factors. This contextual variation is important to the present project, as it suggests that the barriers to compliance and therefore the strategies required to address them differ between clinical settings.

These findings reveal a gap between evidence and practice. Importantly, this gap is not primarily explained by a lack of knowledge or guidelines, but by behavioural and contextual factors that influence whether healthcare workers apply them in practice. Understanding these factors is therefore essential for designing effective implementation strategies and is discussed in the following section.

2.2 Determinants of physician (non-)compliance

If low compliance is not primarily a knowledge problem, the question becomes what does drive physician behaviour? A scoping review of 56 studies mapped the reported determinants onto the Theoretical Domains Framework (TDF), identifying environmental context and resources, social influences, beliefs about consequences, memory and decision-making, knowledge, and skills as the most prominent domains (Schutte et al., 2024). Interestingly, the single most frequently reported group of determinants fell outside the TDF: socio-demographic factors such as seniority, specialty, and sex. Though generally considered unmodifiable, their strong influence is thought to reflect an underlying role of local culture (Schutte et al., 2024). An interview study with 42 physicians and residents confirmed this breadth, identifying determinants across nine TDF domains (Squires et al., 2014). Recurring barriers include high workload, time pressure, limited access to materials, and a culture in which non-compliance is normalised, while individualised feedback, role models, and a supportive ward culture act as enablers (Erasmus et al., 2010; Squires et al., 2014).

Although these determinants are broad, several studies show that their relative importance differs between clinical contexts and infection prevention behaviours. Determinants also differ by specialty and measure. A COM-B–based survey of orthopaedic physicians, who must comply with more than twenty separate surgical-site-infection measures, found capability and planning to be the strongest correlates of compliance, while motivation, though high, was not decisive (Tomsic et al., 2021). This reinforces a recurring theme: physicians are generally motivated to prevent infection, so interventions targeting motivation alone are unlikely to close the gap. The more decisive levers lie in

capability, planning, and context. Qualitative work adds that decisions are also shaped by visibility, social expectation, and habit. Practices are sometimes performed as much for how they appear to others as for their clinical rationale (Jackson et al., 2013; Lee, 2013).

While previous studies identify a broad range of determinants, recent work within the STIP programme provides more detailed insight into the behavioural mechanisms underlying physicians' decisions. A qualitative study within the STIP programme examined how physicians actually make these decisions. It found heterogeneity in how physicians relied on heuristics and habit, others on deliberate reasoning. These choices were further shaped by personal beliefs and the social and physical environment (Schutte et al., 2025). Of 26 mechanisms of action that can underlie behaviour, 18 were matched to physicians' IPC behaviour. Six stood out as the most promising intervention targets: behavioural cueing, beliefs about consequences, values, social norms, social learning and imitation, and environmental context (Schutte et al., 2025).

Taken together, physician compliance depends less on what physicians know than on how they decide and the context they work in. Because the dominant determinants are social, contextual, and cultural, and because they vary between wards, specialties, and individual measures, interventions cannot be uniform. They must be matched to the determinants that actually play a role in a given setting. The next sections discuss how the theoretical and tailored approaches make this possible.

2.3 Different IPC behaviours require different approaches

Although hand hygiene is the most frequently studied infection prevention behaviour, it should not be assumed that the same behavioural mechanisms apply to every IPC measure. Different infection prevention behaviours are influenced by different determinants and may therefore require different implementation strategies. Glove use provides a clear example of this. While gloves are intended to reduce infection risk, incorrect glove use such as failing to change gloves between procedures or wearing them when not indicated creates missed opportunities for hand hygiene and can increase the risk of cross-transmission (Loveday et al., 2013).

The extent of inappropriate glove use is substantial. In an observational study of 163 glove-use episodes, glove use was inappropriate in 42% of cases, with gloves worn unnecessarily for low-risk procedures in 92% of those instances; in 37% of episodes there was a genuine risk of cross-contamination (Loveday et al., 2013). Crucially, the drivers of this behaviour differ from those of hand hygiene. Glove use was found to be shaped by socialisation and emotion, with disgust and fear as key motivators. Other drivers included the assumption that patients expect gloves, confusion about when gloves are indicated, and social norms and peer pressure (Loveday et al., 2013). A survey across two healthcare organisations confirmed this pattern: the primary motivation for wearing gloves was personal protection rather than patient safety. Staff were influenced by a desire to project professionalism, and they tended to follow the lead of senior colleagues in their own profession (Flores et al., 2020).

This has two important implications for the project. First, it demonstrates that infection prevention behaviours cannot be treated as interchangeable. A decision-support tool that applies the same implementation strategy to every behaviour risks overlooking the specific determinants that drive that behaviour. Second, it reinforces the central argument of this review: improving infection prevention requires strategies that are tailored not only to the local context, but also to the specific behaviour being targeted.

2.4 Theoretical frameworks for behaviour change

The previous sections showed that improving infection prevention is primarily a matter of changing behaviour rather than increasing knowledge. To change behaviour systematically, it first needs to be understood using established behaviour change theories. However, a recurring criticism of infection prevention research is that interventions are often developed without an explicit theoretical foundation (Schutte et al., 2024). Several well-established frameworks from behaviour change science provide a structured way to understand behaviour and design tailored interventions.

The COM-B model is one of the most widely used frameworks for understanding and changing behaviour. It proposes that behaviour (B) results from the interaction between three components: capability, opportunity, and motivation (Michie et al., 2011). Capability refers to an individual's physical and psychological ability to perform the behaviour, opportunity to the physical and social environment that enables it, and motivation to the reflective and automatic processes that drive behaviour. COM-B forms the core of the Behaviour Change Wheel (BCW), which links these behavioural determinants to appropriate intervention functions and policy categories, providing a structured approach for designing behaviour change interventions (Michie et al., 2011). Figure 3 presents the Behaviour Change Wheel, with the COM-B model at its core.

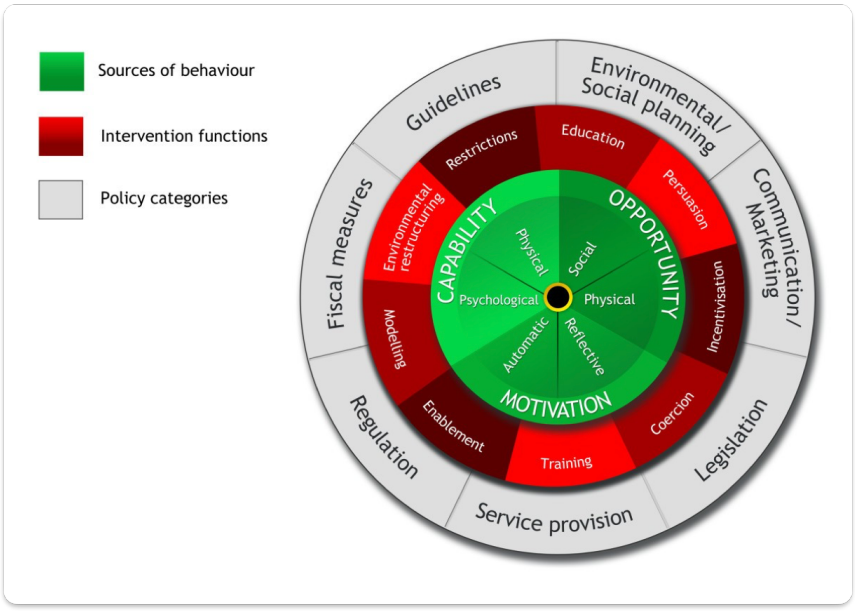


Figure 3. *The Behaviour Change Wheel (Michie et al., 2011)*

To gain a more detailed understanding of the determinants of behaviour, the Theoretical Domains Framework (TDF) is often used alongside COM-B. The TDF synthesises constructs from multiple behaviour change theories into fourteen domains, including knowledge, skills, social influences, environmental context, and beliefs about consequences, providing a comprehensive framework for identifying barriers and enablers of behaviour (Atkins et al., 2017). The TDF domains map directly onto the COM-B components, making the two frameworks complementary. While the TDF helps identify the specific factors influencing behaviour, COM-B and the Behaviour Change Wheel translate these insights into appropriate intervention functions and behaviour change strategies.

A key challenge is translating behavioural determinants into appropriate intervention strategies. Early research addressed this by systematically mapping behavioural determinants to evidence-based behaviour change techniques (Michie et al., 2008). This work evolved into structured tools that connect mechanisms of action to specific behaviour change techniques. One example is the Theory and Techniques Tool, which is used within the STIP programme to support evidence-based strategy selection (Schutte et al., 2025). Rather than relying on intuition, these tools provide a transparent and grounded in theory to move from identifying what needs to change to selecting how to change it. More recently, this logic has been embedded in digital decision-support tools. For example, Kelly et al. (2024) developed a co-designed prototype that coded behavioural barriers using COM-B and linked them to Behaviour Change Wheel intervention functions, demonstrating that these frameworks can be successfully operationalised in an interactive decision-support tool.

These frameworks provide the theoretical foundation for both the STIP programme and the present project. Together, they offer a structured approach for understanding behaviour, identifying its underlying determinants, and selecting appropriate implementation strategies. They also explain why tailoring is essential: because behaviour is shaped by a specific combination of capability, opportunity, and motivation within a given context, effective implementation strategies must be matched to that context rather than applied uniformly. The next section examines how this principle of tailoring is translated into implementation practice.

2.5 Tailored implementation and self-guided toolkits

The previous section described the theoretical frameworks used to understand behaviour and select appropriate intervention strategies. The next challenge is translating these frameworks into implementation practice. Tailored implementation addresses this challenge by identifying the barriers and enablers within a specific context and selecting strategies that target those determinants, rather than applying the same intervention across all settings.

A Cochrane review provides the strongest evidence supporting this approach, reporting that tailored interventions were 1.56 times more likely to improve professional practice than non-tailored interventions (von Lengerke et al., 2023; Albers et al., 2025). However, the evidence should be interpreted with caution. The reported effects are generally small to moderate, and most studies

have evaluated interventions aimed at changing clinicians' behaviour rather than supporting the infection prevention professionals responsible for implementing those interventions (Vis et al., 2022).

This distinction is important for the present project. Von Lengerke et al. (2023) distinguish between first-order interventions, which target clinicians directly to improve compliance, and second-order interventions, which support infection prevention professionals in delivering those interventions. STIP-assist belongs to the latter category. It is designed to support infection prevention professionals in selecting tailored implementation strategies, while the behaviour change logic embedded within the tool ultimately targets physicians' infection prevention behaviour.

Most tailored implementation approaches currently rely on experienced implementation experts to identify barriers and match strategies. Although effective, this expert-driven model presents practical challenges, including limited scalability, higher costs, and reduced sensitivity to local context (Vis et al., 2022). Self-guided implementation toolkits have therefore been proposed as an alternative. The best-studied example is the ItFits-toolkit, developed within the ImpleMentAll project. It structures the tailoring process into four iterative modules: identifying barriers, matching strategies, designing an implementation plan, and evaluating progress, while encouraging users to work pragmatically, focus on one barrier at a time, and adapt their approach when necessary (Finch et al., 2024).

Evidence suggests that this self-guided approach is feasible and acceptable. A controlled trial found that the ItFits-toolkit resulted in a small but significant improvement in the normalisation of the target intervention, without increasing the time required from implementers (Vis et al., 2022). Process evaluations further showed that implementation teams were able to carry out tailored implementation independently, although participants consistently identified the tailoring process as detailed and time-intensive (Finch et al., 2024).

These findings provide important implications for STIP-assist. First, they demonstrate that self-guided tailored implementation is feasible, provided the process remains practical and manageable. Second, they highlight the importance of reducing complexity rather than adding to it. Similar conclusions were reached in the co-design study by Kelly et al. (2024), where users preferred plain language, a clear step-by-step workflow, and guidance in translating behavioural analyses into implementation strategies. These insights directly informed the design of STIP-assist. Finally, evidence from a cluster-randomised trial showed that interventions tailored to identified psychological determinants resulted in sustained improvements in physician hand hygiene compliance and reductions in multidrug-resistant organism infections compared with a standard untailored campaign (von Lengerke et al., 2019). Together, these findings support both the rationale for tailored implementation and the development of self-guided decision-support tools such as STIP-assist.

2.6 Synthesis and implications for design

The literature reviewed in this chapter provides a clear foundation for the design of STIP-assist. Together, the studies demonstrate that physician compliance is primarily influenced by behavioural and contextual factors, that implementation strategies should therefore be tailored to local determinants, and that structured decision support can help translate implementation science into daily practice. Table 1 summarises the main findings from the literature and the resulting design implications for this project.

Finding from literature	Supporting literature	Design implication for STIP-assist
Physician compliance is primarily influenced by behavioural and contextual factors rather than knowledge alone.	Erasmus et al. (2010); Squires et al. (2014); Schutte et al. (2024)	Provide structured identification of behavioural barriers.
Behavioural determinants differ between wards, specialties, and IPC behaviours.	Cantrell et al. (2008); Loveday et al. (2013); Flores et al. (2020)	Create context-specific rather than generic recommendations.
COM-B, TDF and the BCW provide a structured framework for diagnosing behaviour and selecting interventions.	Michie et al. (2011); Atkins et al. (2017)	Structure the tool around behavioural diagnosis and strategy selection.
Theory and Techniques Tool links behavioural determinants to implementation strategies.	Michie et al. (2008); Schutte et al. (2025)	Base recommendations on evidence-informed strategy matching.
Tailored implementation is more effective than generic implementation.	von Lengerke et al. (2023); Albers et al. (2025)	Recommend tailored strategies rather than one-size-fits-all interventions.
Self-guided implementation toolkits are feasible but should minimise effort and complexity.	Vis et al. (2022); Finch et al. (2024); Kelly et al. (2024)	Design a simple, guided workflow using clear language and step-by-step support.

Table 1. *Main research findings*

The synthesis highlights a clear gap between behavioural theory and implementation practice. While existing research provides strong theoretical frameworks and evidence-based implementation strategies, infection prevention professionals currently lack practical decision support to apply these insights consistently in daily practice. STIP-assist addresses this gap by guiding users through barrier identification, tailored strategy selection, and implementation planning. The resulting design implications form the foundation for the design requirements presented in the following chapter.

3 Interviews & design requirements

Chapter 2 showed that physician compliance is a behavioural and contextual problem, and that interventions only work when they are tailored to the local situation. To see what this means inside Amsterdam UMC and to find out what a supporting tool would actually need to do, I interviewed infection prevention professionals and others involved in their work. This chapter reports what they told me, grouped by theme, maps the stakeholders around the tool, and turns the findings into the design requirements for the toolkit.

3.1 Method in brief

To better understand current infection prevention practice and user needs, qualitative data were collected through semi-structured interviews and contextual observation. As summarised in Table 2, four interviews were conducted with key stakeholders: two infection prevention professionals (Deskundigen Infectiepreventie), one hand hygiene observer, and one of the developers of the ItFits-toolkit. These participants were selected based on their expertise and direct involvement in infection prevention implementation. In addition, an observational visit was conducted during a pilot at one of the participating hospitals. To protect participants' privacy, all interview findings are presented anonymously throughout this report.

The interviews focused on how infection prevention works in practice, which behavioural and implementation challenges keep recurring, and what a supporting tool should provide. The interview data were analysed by identifying and grouping recurring observations into themes, which formed the basis for the design requirements presented later in this chapter.

Method	Participants	Purpose
Semi-structured interviews	2 infection prevention professionals	Understand workflows, barriers and user needs
Semi-structured interview	1 hand hygiene observer	Understand observation practices and implementation challenges
Semi-structured interview	1 ItFits developer	Understand self-guided implementation toolkits
Contextual observation	Pilot at Ziekenhuis Gelderse Vallei	Observe pilot study

Table 2. *Interview overview*

Several additional expert consultations were conducted on the technical feasibility, privacy, and architecture of an AI-supported concept. As these discussions concerned implementation constraints rather than user needs, they are reported separately in Chapter 6.

.3.2 Findings

The hospital measures a lot, but struggles to turn data into action. One DI described monitoring tools, including Power BI dashboards that break hand hygiene compliance down by department, professional group, and indication. Overall compliance sits around 55 to 60%, with a lot of variation underneath. But turning that data into a concrete strategy still depends almost entirely on the professional's own interpretation. The same DI sits at the centre of a wide network, coordinating between physicians, nurses, departmental contact persons, observers, microbiologists, and researchers. The other DI made a related point: infection data is scattered across departments and not linked to behaviour, so it is hard to tell what actually works. The problem is not a lack of data. It is the missing step between knowing there is a problem and knowing what to do about it. This is the same gap the literature pointed to in Section 2.6.

Feedback does not reach the people who could act on it. The observer described the digital registration system (K2) as slow and difficult to use, with no feedback feature. Observers never see the overall results, and departments cannot see their own scores. Monitoring ends up being used mainly for reporting, so the chance to learn from it is lost. The literature treats timely, individualised feedback as a behaviour change tool in its own right (Section 2.6), which makes this a real missed opportunity.

Every department is different. Both the observer and one DI stressed that infection prevention is not applied uniformly. On the ICU, computers count as part of the patient environment, which adds hand hygiene moments. Some departments, like the NICU, use stricter rules than the official guidelines. Even the observation schedule is adapted in practice, spread across the year instead of run rigidly. Tailoring already happens informally. A tool therefore has to support context-specific choices that fit each setting.

Behaviour follows perceived risk, not knowledge. One DI explained that visibly dirty surfaces get cleaned, but invisible contamination does not, because the risk does not feel real. The observer saw the same logic with gloves: people feel protected, so they walk from room to corridor without disinfecting. Phones get handled right after patient contact, and the phones themselves are rarely cleaned. These are habits and risk misperceptions, not gaps in knowledge, which matches the glove-use literature in Section 2.3.

Physicians behave differently from nurses. Some of the strongest findings were about physicians. During ward rounds, compliance followed the senior consultant: if the lead physician disinfected, the rest did too; if not, compliance dropped. One DI described physicians as harder to reach through the usual channels, more likely to question whether a protocol applies, and more reliant on their own judgement. At the pilot, the physician I observed already saw himself as engaged and doing well, while staying open to new ideas. Together this shows that reaching physicians is about social modelling, autonomy, and framing, and needs different channels than reaching nurses.

A useful tool guides, it does not prescribe. The ItFits developer gave the clearest lessons about form. The hard part of building ItFits was making abstract implementation theory usable. Real users rarely work in a straight line: many think they already understand the problem and want to jump to solutions. So ItFits lets people enter at different points, simplifies the language, and guides decisions without forcing fixed steps, an approach grounded in supporting local implementers to identify barriers, match them to strategies, and build an actionable plan while keeping the decisions with the team (Vis et al., 2022; Finch et al., 2024). The developer also noted several points. People use these tools during specific phases, not continuously. The tool has to fit existing routines. And simplifying the theory cannot come at the cost of its credibility.

Engagement helps adoption. One DI found interactive formats, like black light demonstrations, cleaning markers, and an infection prevention escape room, far more effective than clinical lessons given to tired staff at the end of a shift. They got people involved and kept the conversation going afterwards. The downside is that they cost time and resources, which limits how far they scale. So engagement is worth designing for.

3.3 Stakeholders

Infection prevention involves many actors, so before defining requirements I mapped them in two ways. The first map (Figure 4) shows how they relate. The DI sits at the centre as the linking pin: advising physicians and nurses, receiving policy responsibility from the clinical microbiologist, and exchanging data and feedback with researchers about the research that has to be implemented. The observer informs the DI and observes physicians and nurses during rounds. Physicians and nurses, in turn, deliver care to patients at the end of the chain.

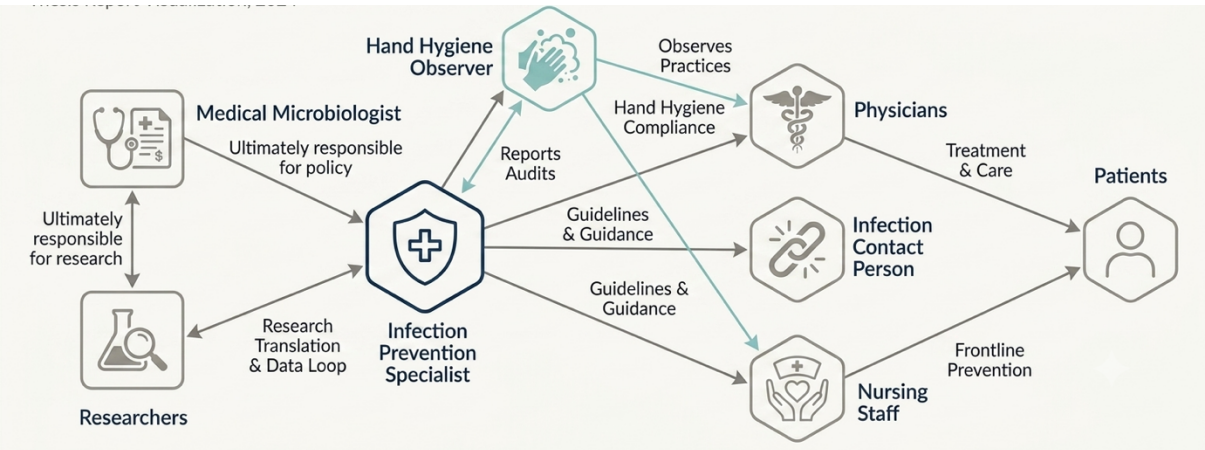


Figure 4. *Stakeholder relationship map*

The second map (Figure 5) ranks the actors by how central they are to the tool, in three rings: core, direct, and indirect. This places the DI in the core as the primary user, with physicians, nurses, contact persons, observers, microbiologists, and researchers in the surrounding rings. Although many stakeholders are involved in infection prevention practice, the infection prevention professional is the primary user of the toolkit, while physicians represent the primary behaviour change target group. This distinction runs through the rest of the project: the toolkit is designed for the DI, in order to change the behaviour of physicians.



Figure 5. *Stakeholder rings: core, direct, indirect*

3.4 Design requirements

The findings above, together with the theory from Chapter 2 and input from the project's stakeholders, lead to the following requirements (Table 3). They are grouped into six themes. DR13 was not in the original set, but the interviews pointed to it clearly enough to add.

#	Requirement	Description	Source
Context			
1	Fits a busy clinic	Usable in time-constrained settings	Interviews (IP professionals, observer)
2	Adapts to clinical departments	Supports adaptation to departments via context analysis.	Interviews (IP professionals, observer)

User			
3	Guides strategy selection	Helps professionals select appropriate strategies.	Interview (IP professional)
4	Structures reflection	Supports structured reflection; judgement stays with user.	Interview (ItFits developer)
Behaviour change			
5	Surfaces determinants	Helps identify determinants behind non-compliance.	Literature (COM-B, TDF)
6	Matches to strategies	Matches determinants to suitable strategies.	Literature (COM-B, TDF)
7	Makes consequences visible	Increases visibility of behavioural consequences.	Literature + interview (IP professional)
Implementation process			
8	Works iteratively	Supports an iterative rather than fixed process.	ItFits literature + interview (developer)
9	Supports collaboration	Supports working with multiple stakeholders.	Interviews (IP professionals)
Usability			
10	Needs no training	Understandable without extensive training.	Stakeholder (project team)
11	Translates theory	Turns complex theory into intuitive interactions.	Interview (ItFits developer) + stakeholder
Engagement			

12	Is engaging to use	Supports active rather than passive approaches.	Interview (IP professional)
Physician engagement			
13	Tailors physician engagement	Helps tailor how physicians are engaged.	Interviews (IP professionals, observer) + pilot

Table 3. *List of design requirements*

Where each requirement comes from

To keep the reasoning traceable, the table above links each requirement to its main source: a finding from the interviews (Section 3.2), the literature (Chapter 2), or a requirement raised by the project team.

One further piece of existing work feeds directly into the toolkit's content. Within the STIP programme, a discrete choice experiment was conducted to study physicians' preferences for strategy attributes. Its results inform the toolkit's strategy recommendations, allowing them to be tailored more precisely. The experiment shapes the tool's content rather than forming a design requirement in itself.

These requirements also guide the rest of the project. In the later concept and evaluation chapters, each concept is assessed against them to judge how well it meets the needs identified here.

3.5 Scope and delivery constraints

The stakeholders from Amsterdam UMC also set boundaries for the final product. These are less about needs to meet than about the space the design has to fit within:

- The toolkit should be a finished, open-access product that other hospitals can use without paying a developer for support or monthly upkeep. It should not need to be built into hospital systems.
- It should be able to fall back on a basic paper worksheet method. This digital tool is the first deliverable, with room to develop it further under future funding.
- Stakeholders preferred a mix of digital and physical or interactive elements, a preference echoed at the VHIG session.
- The tool must ask little of physicians' time while still meeting their needs. It should add real value, not just be enjoyable for the DI to use.

- Privacy and infrastructure conditions apply as design constraints: avoid sensitive data, and stay deployable across hospitals. The reasoning behind these is covered in Chapter 6.

4 Brainstorm, sketches & design process

With the design requirements in place, the next step was to generate ideas for a tool that could meet them. This chapter shows how I moved from those requirements to concrete concepts: first by generating a broad set of ideas, then by clustering them into directions, and finally by selecting the direction with the most potential to develop further.

4.1 From requirements to ideas

I generated ideas through a mix of methods. The starting point was sketching and brainstorming on paper, working from "how might we" questions derived from the requirements, for example how might we help a DI move from a low compliance number to a concrete action, or how might we make a tailored strategy easy to choose under time pressure. Alongside this, I sparred informally with my housemate and brothers. They are not infection prevention experts, which made them useful: if an idea was hard to explain to an outsider, it was probably not intuitive enough, which connects directly to the usability requirements (DR10, DR11).

This produced around fifteen ideas, ranging from a digital AI chatbot that analyses a conversation and suggests strategies, to a workbook, a decision tree, a card deck, a magnetic mapping board, persona sheets, and a workshop game (Figure 6). The ideas deliberately spanned a wide space, from fully digital to fully physical, and from tools used by a DI alone to tools used together with physicians.

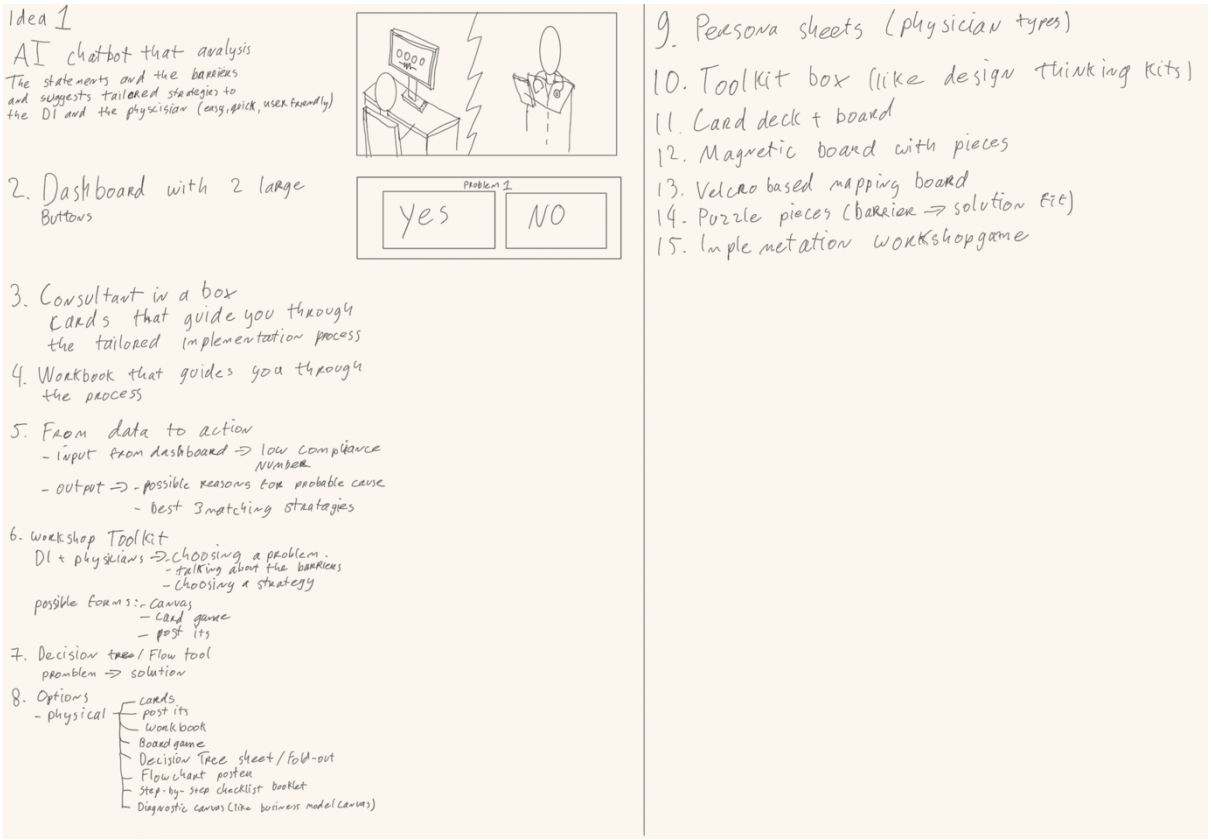


Figure 6. *Idea generation sheet: the fifteen initial ideas*

4.2 Clustering the ideas

To make sense of the fifteen ideas, I grouped them into clusters that shared a similar form and logic (Figure 7). Six directions emerged:

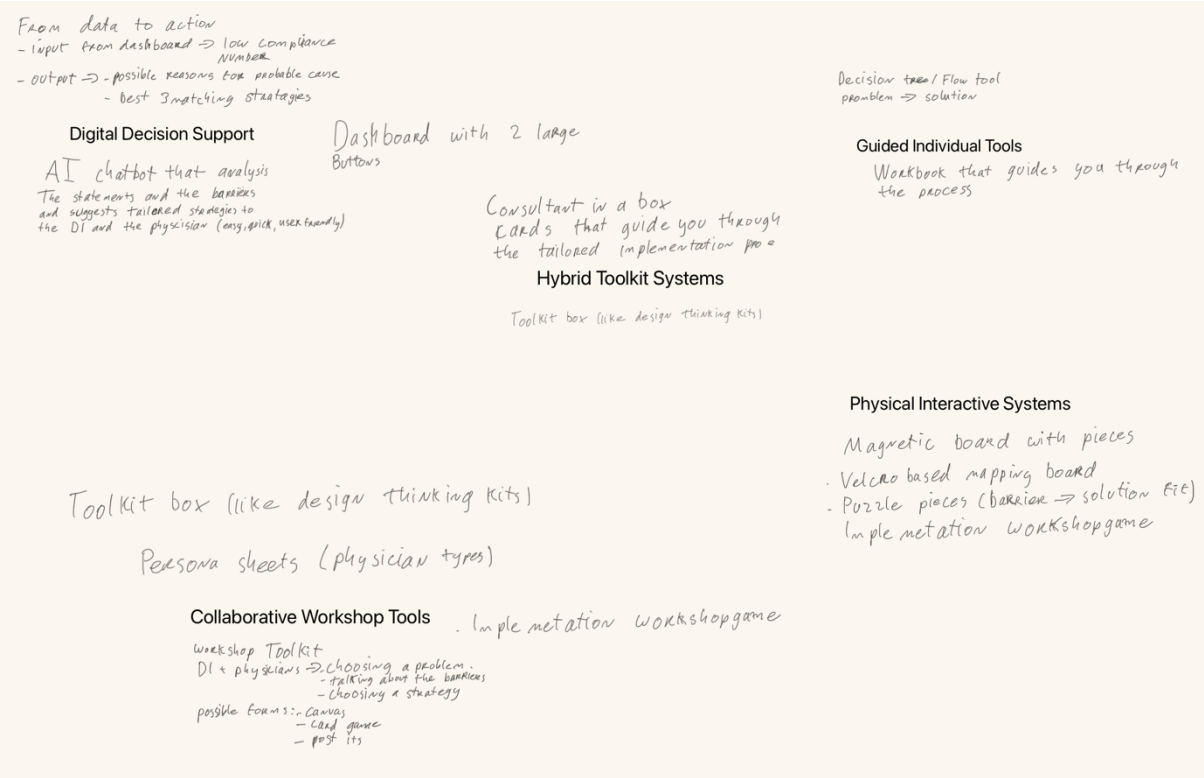


Figure 7. *Clustering of the ideas into six directions*

Digital Decision Support: an AI chatbot, a dashboard, and a "from data to action" tool that turns a low compliance number into probable causes and matched strategies.

Hybrid Toolkit Systems: a physical toolkit box, similar to a design thinking kit, combined with guiding elements.

Guided Individual Tools: a workbook or decision tree that walks a single user through the process.

Physical Interactive Systems: a magnetic board, a velcro mapping board, puzzle pieces matching barriers to solutions, and a workshop game.

Collaborative Workshop Tools: a workshop format for the DI and physicians together, supported by persona sheets representing physician types.

Persona-based tools: sheets describing physician types, to deepen the conversation about who the strategy is for.

Mapping the ideas this way made the design space visible. It runs along two axes: digital to physical, and individual to collaborative. This matters because the interviews and the VHIG session pointed in two directions at once. The data-to-action gap and the time pressure argue for something digital and efficient, while the engagement findings and the stakeholder preference for physical, interactive formats argue for something collaborative and tangible.

4.3 From clusters to two concepts

Rather than developing all six clusters, I combined them into two concept directions that together cover the design space.

Concept 1, STIP-assist, came mainly from the Digital Decision Support cluster. It is a digital, AI-supported tool that helps the DI turn a conversation with a physician into identified barriers and a small set of tailored strategies, grounded in the STIP research.

Concept 2, the Implementation Co-Creation Toolkit, came from the Collaborative Workshop, Physical Interactive, and Hybrid clusters. It is a physical workshop tool, built around barrier cards, a mapping board, and persona sheets, that helps the DI and physicians identify barriers and match strategies together.

The two concepts share the same underlying logic, the six-step flow from preparation to implementation that runs through the STIP approach and the ItFits structure (Chapter 2). What differs is the form: one supports implementation through technology and efficiency, the other through social interaction and shared ownership.

Selecting a direction. I did not score every idea against the requirements in a formal matrix. The selection was driven by design judgement, and it was made under real time pressure: a working prototype was needed for the pilot and her PhD research, which meant I had to commit to a direction quickly. Of the two concepts, I saw the most potential in STIP-assist, because it concentrates on the qualities the research had shown to matter most: ease of use, efficiency, data collection, and stakeholder engagement. It also responds directly to the central finding that the hospital has data but lacks the step from data to action, and it fits the requirement to be usable under time pressure (DR1) and scalable to other hospitals as an open-access product.

The Co-Creation Toolkit scored strongly on engagement and shared ownership (DR12, DR13), but its main weakness is the one a DI already named about experiential formats: it costs time and resources, which limits how far it scales. For that reason STIP-assist became the main direction to develop, while the strengths of the toolkit, its emphasis on collaboration and tangible interaction, were kept in view for integration into the concept later (Chapter 5).

The two concepts themselves are described in full in the next chapter.

5 Two concepts: comparison and development

The previous chapter ended with two concept directions. This chapter describes them in full, compares their strengths and weaknesses, and explains why I carried one of them forward into development.

5.1 Concept 1: STIP-assist

STIP-assist supports the DI in turning practical, context-specific insights into concrete, tailored implementation strategies, using an AI-supported tool. The aim is to turn a rich conversation on the ward into structured, evidence-based recommendations that stay aligned with implementation science.

How it works. The process starts with a conversation between the DI and a physician on the department. A short set of structured questions explores the barriers around infection prevention, such as workload, responsibilities, motivation, and practical constraints. This moment captures the context-specific insights that the rest of the process builds on.

The tool then follows six steps (Figure 8):

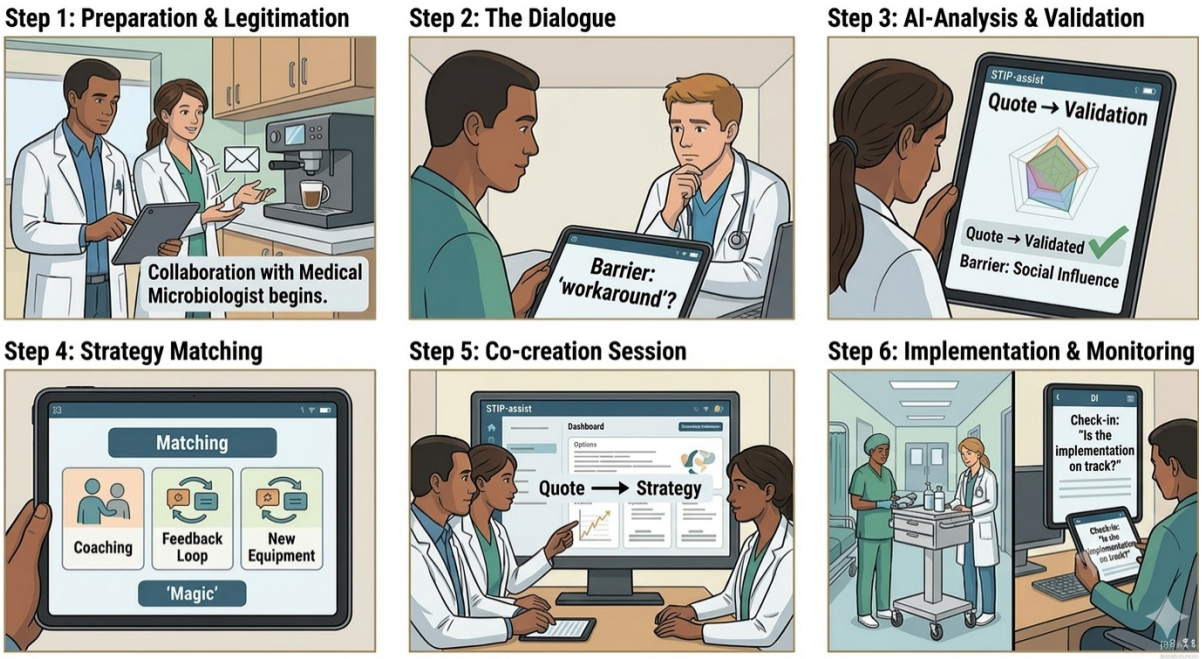


Figure 8. *Storyboard of the STIP-assist concept, illustrating the intended six-step flow. (generated with Gemini)*

Preparation and legitimation. The DI starts with a short "go" from the clinical microbiologist, which gives the work a mandate before it begins.

Dialogue on the ward. The DI has a short conversation of ten to fifteen minutes with a physician. With consent, the tool listens in, so the DI can focus on the conversation instead of taking notes.

AI analysis and validation. The tool turns the audio into text and identifies the likely barriers, linking them to the underlying determinants (capability, opportunity, motivation). The DI can correct the analysis, so a human stays in the loop.

Strategy matching. The barriers are matched to strategies from the STIP research, shown as a barrier profile and a small set of two to three tailored options.

Co-creation session. The DI, the physician, and the microbiologist look at the result together and decide which strategy is most suitable and feasible in their context.

Implementation and monitoring. The chosen strategy goes into practice, and the tool supports a short follow-up to see whether it worked.

Added value. The tool supports both the analysis of complex situations and the communication between stakeholders. It structures the conversation, makes insights explicit, and offers concrete suggestions to guide the decision. Every suggestion stays traceable to the STIP research and the DCE data, so the advice keeps its scientific grounding. By combining AI-supported analysis with human deliberation, it aims to bridge the gap between abstract theory and daily practice.

Considerations. The concept also has limits. It depends on technology and on the privacy and IT conditions that come with handling conversational data. There is also a risk that the work moves "behind a screen," which could weaken the personal, face-to-face engagement that the physical concept creates more naturally.

5.2 Concept 2: Implementation Co-Creation Toolkit

The Implementation Co-Creation Toolkit is a physical, workshop-based tool. Where STIP-assist works through technology, this concept works through interaction, reflection, and shared ownership. Its core idea is that implementation is a social process: instead of analysing barriers alone, the DI and the care team explore them together and decide on actions as a group. Making the process physical and visual supports discussion and mutual understanding.

How it works. The toolkit is used in a short workshop of thirty to forty-five minutes, in which the DI and the care team work towards concrete actions through six steps (Figure 9):

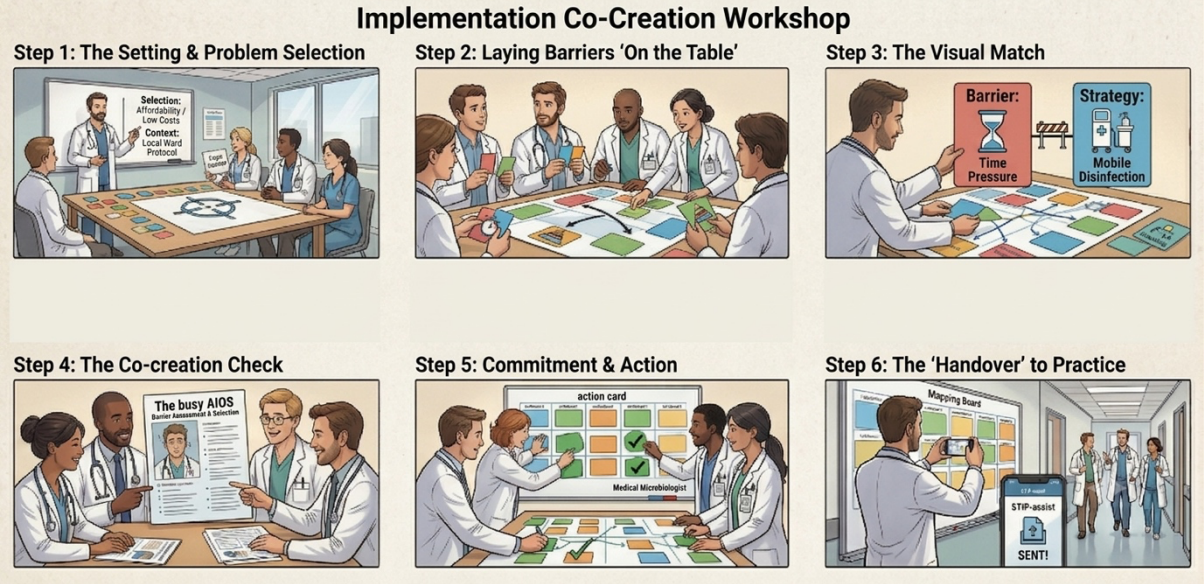


Figure 9. *Storyboard of the Implementation Co-Creation concept.(generated with Gemini)*

Setting and problem choice. The group picks one concrete problem, for example low adherence to a particular guideline.

Putting barriers on the table. Participants choose physical barrier cards based on COM-B (time pressure, social norms, lack of resources). Because they pick the cards themselves, they recognise the barriers and feel involved.

The visual match. Strategy cards from the STIP research are laid next to the barriers on a mapping board, making the link from problem to solution visible and tangible.

The co-creation check. Persona sheets representing physician types help the group discuss a barrier through someone else's eyes, which makes it easier to be honest about it.

Commitment and action. The group chooses one concrete action and assigns responsibility, marked with a commitment sticker.

The handover to practice. The DI photographs the completed board as a record, and the action moves to the ward.

Key elements. The toolkit consists of barrier cards (based on capability, opportunity, and motivation), strategy cards (evidence-based strategies from the STIP research), a mapping board (a visual structure connecting barriers and strategies), and persona sheets (representations of different healthcare professionals to support discussion).

Added value. The toolkit creates value in three ways. It increases ownership, because stakeholders contribute to the decisions themselves. It reduces hierarchy, because the physical, visual process invites more equal participation. And it leads to concrete action, because abstract challenges are translated into tangible steps. This responds directly to insights from the research phase: the

importance of stakeholder engagement, the influence of social dynamics, and the difficulty of reaching physicians, who are given an active role here.

Considerations. The concept also has clear limits. It takes time and coordination to organise a workshop, its effect depends on active participation, it is less efficient than a digital tool, and it is harder to scale across many departments.

5.3 Comparing the two

Both concepts rest on the same six-step logic and the same STIP research. What differs is the form. STIP-assist supports implementation through technology and efficiency, while the Co-Creation Toolkit supports it through social interaction and shared ownership. Together they represent two complementary approaches: technology-driven support on one side, social and interactive support on the other. The table (Table 4) compares them against the qualities that matter most for this project.

Criterion	STIP-assist (digital)	Co-Creation Toolkit (physical)
Efficiency / time pressure (DR1)	Strong: fast analysis, little time needed	Weaker: a workshop takes 30–45 min from several people
Ease of use (DR10)	Strong: if kept intuitive	Strong: physical and low-threshold
Buy-in / co-creation (DR12, DR13)	Weaker: risk of working behind a screen	Strong: joint, tangible ownership
Scalability (open-access goal)	Strong: digital, deployable across hospitals	Weaker: costs time and materials per session
Data collection	Strong: transcript and analysis, traceable	Weaker: data stays on paper and the board
Feasibility within the deadline	Strong: buildable as a prototype for the pilot	Weaker: harder to deliver quickly
Technical and privacy dependence	A weak point: privacy and IT conditions apply	Strong: needs no technology

Table 4. *Comparison between Concept 1 and Concept 2.*

The pattern is clear. STIP-assist is stronger on efficiency, scalability, data, and feasibility, the qualities the research and the requirements pointed to as decisive. The Co-Creation Toolkit is stronger on engagement, buy-in, and independence from technology, which are real strengths but harder to scale and slower to deliver.

This is why I carried STIP-assist forward. It concentrates on ease of use, efficiency, data collection, and stakeholder engagement, and it answers the central finding that the hospital has data but lacks the step from data to action. The toolkit's strengths stay worth keeping, which is why some of its elements return inside STIP-assist, as described below.

5.4 Developing concept 1

From here I developed STIP-assist further. I showed it to a range of stakeholders, including the sounding board group, a consortium meeting, the VHIG congress, and the project team at Amsterdam UMC, and used their feedback to iterate on the design. That feedback, and what it changed, is discussed in Chapter 6.

A few elements of the Co-Creation Toolkit have already found their way into STIP-assist. The DI can download the project's physical worksheets from within the tool, and several pages allow the transcript or other information to be exported as a PDF. The tool also builds in moments of collaboration: after certain steps, it prompts the DI and physician to meet before continuing, which brings the co-creation idea of concept 2 into the digital flow. A fuller integration of the two concepts is a future step, described in the roadmap (Chapter 8).

6 Prototype, feedback and feasibility

This chapter follows STIP-assist from a first prototype to a working tool. I start with the prototype I built to show the idea, then report the feedback I gathered from stakeholders and the feasibility work that checked whether the tool could actually be built. I then describe the design iteration that followed from both, and end with how I built that iteration for real use inside the hospital's research environment.

6.1 The first prototype

To explore the flow and usability of STIP-assist, I built a first prototype. It was a mid-fidelity, clickable mock-up: it showed the structure, the screens, and the key interactions, but the analysis was simulated with example data rather than driven by real AI. I built it before the hospital's AI environment was available to me, so its purpose was proof of concept, to make the idea tangible enough to show to stakeholders and to gather feedback. The prototype walks through six screens.

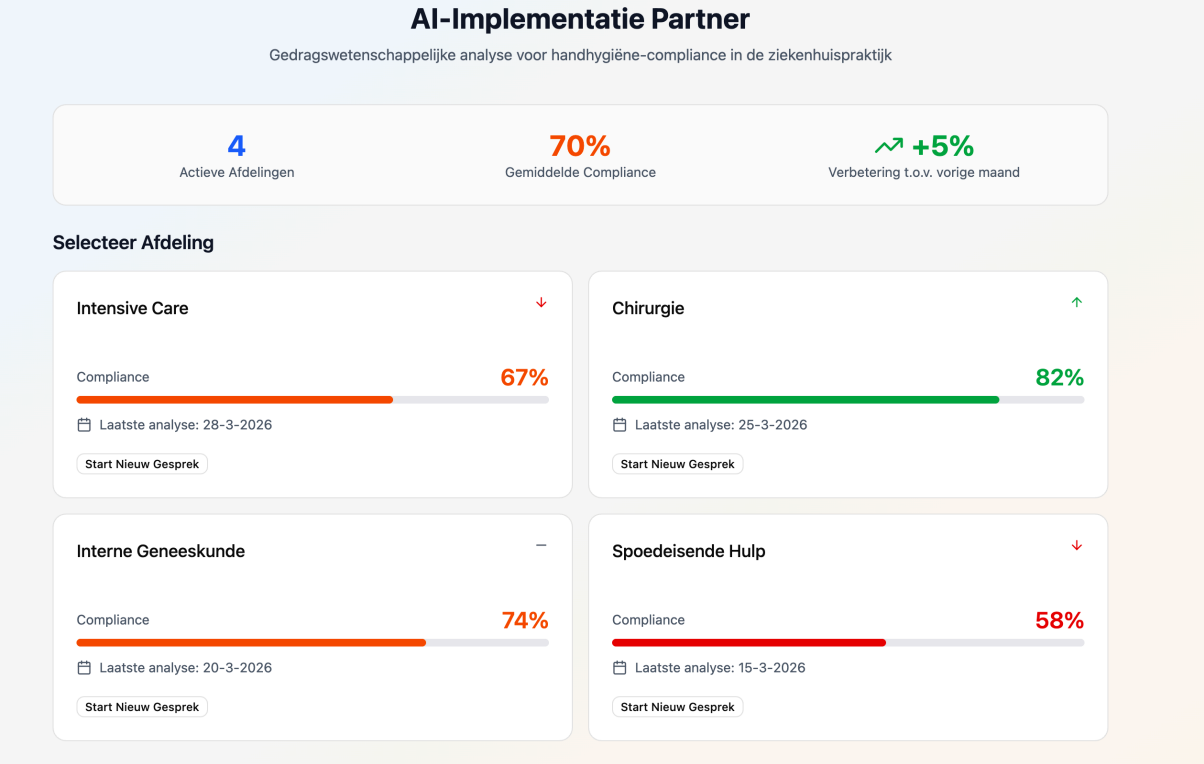


Figure 10. *Dashboard*

Dashboard. (Figure 10) The tool opens on a dashboard showing the hospital's departments and their hand hygiene compliance. Each department appears as a card with its current compliance percentage, a trend indicator, and the date of the last analysis. From here the DI picks a department and starts a new conversation. The dashboard gives the DI an overview of where attention is needed and acts as the entry point into the workflow.

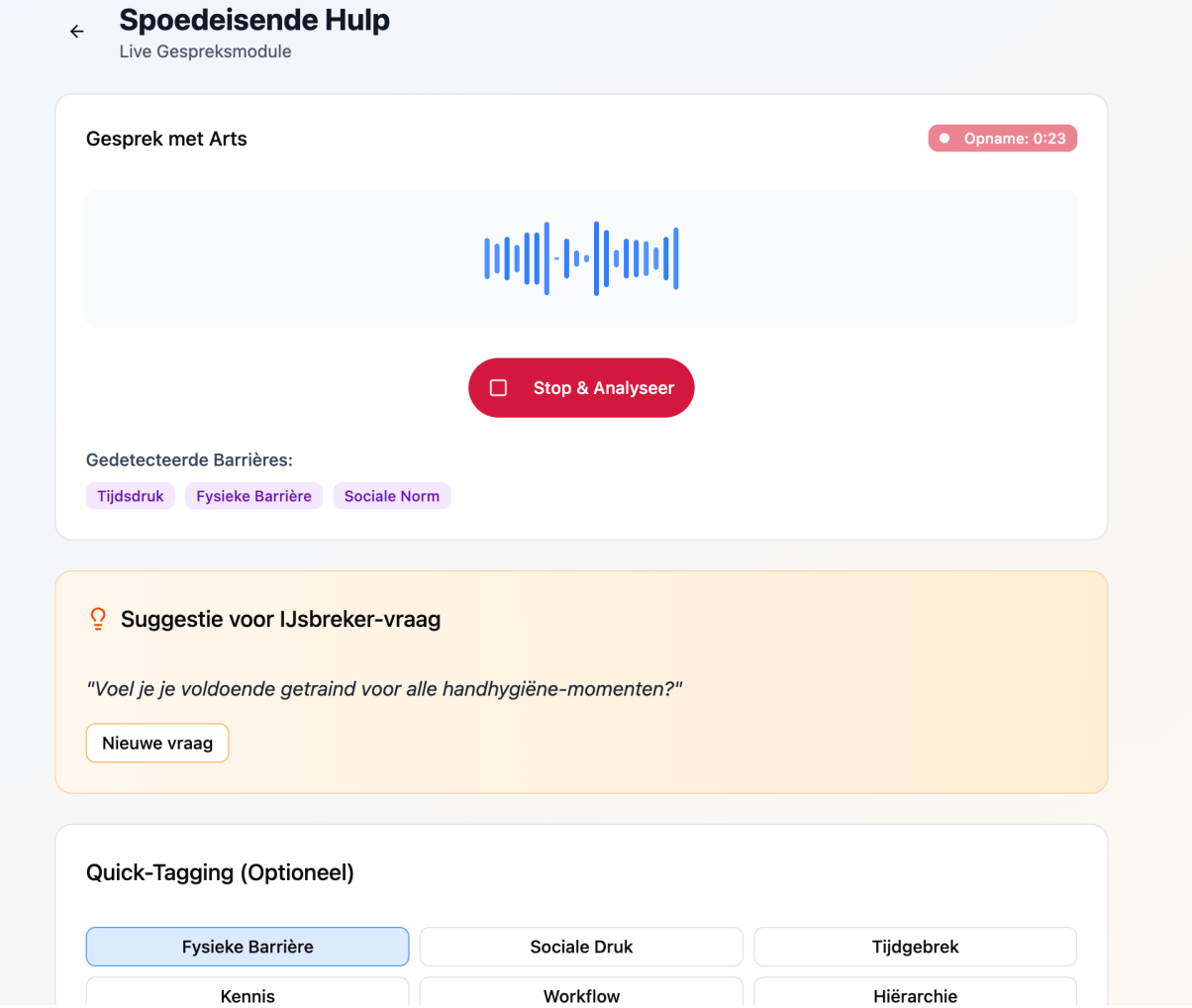


Figure 11. *Conversation module*

Conversation module. (Figure 11) After choosing a department, the DI enters the conversation screen, used during a live conversation with a physician. The DI can record the conversation, which removes the need for note-taking, and the screen offers suggested icebreaker questions based on the STIP methodology. In the mock-up, example barriers appear on screen as if detected from the conversation, and the DI can add or adjust them with a quick-tagging function. This screen combines capturing the conversation, guiding it, and a first look at the barriers in one place.



Figure 12. *Transcript screen*

Transcript. (Figure 12) Once the conversation is done, the DI sees a transcript. This lets the DI review what was said and keeps the later analysis grounded in the real conversation. Showing the full transcript was a deliberate choice to support transparency and trust: the DI can trace any insight back to a specific statement.

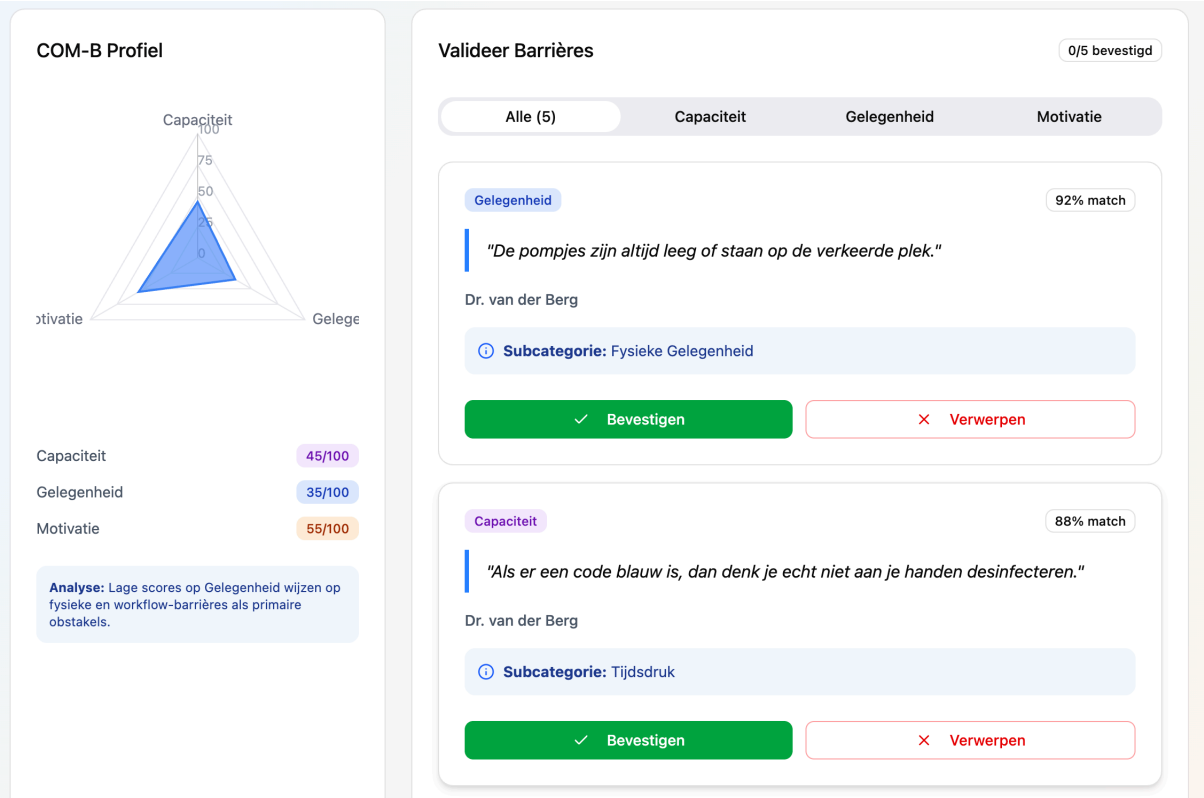


Figure 13. *Barrier analysis with COM-B profile*

Barrier analysis. (Figure 13) The tool then presents the barriers in a structured way, categorised by the COM-B determinants (capability, opportunity, motivation). On the left, a visual profile shows how the barriers are spread across these determinants. On the right, each barrier appears with a supporting quote, a suggested category, and a match percentage. The DI can confirm or reject each one. This validation step keeps the professional in control and makes the analysis a shared result rather than an automated verdict. The match percentages here were example values, since the analysis was simulated.

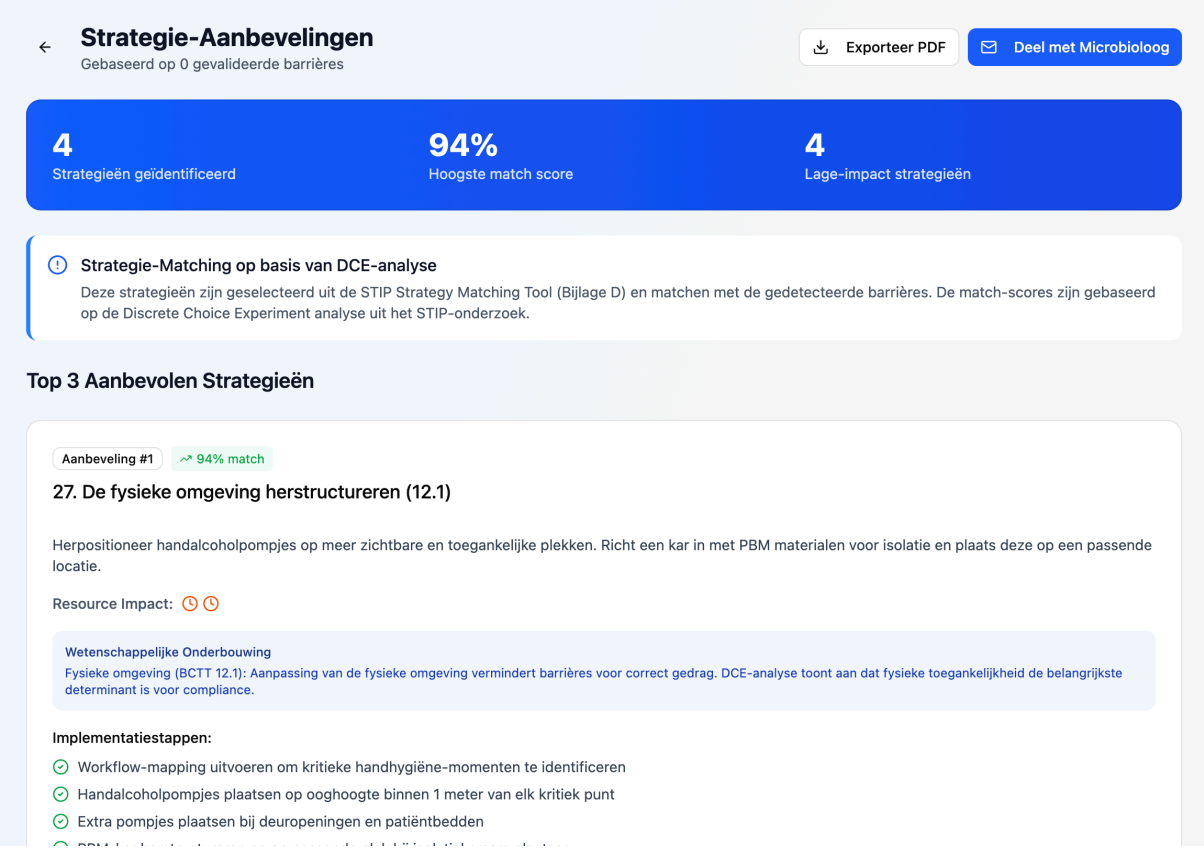


Figure 14. *Strategy recommendations*

Strategy recommendation. (Figure 14) Based on the confirmed barriers, the tool suggests a small set of tailored strategies, drawn from the STIP Strategy Matching Tool and the DCE. The screen highlights the top two to three options, each with a match score, a short description, an indication of the effort required, a link to the underlying behavioural theory, and concrete implementation steps. The results can be exported or shared with other stakeholders, such as the microbiologist.

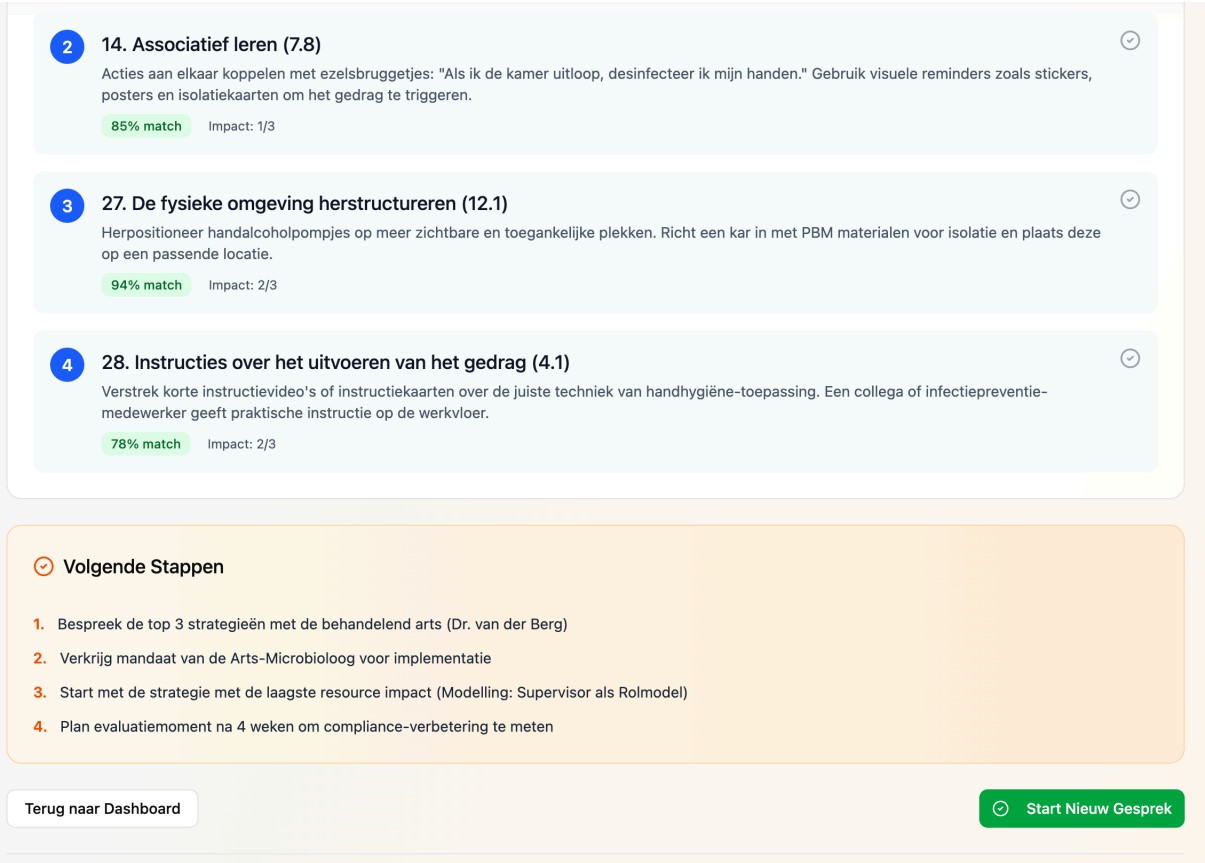


Figure 15. *From strategy to implementation*

From strategy to implementation. (Figure 15) The final screen turns the chosen strategy into next steps: discussing it with the physician, getting a mandate from the microbiologist, choosing a feasible starting point, and planning a moment to evaluate. From here the DI can return to the dashboard or start a new conversation, which supports an iterative way of working across departments.

6.2 Feedback from stakeholders

I showed the prototype to a range of stakeholders to test whether the concept made sense and where it needed to change. These included the sounding board group, a consortium meeting, the VHIG congress, and my project supervisors at Amsterdam UMC. The feedback was broadly positive about the direction, and it pointed to a clear set of improvements. I group it here by theme.

Simplify the flow. The most direct feedback, from the sounding board group, was that the first three screens could be merged into one, and that the conversation should use five focused questions rather than ten. This matches the usability requirements (DR10, DR11): the tool has to stay light and quick to use.

Add the physicians' voice through a simplified ranking. The strongest and most repeated point came up in the consortium meeting and from two supervisors independently. Alongside the DI's conversation, physicians should give their own input through a simple ranking, indicating which

strategy attributes they find important and which version they prefer. The aim is to capture the spread of preferences on a ward and give direction, rather than statistical significance. This is a lightweight, accessible version of the discrete choice experiment, and its repeated mention across three sources made it a priority for the next iteration.

Connect strategies to several sources. Related to this, the consortium and one supervisor asked for the strategy recommendations to draw on more than one source. A match should reflect the direct fit with the identified barriers, the local physician preferences from the ranking, and the findings from the DCE research ("research shows that..."). Each strategy should show how it links to what the physicians indicated and to the evidence, so the advice is grounded in both the local context and the research.

Offer information in layers. Two supervisors asked for a layered design: a simple base version that guides the DI through the core process, with an optional deeper layer for those who want more nuance and more options to consider. In practice this means short pop-ups with research-based guidance at the relevant points, for example next to a barrier or a strategy. This responds directly to the requirement to serve DI professionals of different backgrounds (DR11).

Integrate the surrounding process. A recurring point was that the web tool currently covers only the matching itself, the path from context analysis to barriers to strategy advice. The surrounding process, captured in the project's worksheets (preparation, goal setting, action planning, the do-and-evaluate phases), still sits outside the tool. One supervisor suggested a tabbed interface organised by phase, with icons that surface the worksheet guidance and pop-ups that prompt the DI to act when needed. This raised a design question I carried into the iteration: how much of the worksheet process belongs inside the tool, and what stays as a physical complement.

Bring co-creation back in. The consortium connected this concept to the second concept from Chapter 5. They suggested building the co-creation idea into STIP-assist as a step, for example a prompt to check the result with the physicians, which the DI could do through a co-creation session. This supports the moments of collaboration already planned in the tool and keeps the strengths of the physical concept in view.

Reframe barriers as neutral factors. A smaller but useful point: the tool speaks of "barriers", while a factor can be either a barrier or a help. Framing these as neutral "factors" fits the findings on framing from the research, where physicians respond better to "we are doing well, what can be better" than to a deficit message.

Privacy and licensing. The consortium raised an important constraint: how do licensing and privacy work for an AI tool inside a hospital, where the rules are strict. They suggested meeting the data protection officer to discuss it. This question fed directly into the feasibility work described in the next section.

The VHIG questionnaire. At the VHIG congress I used an interactive Mentimeter to gather quantitative input from the audience of infection prevention professionals. The results, summarised below (Table 5), supported several of the design choices.

Question (VHIG, n ≈ 65–72)	Most chosen answer	Share
How do you prefer to work on improvement projects?	A mix of alone and with physicians	74%
Which format supports you best?	A combination of digital and physical	82%
What matters more in improvement projects?	Middle, no clear preference	avg. 5.1/10
How do you prefer to gather physicians' input?	A combination of questionnaire and conversation	49%

Table 5. *Results of Mentimeter VHIG congres*

Two results stand out. Almost nobody wanted to work fully alone, which supports building collaboration into the tool (DR9, DR13). And 82% preferred a combination of digital and physical, which is the clearest argument for keeping physical elements such as the worksheets inside the digital tool rather than choosing one form over the other. The even split on efficiency versus thoroughness suggests the tool should be able to serve both, which fits the layered design above.

6.3 Feasibility

The prototype showed the intended flow, but its core promise, an AI that turns a conversation into structured barriers and matched strategies, was still simulated. Before building it for real, I needed to know whether that was technically possible, whether it was allowed under the hospital's privacy rules, and how it could be deployed across hospitals. I explored these three questions with a PhD researcher working on AI in healthcare, the hospital's data protection officer, and an AI architect at Amsterdam UMC.

Technical feasibility. Extracting structured insights from text is an established use of machine learning in implementation science: it has been used to extract information from trial reports and predict intervention outcomes (West et al., 2023; Hastings et al., 2025), and a recent framework sets out how AI can be used responsibly in implementation practice (Fontaine et al., 2026). The PhD researcher confirmed that the core function, turning a transcript into candidate barriers, is feasible in principle. The qualification was that feasibility depends heavily on the input: analysing raw, free-form conversation is far more complex than working with structured or semi-structured input.

Privacy. The data protection officer made clear that feasibility depends on the design of the data handling rather than on a simple yes or no. The key questions are where the data comes from, what it contains, and how traceable it is to individuals. Conversational data, and especially audio recordings, raise the sensitivity sharply. The conclusion was that the concept is feasible within Amsterdam UMC, provided privacy is treated as a design parameter from the start, and that working with structured or summarised input rather than raw audio reduces the risk considerably.

Architecture and scalability. The AI architect confirmed that the tool could be developed within the hospital's internal environments, which allow controlled experimentation while keeping sensitive data inside institutional boundaries. Because the tool is meant to be tested across several hospitals, he pointed to secure interfaces such as APIs that let other hospitals use the system without direct access to internal infrastructure, while noting that this adds its own questions of security and data governance.

What this meant for the design. Together, these conversations pointed in the same direction. The technology is feasible, but raw audio is both the hardest input to analyse and the riskiest for privacy. This is why the design moved towards more structured input, such as a transcript or an interview guide built around fixed topics, rather than fully automated analysis of free conversation. Feasibility was therefore not a check at the end, but a factor that actively shaped the concept.

6.4 The last iteration

The feedback and the feasibility work led to a substantial revision of STIP-assist. Most of the changes followed directly from what stakeholders asked for, and on two points I made a deliberate choice to do something different. The biggest shift, though, runs underneath all of them: where the first prototype simulated its analysis with example data, this version runs on real AI through the hospital's AI environment, so the tool now genuinely turns input into factors and matched strategies.

A simpler, phased interface. Following the sounding board group, I merged the first three screens into one, which makes the entry into the tool quicker and lighter (DR10). I also restructured the tool around the phases of the implementation process: preparation, conversation, analysis, strategies, operationalisation, and the do-phase. (Figure 16) The worksheets that used to sit outside the tool are now built in as screens within these phases, each broken into clear steps that unlock as the previous one is completed. This answers the feedback that the tool previously covered only the matching step, and it brings the surrounding process into one place.

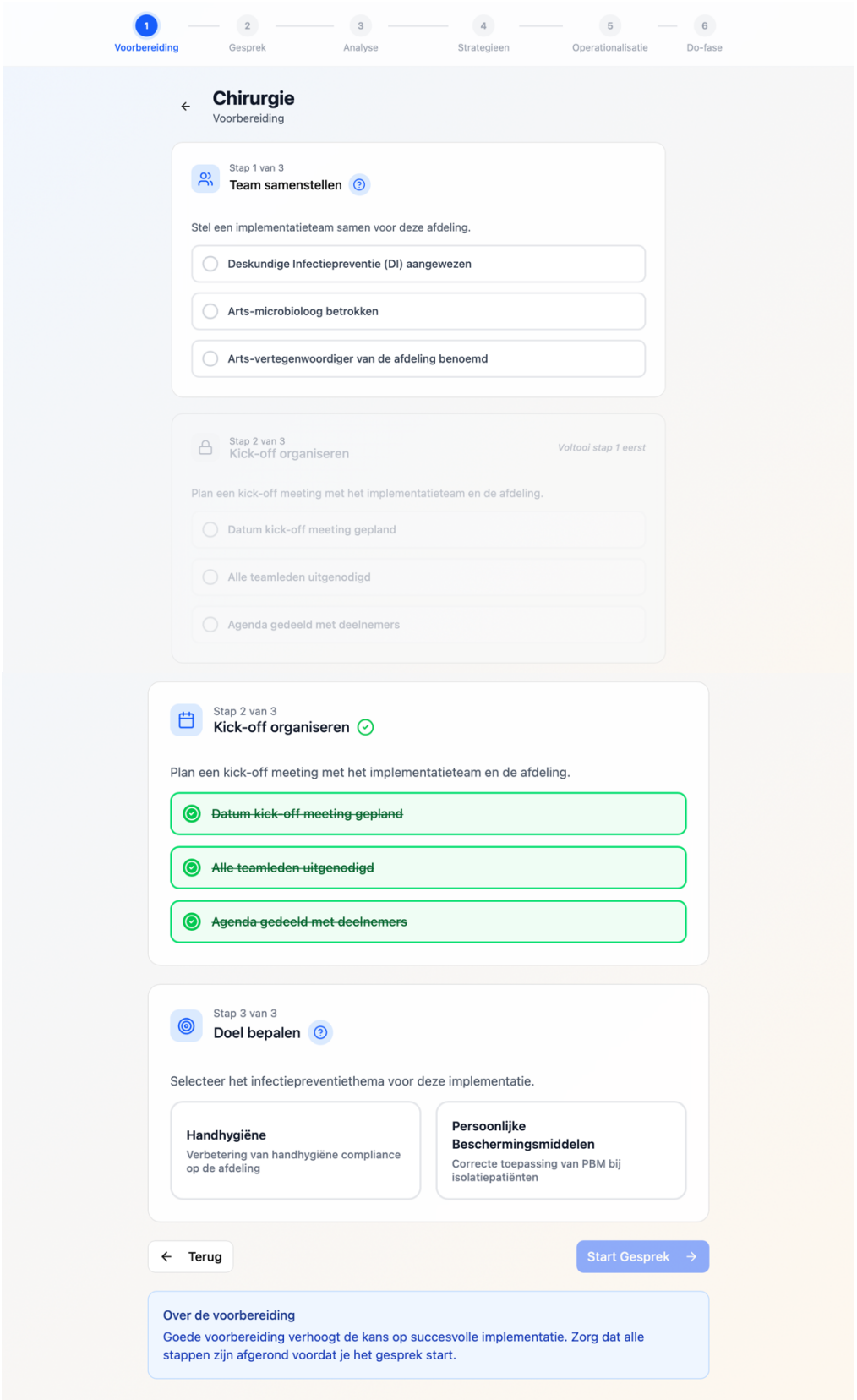


Figure 16. *The six-phase structure, shown here in the Preparation phase: composing the implementation team and organising the kick-off. The worksheets are now integrated as guided steps.*

Physicians' input through a simplified ranking. I built in the ranking that the consortium and two supervisors asked for. Physicians indicate which strategy attributes they find important and which version they prefer, which gives a picture of the spread of preferences on a ward. (Figure 17) This is a lightweight, accessible form of a discrete choice experiment, a method designed to elicit preferences by having people choose between alternatives with varying attributes (Lancsar & Louviere, 2008). The aim here is direction for a specific department rather than statistical significance.

On the number of questions, I went the other way. The sounding board group suggested reducing the set to five questions. I chose to keep it larger, because a larger set gives a more reliable result. The cost is small and manageable: in a conversation the DI decides how many questions to ask, so the time pressure there stays in the DI's hands, and as a questionnaire it adds only about two minutes for the physician. I judged that small addition to be worth the better result.

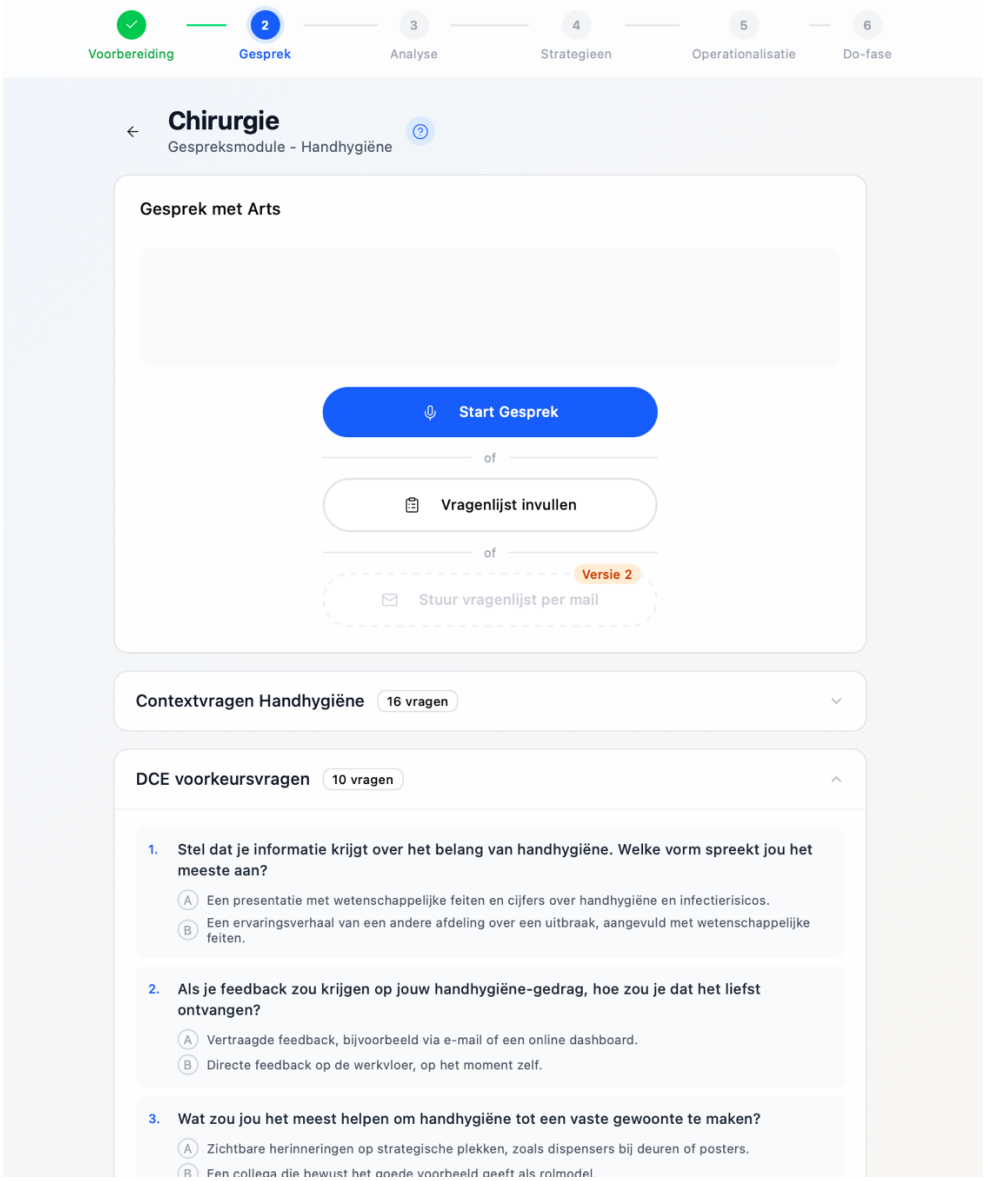


Figure 17. *The simplified ranking that captures physicians' preferences for strategy attributes.*

Factors instead of barriers. Following the feedback, the tool now speaks of "factors" rather than "barriers", since a factor can either block or help a behaviour. (Figure 18) This neutral framing fits the research on how physicians respond better to a constructive message than to a deficit one.

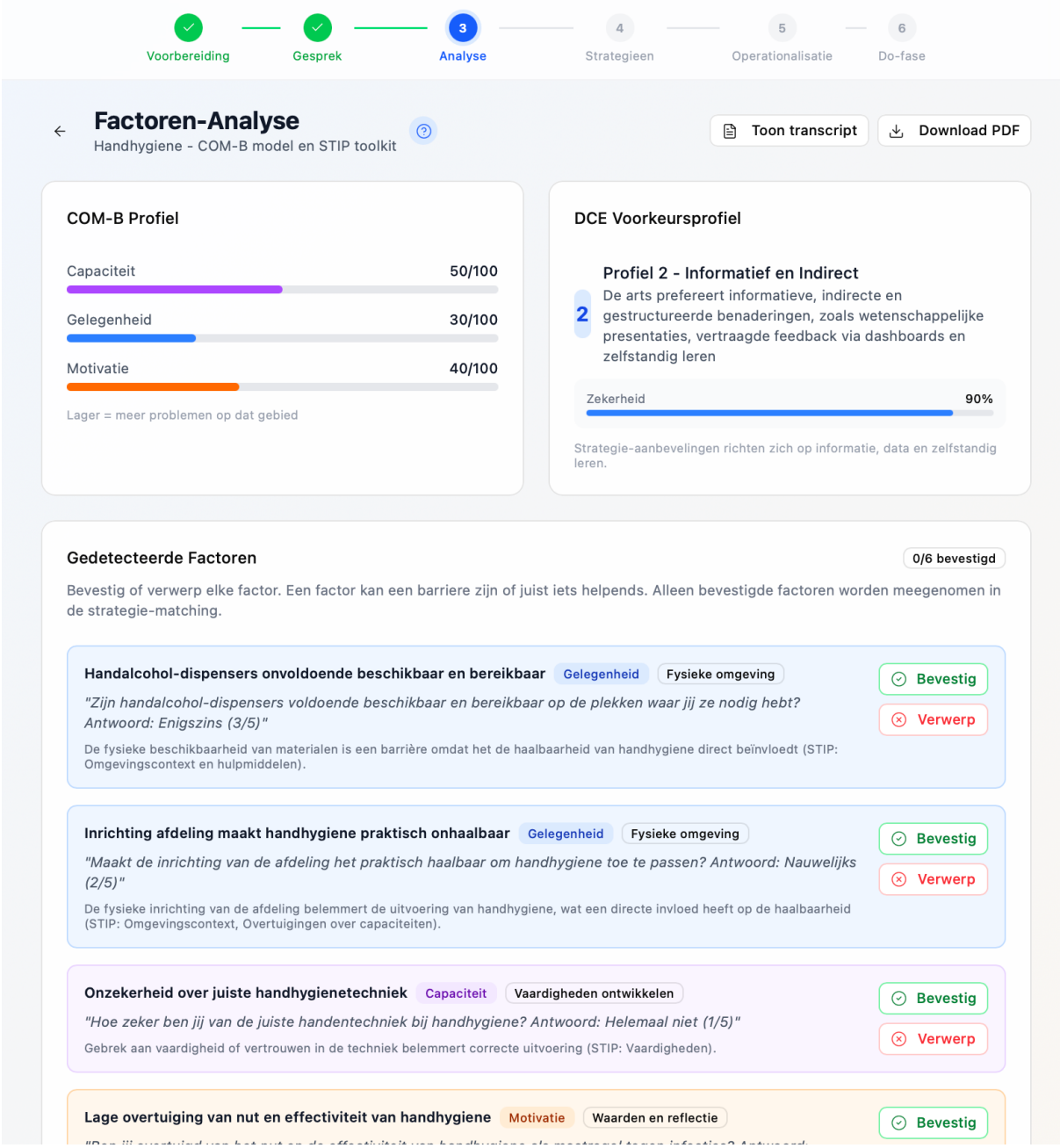


Figure 18. *The analysis screen, where the AI presents the identified factors, categorised by COM-B, for the DI to validate.*

Strategies linked to several sources. The strategy recommendations now draw on more than one source, as the consortium and one supervisor asked. A match reflects the direct fit with the identified factors, the local physician preferences from the ranking, and the findings from the STIP research. (Figure 19) This grounds the advice in both the local context and the evidence.

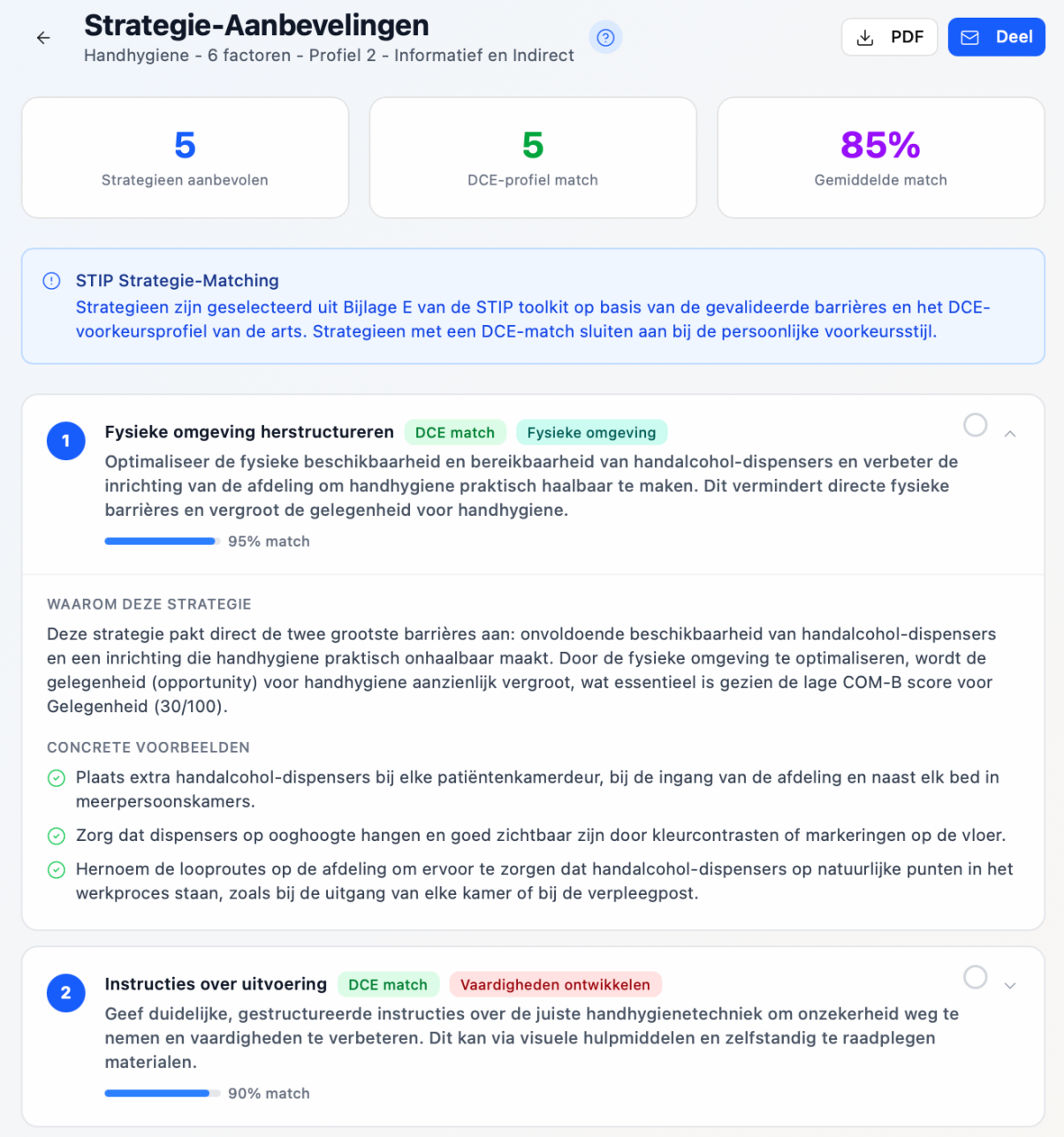


Figure 19. *The strategy screen, where each recommendation combines the fit with the factors, the local physician preferences, and the research findings.*

Built-in moments of collaboration. The tool now prompts the DI at certain points to check the result with the physicians or the team before continuing. (Figure 20) This brings the co-creation idea of the second concept into the digital flow. Alongside this, the DI can download the project's worksheets from within the tool and export the transcript or other information as a PDF. This keeps a physical and shareable element in a digital tool, in line with the strong VHIG preference for combining the two.

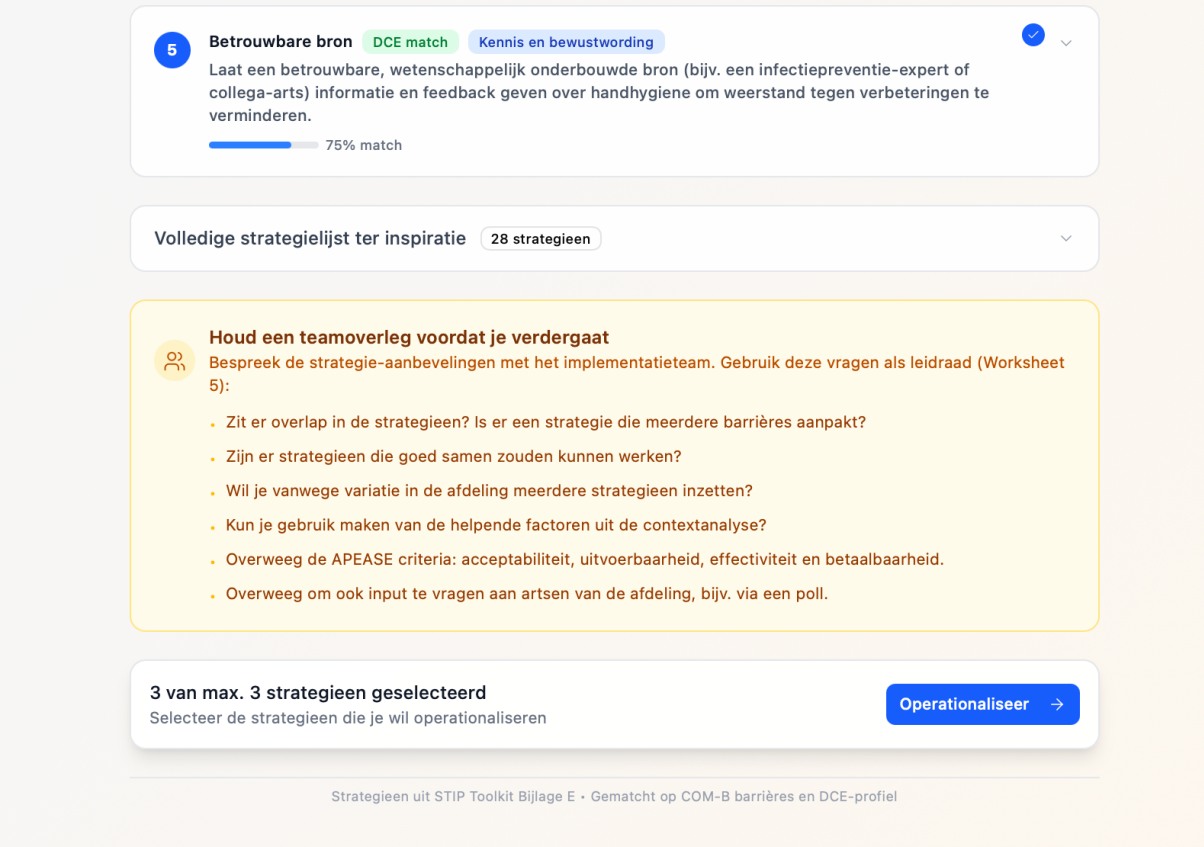


Figure 20. A collaboration prompt, asking the DI to discuss the result with the team before continuing.

Structured input. Following the feasibility work, the tool can take input as a structured questionnaire as well as a conversation, which is both easier to analyse and safer for privacy than free-form audio (Figure 17).

What I chose not to do, for now. Two points from the feedback are not in this version. I kept the questionnaire larger rather than shorter, as described above. And I did not build the layered information as pop-ups: I am not convinced that pop-ups are the right form for it, and the deeper layer of research-based guidance still needs to be designed and implemented. This is a planned step, described in the roadmap (Chapter 8).

6.5 Building it for real use inside myDRE

The iteration above describes the design. This section describes how I turned that design into a tool that actually runs. Where the first prototype simulated its analysis, the final version is a working application that the DI can open, use end to end, and that processes real input. Building it raised a set of practical questions that the design alone did not answer: where the tool should run, how it should handle several physicians on one ward, and how it should keep a DI's work safe over the course of a pilot. Each of these shaped the tool in a concrete way.

A tool that runs inside the hospital. The feasibility work pointed clearly to one requirement: sensitive conversation data must stay inside the hospital. I therefore built the final version to run entirely within myDRE, the Trusted Research Environment of Amsterdam UMC. The tool lives on a protected virtual machine, and all of its processing happens there. The transcription of a recorded conversation and the analysis that turns input into factors and matched strategies both run through the hospital's own internal AI service, so no conversation data leaves the institutional boundary. This is the difference between the prototype and the final tool. The promise that was simulated before is now real, and privacy is built in from the start. Inside this environment the DI captures input in the way that fits the situation, by recording the conversation, uploading an audio file, or filling in the questionnaire. (Figure 21)

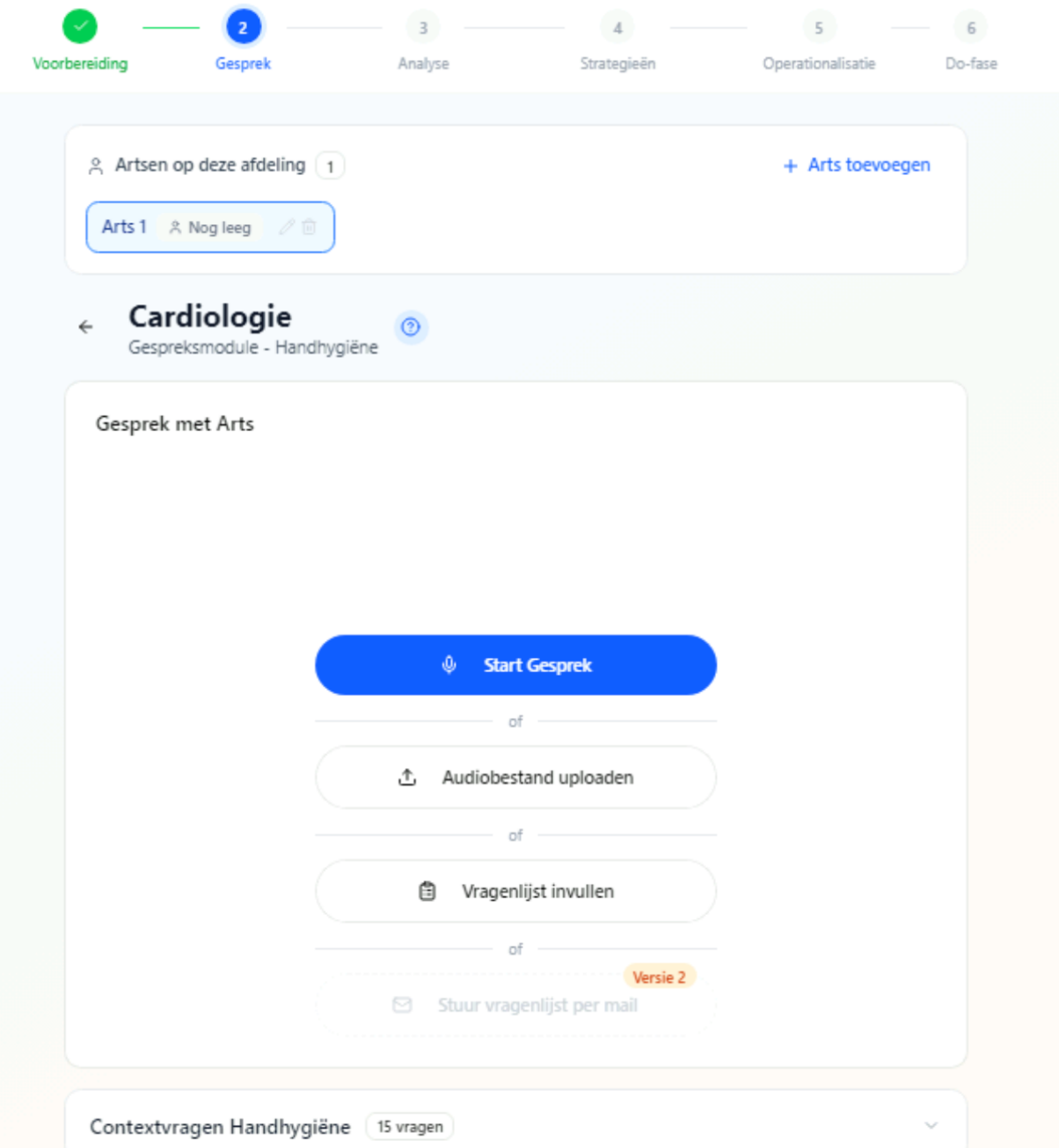


Figure 21. The conversation module inside myDRE. The DI can add several physicians to one department, and capture each conversation by recording live, uploading an audio file, or filling in the questionnaire.

Several physicians on one ward. A ward is not one physician, and improving compliance there means understanding a group. I therefore extended the tool so a DI can add several physicians to one department trajectory. Each physician has their own conversation or questionnaire, their own preference ranking, and their own COM-B profile, so an individual picture is never lost. The factors from all of these are then pooled into one departmental list that the DI validates once, which keeps the analysis manageable and reflects that a barrier on a ward is rarely the property of a single person. (Figure 22) The COM-B picture is shown both per physician and as a departmental view, so the DI can move between the individual and the group. I made one limitation explicit here, in

discussion with Miriam: the departmental view currently averages the scores, and an average can hide disagreement, since a ward where half the physicians score high and half score low looks the same as a ward where everyone sits in the middle. Showing the spread, and flagging strong disagreement, is a known refinement for a later version. For the pilot, the averaged view is enough to give the DI direction.

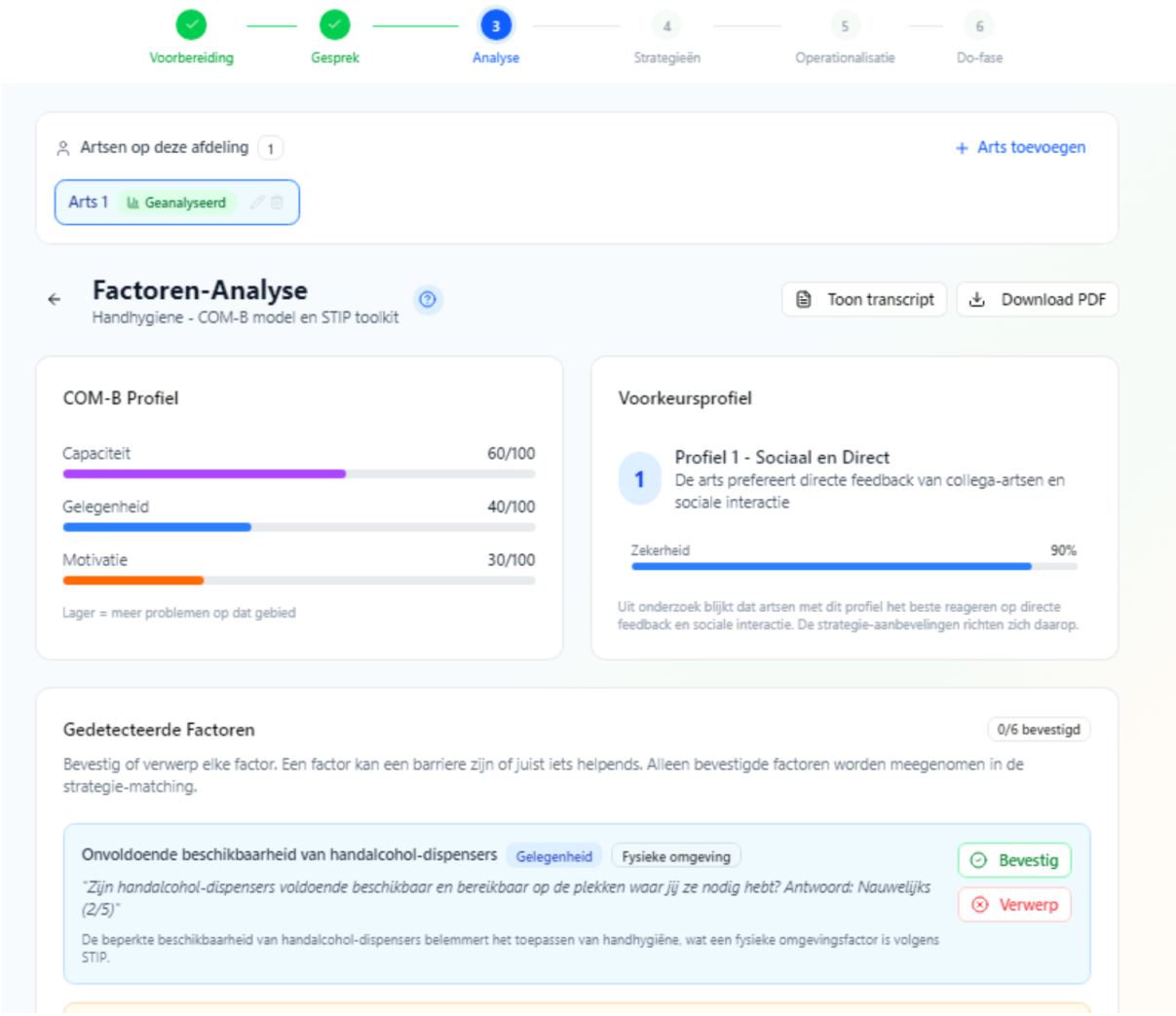


Figure 22. The analysis screen running on real AI. Per physician it shows a COM-B profile and a preference profile, and below it the detected factors that the DI confirms or rejects. The bar at the top holds the physicians on the department, each with their own analysis status.

Keeping the DI's work safe over a pilot. A pilot runs for months, and a DI builds up a trajectory across many sessions. The tool therefore has to remember what was done. The environment made this harder than it sounds, because the virtual machine resets every night, so anything held only in the browser would be lost. I solved this by writing all progress to a file on the machine, which the tool reads back when the DI opens it again. The DI can close the tool, return the next day, and pick up where they left off. The home screen surfaces this directly, with a card that lets the DI continue an active trajectory in one click, and it shows the progress across the departments on the ward. (Figure

23) Each DI works in their own account within myDRE, so one DI's work is separated from another's, and a department's data stays with the person handling it.

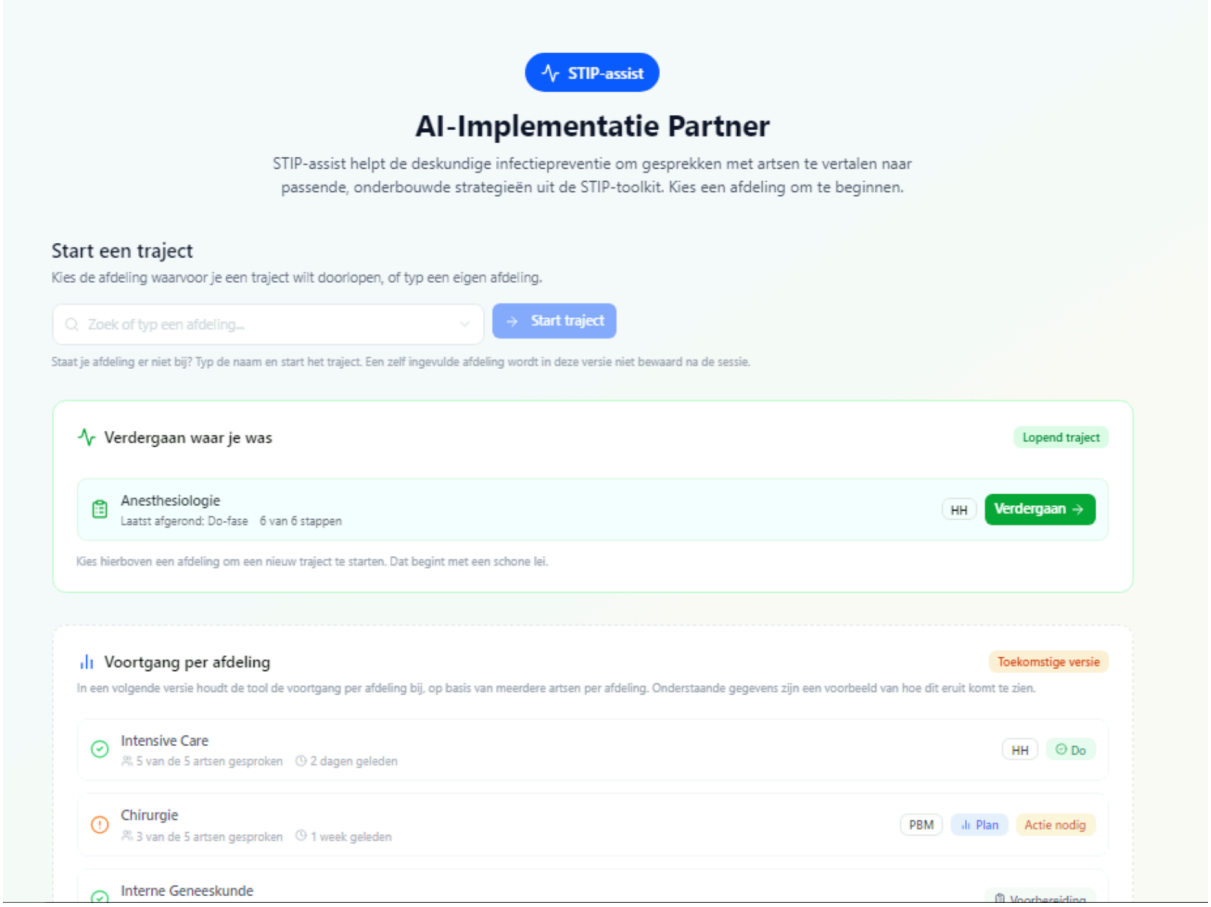


Figure 23. The home screen. The card at the top lets the DI continue an active trajectory in one click, and the overview below tracks the progress per department across the ward.

Working around the limits of the environment. Building inside a protected environment also meant working around its limits. The virtual machine has no microphone, so live recording is not always possible there. This is why I added the option to upload an existing audio file alongside the live recording, so a DI can record a conversation on a separate device and bring it into the tool for transcription (Figure 21). It keeps the conversation route open within the constraints of the environment, next to the questionnaire route built for the same reason of easier and safer input. The same thinking shapes the end of the workflow. In the do-phase the strategies become a concrete action plan, and the DI can export that plan to Excel to share it and work with it outside the tool. (Figure 24) This keeps a shareable, physical element in a digital tool, in line with the strong VHIG preference for combining the two.

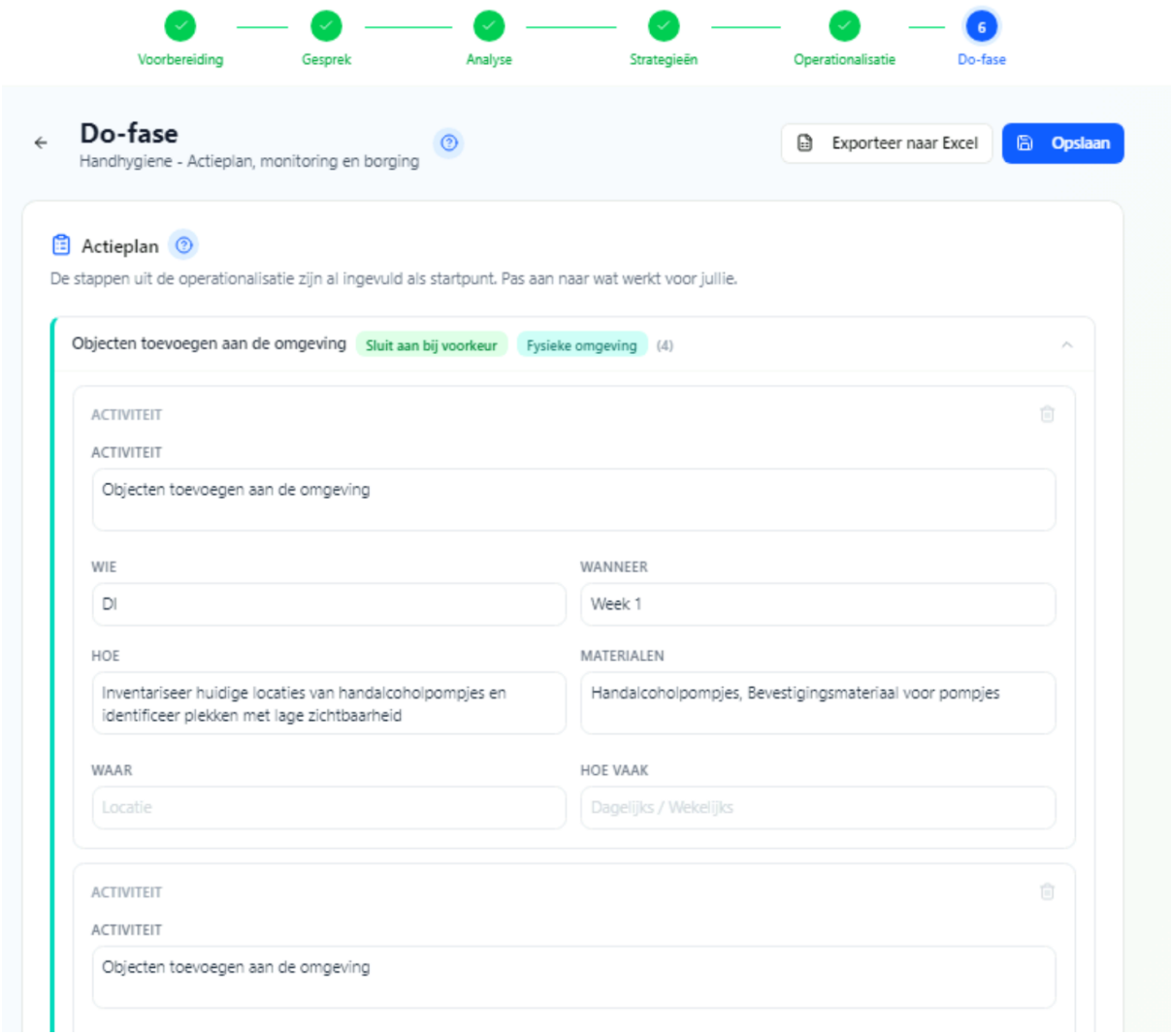


Figure 24. The do-phase. The chosen strategies are turned into an editable action plan, which the DI can export to Excel to share and use outside the tool.

Together these steps turned the design into a tool that is ready to be used in a real pilot. It runs inside the hospital, it handles a ward rather than a single physician, and it keeps a DI's work safe across the months a pilot lasts. The decisions here were practical, but each one traces back to the same principles that shaped the design: keep the DI in control, keep the data safe, and keep the tool light enough to use.

This version has not yet been tested with users. What it achieves against the design requirements, and what still needs to be evaluated, is the subject of the next chapter.

7 Final evaluation

This chapter evaluates STIP-assist against the design requirements set out in Chapter 3, which serve as the criteria for judging the concepts. The evaluation here is analytical. It assesses how well the current tool meets each requirement, and it is honest about what has been realised, what is designed but still needs testing, and what is still to be built. A full usability test was designed for this purpose, but its execution falls outside the project window. The protocol for that test is presented at the end of this chapter.

7.1 Evaluation approach

I evaluate the tool on two levels. First, against the thirteen design requirements (DR1–DR13), which give a detailed, traceable check. Second, against the two overarching criteria: practicality (does it fit and work in clinical practice) and relevance (does it address the real needs identified in the research).

The main limitation of this evaluation is that the tool has not yet been tested with users in practice. A formal usability test was designed for this, but its execution falls outside the timeframe of the graduation project, so no user results are reported here. Each requirement is therefore marked as realised, designed but not yet tested, or still to do. This distinction matters: the tool is designed to meet all the requirements, but meeting them in design is not the same as demonstrating it in practice. The test protocol in Section 7.4 sets out exactly how that practical demonstration would be carried out.

7.2 Evaluation against the requirements

The table below assesses each requirement and gives it one of three statuses: realised (built and demonstrable), designed, to be tested (built into the design but not yet validated with users), or still to do (not yet built).

DR	Requirement	Status	Basis / what still needs testing
1	Fits a busy clinic	Designed, to be tested	The DI controls how many questions to ask; actual time use not yet measured in practice
2	Adapts to the ward	Realised	The tool runs a context analysis per department

3	Guides strategy selection	Realised	The matching step produces tailored strategies per case
4	Structures reflection	Realised	The phased flow and the validation step structure the DI's thinking
5	Surfaces determinants	Realised	The AI identifies factors and categorises them by COM-B
6	Matches determinants to strategies	Realised	Strategies are matched to the validated factors
7	Makes consequences visible	Designed, to be tested	The tool presents factors and rationale; whether this makes consequences feel tangible is not yet tested
8	Works iteratively	Realised	The phased structure allows moving between steps and revisiting
9	Supports collaboration	Realised	Built-in prompts and shareable PDF/worksheet exports support working with stakeholders
10	Needs no training	Designed, to be tested	The interface was simplified; ease of use without training not yet validated
11	Translates theory, layered	Partly; deeper layer still to do	The base process works; the optional deeper layer of guidance still needs to be designed and built
12	Is engaging to use	Designed, to be tested	The phased, interactive design aims for engagement; not yet tested with users
13	Tailors physician engagement	Realised	The ranking captures physician preferences; collaboration prompts support engagement

Table 6. *Evaluation overview*

Table 6 shows that the core functionality is realised, several requirements that concern the user's experience are designed but await a user test to confirm, and one requirement (DR11, the layered information) is partly done with a clear remaining step.

7.3 Practicality and relevance

Beyond the individual requirements, the tool can be judged against the two overarching criteria. This assessment is analytical, drawing on the design itself and on the stakeholder feedback from Chapter 6.

On relevance, the tool addresses the central need identified throughout the research: the gap between having compliance data and knowing what to do with it. By turning a conversation and a simplified ranking into tailored, evidence-grounded strategies, it targets exactly the step that the interviews showed was missing. The repeated stakeholder requests for the ranking, the multi-source matching, and the integration of the worksheets, all of which are now built in, suggest the concept matches what the users said they needed.

On practicality, the evidence is more conditional. The tool is designed to fit clinical time pressure, to work without training, and to be deployable across hospitals, but these are the qualities that depend most on real use, and they are the ones a user test is meant to confirm. The stakeholder enthusiasm during the feedback rounds is encouraging, but it is not the same as evidence from use in practice. This is the gap the test protocol in the next section is built to close.

7.4 A usability test ready to run

The analytical evaluation above shows what the tool is designed to do. The remaining question is whether DIs can actually operate it in practice. To answer that, I designed a full usability test protocol. It is written before any data collection, with the tasks, success criteria, and outcome measures fixed in advance, so that the eventual results can be judged against criteria set without knowledge of the outcome. This pre-specification is itself a piece of methodological work, and it means the test can be run with the design unchanged. The complete protocol is included in Appendix G; this section summarises its logic.

What the test asks. The study concentrates on usability, since desirability and relevance were addressed in the earlier feedback rounds. It asks whether DIs can complete each of the six workflow steps without critical errors or facilitator help, where in the workflow usability problems occur and how severe they are, how DIs rate the tool on a standardised instrument, and whether the tool fits the way DI work is actually carried out. It deliberately leaves three things out of scope: clinical effectiveness, the accuracy of the AI analysis against a gold standard, and long-term adoption. Each of those needs a separate study, and they are placed later in the roadmap.

How it works. The design is mixed-methods and within-subjects, so each participant is their own reference point. Five to eight DIs or DI trainees work through the full six-step workflow with a prepared, realistic scenario, framed as goals rather than click-by-click steps so that navigation problems surface. Two data streams run alongside each other. A concurrent think-aloud with structured observation captures where and why problems occur, since a participant who hesitates reveals a problem that a completion rate alone would miss. Task metrics and the System Usability Scale (Brooke, 1996) provide task completion, time, errors, interventions, and a single comparable usability score. The sample is purposive, selected for the DI role rather than at random, and five participants is the established threshold for surfacing the majority of usability problems in one round (Nielsen & Landauer, 1993).

How the findings would be read. The two streams are read together. A low usability score is interpreted through the observed problems behind it, and a step with many interventions is cross-checked against the think-aloud reasoning for that step. With this sample size the quantitative measures are used descriptively, to characterise the sample and to triangulate the qualitative findings, not to make statistical claims. Stating that limit plainly is part of the rigour: the test is built for problem discovery, not generalisation. The prepared scenario controls the conditions at the cost of some realism, and the think-aloud itself can slow or alter how a participant works. These limits are acknowledged so that the claims stay scoped to what the method can support.

Taken together, the protocol turns the open question of practicality into a concrete, repeatable procedure. The analytical evaluation establishes that the tool is designed to meet its requirements. The test, once run, would establish whether it meets them in the hands of the people it is built for. That step is the first item on the roadmap that follows.

8 Roadmap

This chapter looks beyond the graduation project to the future development of STIP-assist. With the prototype completed and pilot studies about to begin, the focus shifts from developing the concept to considering its long-term implementation and growth. The key question is how a research prototype developed for a single hospital can evolve into a practical tool that infection prevention departments across the Netherlands can adopt, maintain, and trust with patient data.

The roadmap describes this journey step by step. STIP-assist starts as a prototype at Amsterdam UMC, where it is tested and refined through pilot studies with infection prevention professionals. After successful validation, the tool can be developed into a stand-alone product and gradually introduced in other hospitals. The long-term vision is for behaviour-informed, evidence-based infection prevention to become part of everyday practice, without relying on a research project or a dedicated implementation team.

The roadmap is organised around two elements: horizons and tracks. The horizons show how STIP-assist develops over time, from the current prototype to a tool that can be used across the Netherlands. The tracks represent the five areas that need to develop alongside the tool, including technology, funding, governance, and ownership. A successful implementation depends on progress in all of these areas. A well-designed prototype still falls short without long-term funding, clear ownership, and support for implementation.

The chapter first describes the different horizons and then explains the five tracks that support this development. The full roadmap is shown at the end of Section 8.1, with a detail of each horizon alongside its description.

8.1 The journey, horizon by horizon

Each horizon represents a different stage in the development of STIP-assist. Before moving to the next stage, a number of key conditions must be met. These conditions act as decision points, helping to determine whether the project is ready to move forward. This keeps the roadmap realistic and ensures that each stage builds on a solid foundation.

Horizon 0 — Laying the foundation (current stage)

This is the starting point of the roadmap and represents the current situation (2026) (Figure 25). The tool exists as a working prototype with a first end-to-end workflow and a first technical architecture, running securely inside myDRE at Amsterdam UMC. At this stage, the focus lies on demonstrating that the core concept works reliably and provides value for infection prevention professionals. New features come later.

At this stage, three questions need to be answered. First, does the workflow support infection prevention professionals in moving from a conversation to identified barriers and finally to suitable strategies in a logical way? Second, does the technology work as intended by processing audio, generating transcripts, recognising barriers, and linking them to appropriate strategies within the secure myDRE environment? Finally, is the tool ready to handle real conversations safely by meeting requirements for data storage, privacy (AVG), and information security?

At this stage, the main challenge is building trust. Even if the technology works as intended, the tool will only be useful if infection prevention professionals recognise the identified barriers and consider the suggested strategies relevant to their situation.

Before moving to Horizon 1, several conditions must be met. The tool must run reliably within the myDRE environment and the complete workflow must be validated from start to finish. The Data Protection Impact Assessment (DPIA), AVG checks, and pilot protocol must be approved. At least one hospital department must be ready to participate in the pilot, and API key management must be in place. Finally, internal testing should show that infection prevention professionals recognise most of the identified barriers as accurate and relevant.

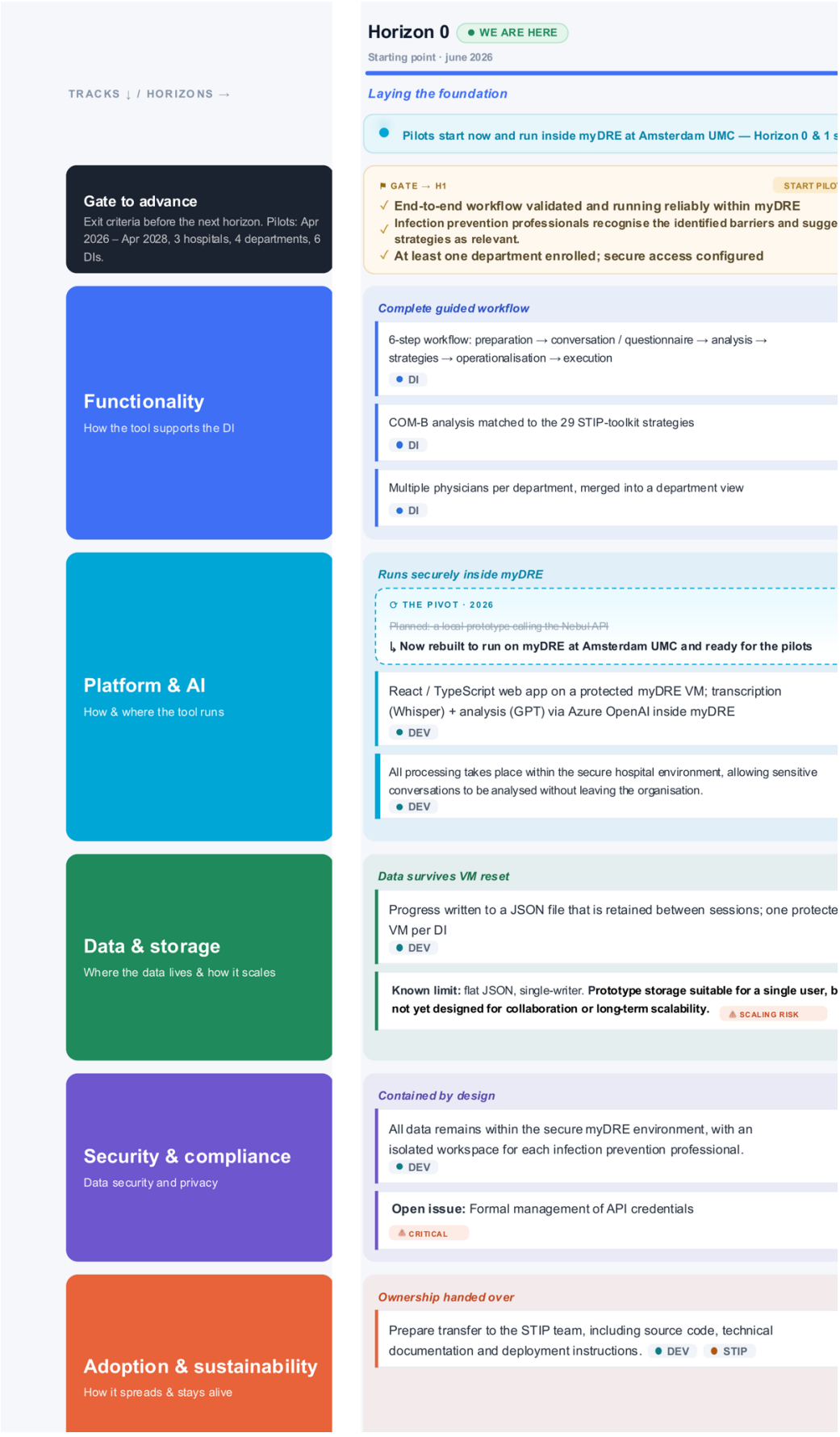


Figure 25. Horizon 0 in the roadmap: laying the foundation.

Horizon 1 — Pilots (starting now)

In this stage (Figure 26), STIP-assist is tested with real users in everyday practice. The focus shifts from whether the technology works to whether it is useful and fits the daily workflow of infection prevention professionals. The pilot studies take place within the same secure myDRE environment as Horizon 0, allowing the tool to be tested safely in a real clinical setting.

The pilot studies run from April 2026 to April 2028 across three hospitals and four departments: ZGV, LUMC, and the intensive care and neurology departments at Amsterdam UMC. In total, six infection prevention professionals participate, each working together with a clinical microbiologist and one or more physician representatives. During the pilots, infection prevention professionals use STIP-assist to support real ward conversations. The evaluation focuses on ease of use, the quality of the analyses, trust in the AI-generated recommendations, and the usefulness of the suggested implementation strategies.

This stage provides the first evidence of how STIP-assist performs in practice. It is also an important opportunity to identify areas for improvement. For example, the pilot studies can show whether relevant barriers are missed, whether the suggested strategies fit the local context, and whether the interface is easy to use. Feedback from users and issues identified during the pilots are used to improve the tool through iterative updates.

The pilot studies are evaluated mainly using qualitative data. The main focus is on whether infection prevention professionals and physicians find the approach valuable, whether the tailored strategies are received more positively than previous improvement projects or standard e-learning, whether collaboration between infection prevention professionals and physicians improves, and whether infection prevention professionals feel confident using the tool. Improvements in compliance are a welcome outcome. They are a secondary signal at this stage, behind the qualitative experience of the users.

At this stage, the main question is whether infection prevention professionals choose to use the tool as part of their daily work. Even if the tool works well during the pilot studies, it will only have a lasting impact if users continue to use it after the research team has stepped back. For this reason, long-term adoption is a key measure of success.

Before moving to Horizon 2, several conditions must be met. The pilot studies must be completed with a positive overall evaluation. Infection prevention professionals and physicians should experience the tool as usable, valuable, and a natural part of their daily work, while infection prevention professionals should feel confident using it independently. The tailored strategies produced with the tool should be received at least as positively as previous improvement approaches. In addition, no critical data protection or security incidents should have occurred during the pilots. Finally, a stand-alone version of STIP-assist with user authentication and support for multiple users must be available outside the myDRE environment.

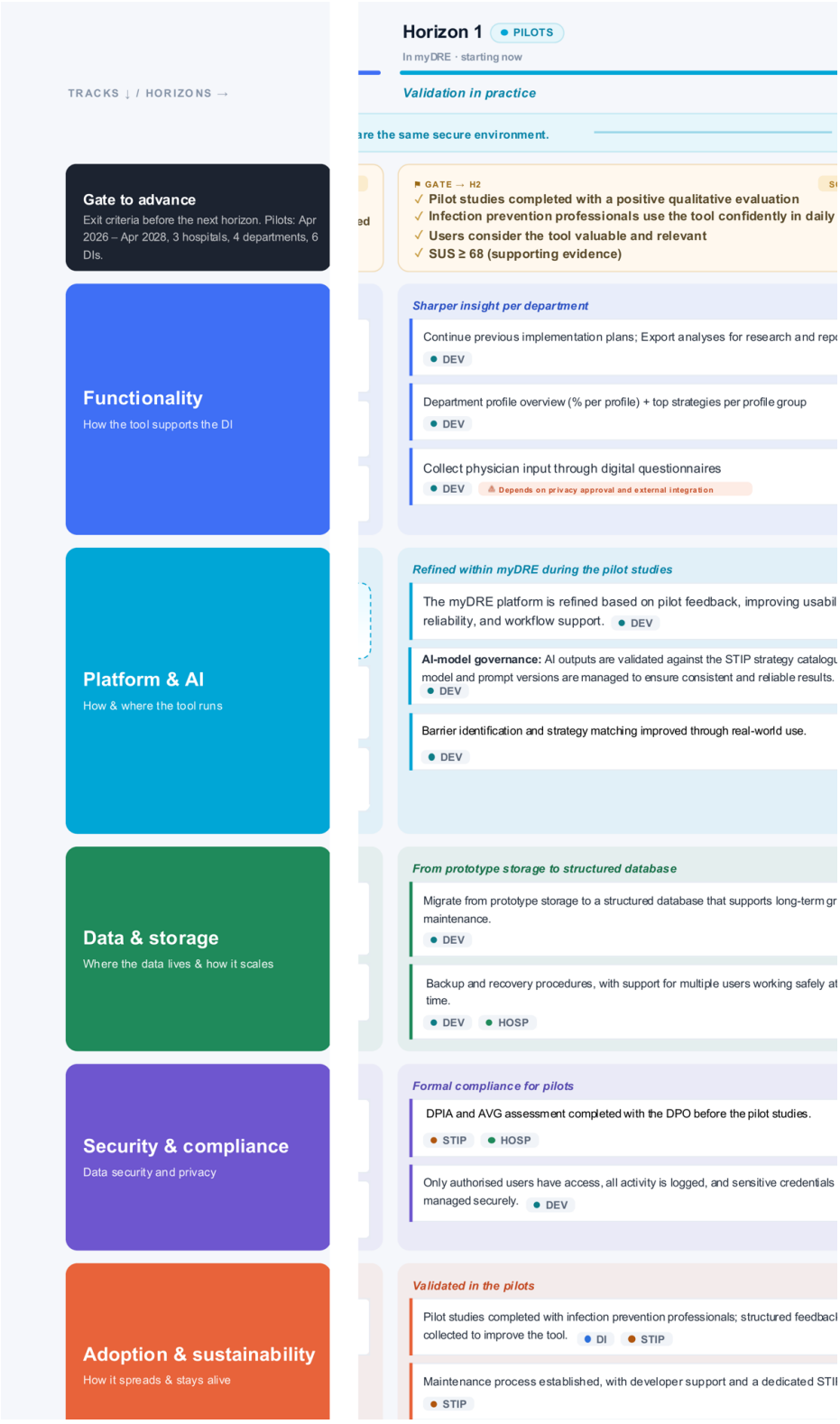


Figure 26. Horizon 1 in the roadmap: validation in practice through the pilots.

The System Usability Scale (SUS) is used to assess usability, with a score of 68 generally indicating acceptable performance. The decision to scale STIP-assist rests mainly on the experiences and feedback of users during the pilot studies. The SUS score supports that judgement; it does not replace it.

Horizon 2 — Scaling for wider adoption (mid-term, indicative 2028)

If the pilot studies are successful, the next challenge is to prepare STIP-assist for use beyond Amsterdam UMC (Figure 27). The focus shifts from demonstrating that the tool works to ensuring that it can be used in different hospitals and healthcare settings.

To achieve this, STIP-assist is developed into a stand-alone product that no longer depends on the myDRE environment. It must support multiple hospitals, users, and accounts, while remaining compatible with different IT systems. The technical architecture is further strengthened by introducing a hosted backend, secure user accounts for infection prevention professionals, and a flexible AI architecture that allows different AI models to be integrated when needed.

This is the phase where the tool changes from a project into a product. Alongside the technical work, harder questions move to the centre: maintenance, governance, ownership and funding.

One of the main challenges at this stage is ensuring that STIP-assist can be maintained after the research project has ended. While ownership of the tool is clear, long-term funding still needs to be established. The development and research team behind STIP will remain responsible for the tool, while infection prevention professionals continue to be its primary users.

The intention to further develop STIP-assist is clear, and an initial budget is available. However, the costs of long-term hosting, maintenance, and further development have not yet been fully determined. For this reason, the roadmap identifies a subscription model for participating hospitals as a possible funding approach, although this has not yet been formally agreed.

Before moving to Horizon 3, several conditions must be met. STIP-assist should be in active use in multiple hospitals outside Amsterdam UMC. A reliable approach for migrating data to the stand-alone platform must be available, and the operational costs per consultation should be sustainable. In addition, long-term maintenance must be organised within the STIP team, and a sustainable funding model should be established, with at least one hospital committed to paying for the service.

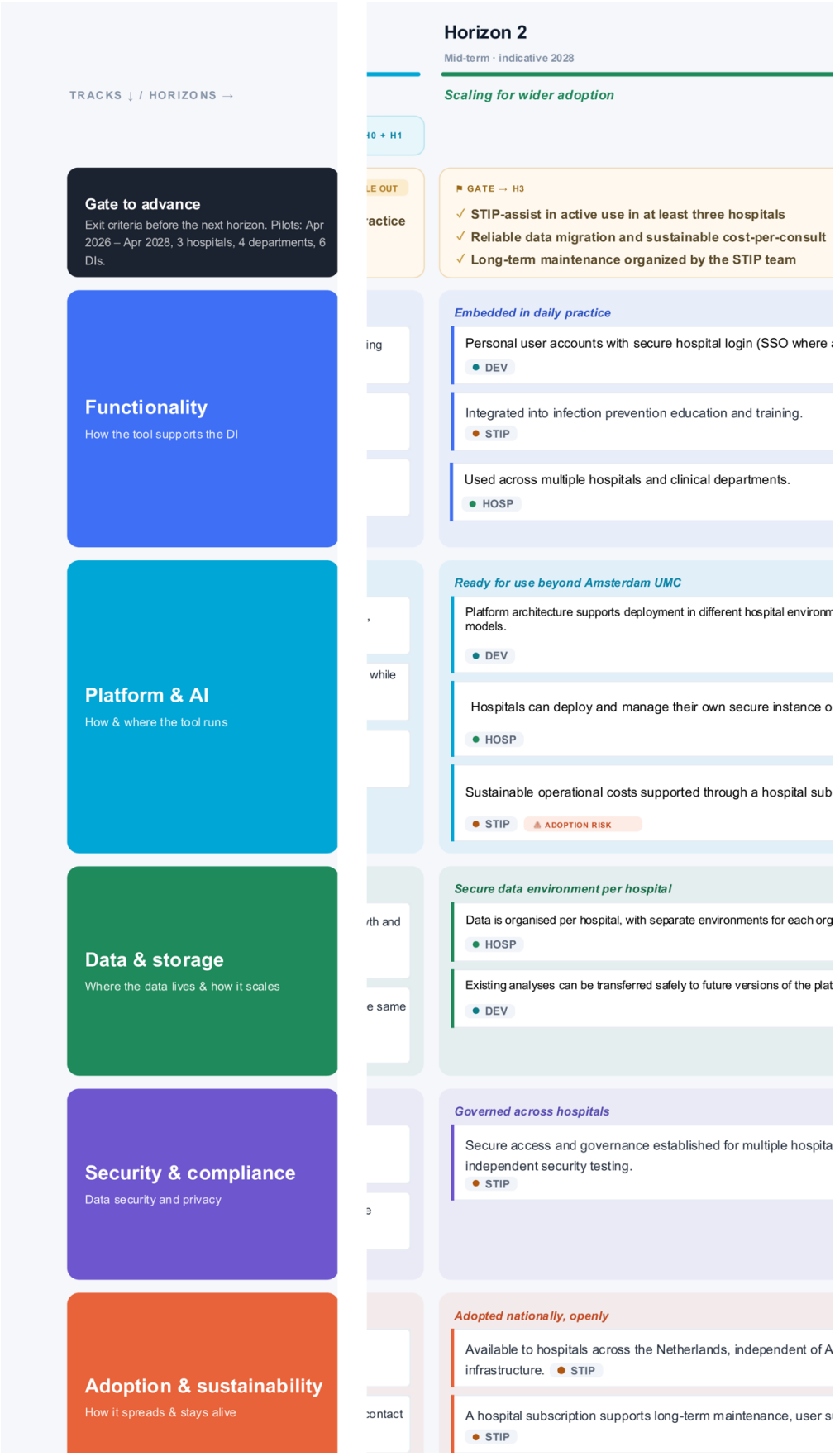


Figure 27. Horizon 2 in the roadmap: scaling for wider adoption.

Horizon 3 — National adoption and long-term sustainability (long-term, indicative 2030 and onwards)

This horizon represents the long-term vision for STIP-assist (Figure 28). By this stage, the tool has evolved beyond a single research project and has become part of routine infection prevention practice. The preferred implementation model is similar to ItFits: a central platform from which Dutch hospitals can obtain STIP-assist as a stand-alone application. Each hospital installs and manages the tool locally. There is no direct integration into an electronic health record system such as Epic or HiX.

This approach reduces the complexity of integration with hospital IT systems while allowing patient data to remain within the hospital's own secure environment. This supports privacy requirements and enables the conversation module to function as intended. Hospitals are responsible for running and maintaining the tool as part of their routine infection prevention activities.

One challenge that remains important in the long term is keeping STIP-assist relevant and up to date. Even a successful tool can lose its value if it is not maintained as clinical guidelines and AI technologies continue to evolve. To address this, STIP-assist should be supported through continuous monitoring, regular updates of AI models and prompts, and periodic reviews to ensure alignment with current infection prevention guidelines. A sustainable funding model, such as a hospital subscription, could provide the resources needed for long-term maintenance without relying on temporary research projects or grant funding.

The long-term vision is reached when STIP-assist is widely adopted by infection prevention departments across the Netherlands. The tool is maintained independently of Amsterdam UMC through a sustainable governance and funding model. By this stage, behaviour-informed, evidence-based infection prevention has become part of routine clinical practice.

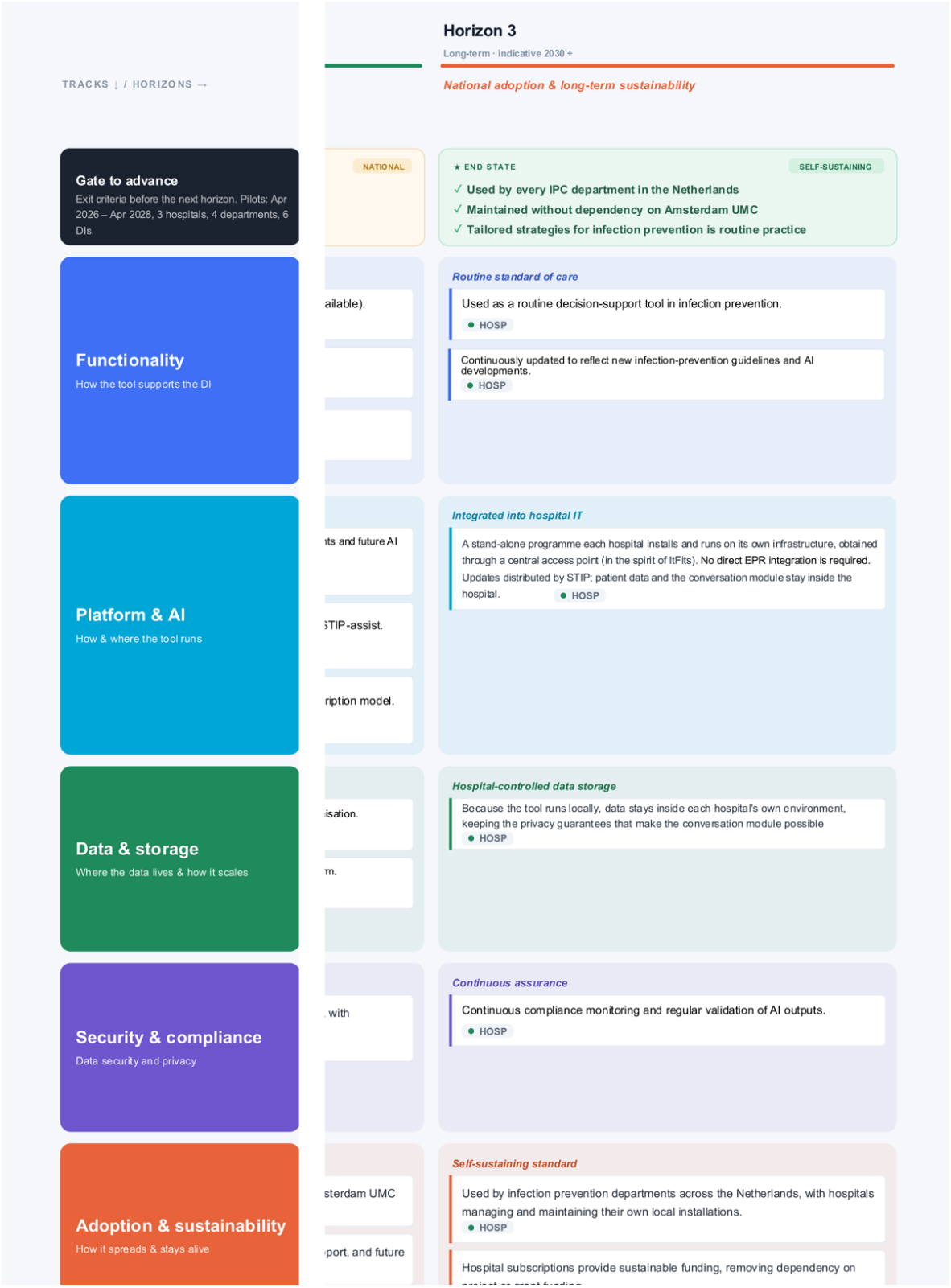


Figure 28. Horizon 3 in the roadmap: national adoption and long-term sustainability.

The complete roadmap is shown below across two pages (Figure 29 + 30): the left side covers the current state and the pilots (Horizon 0 and Horizon 1), and the right side covers scaling and national adoption (Horizon 2 and Horizon 3).

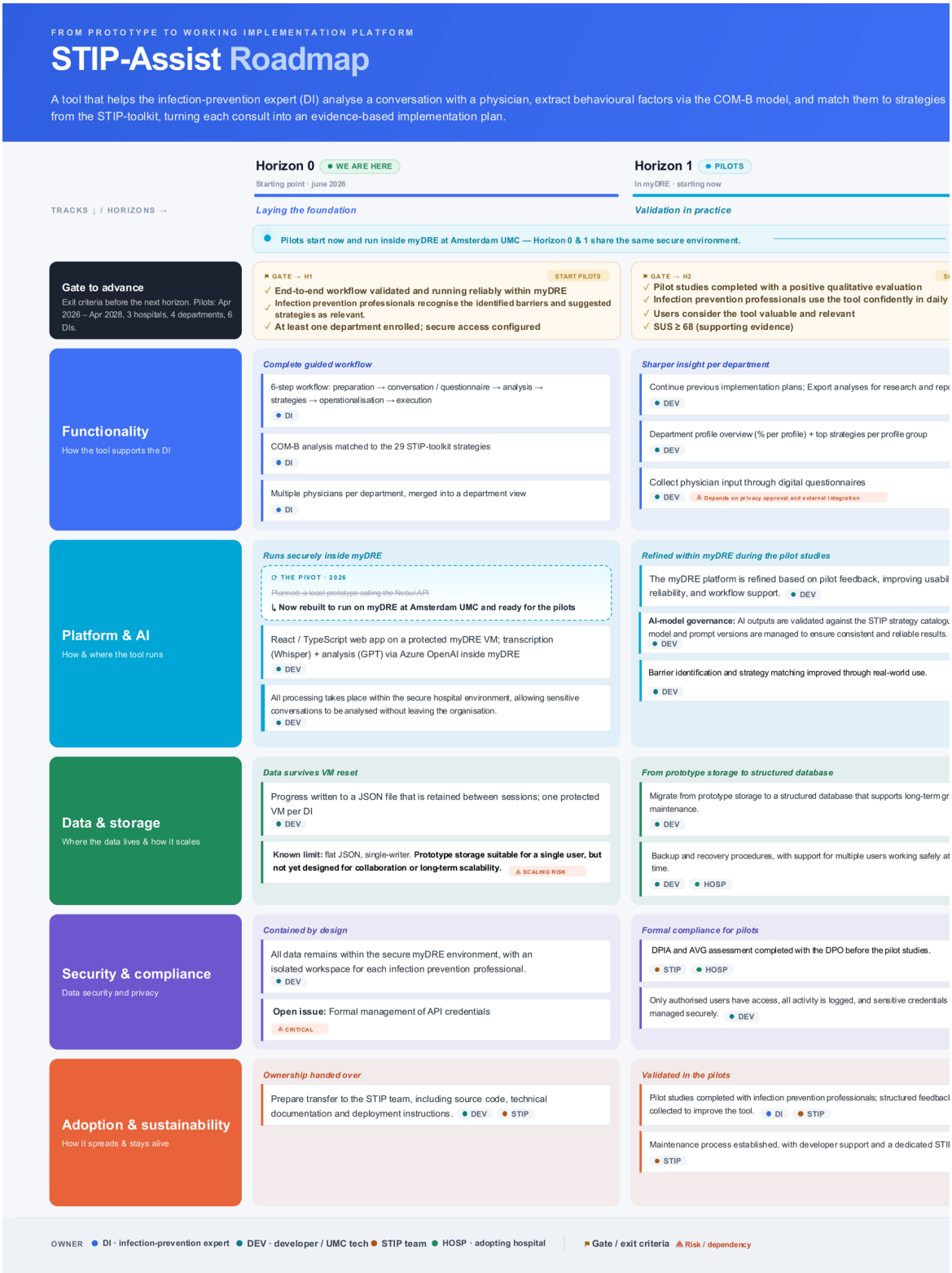


Figure 29. The full STIP-assist roadmap, left half: Horizon 0 and Horizon 1.

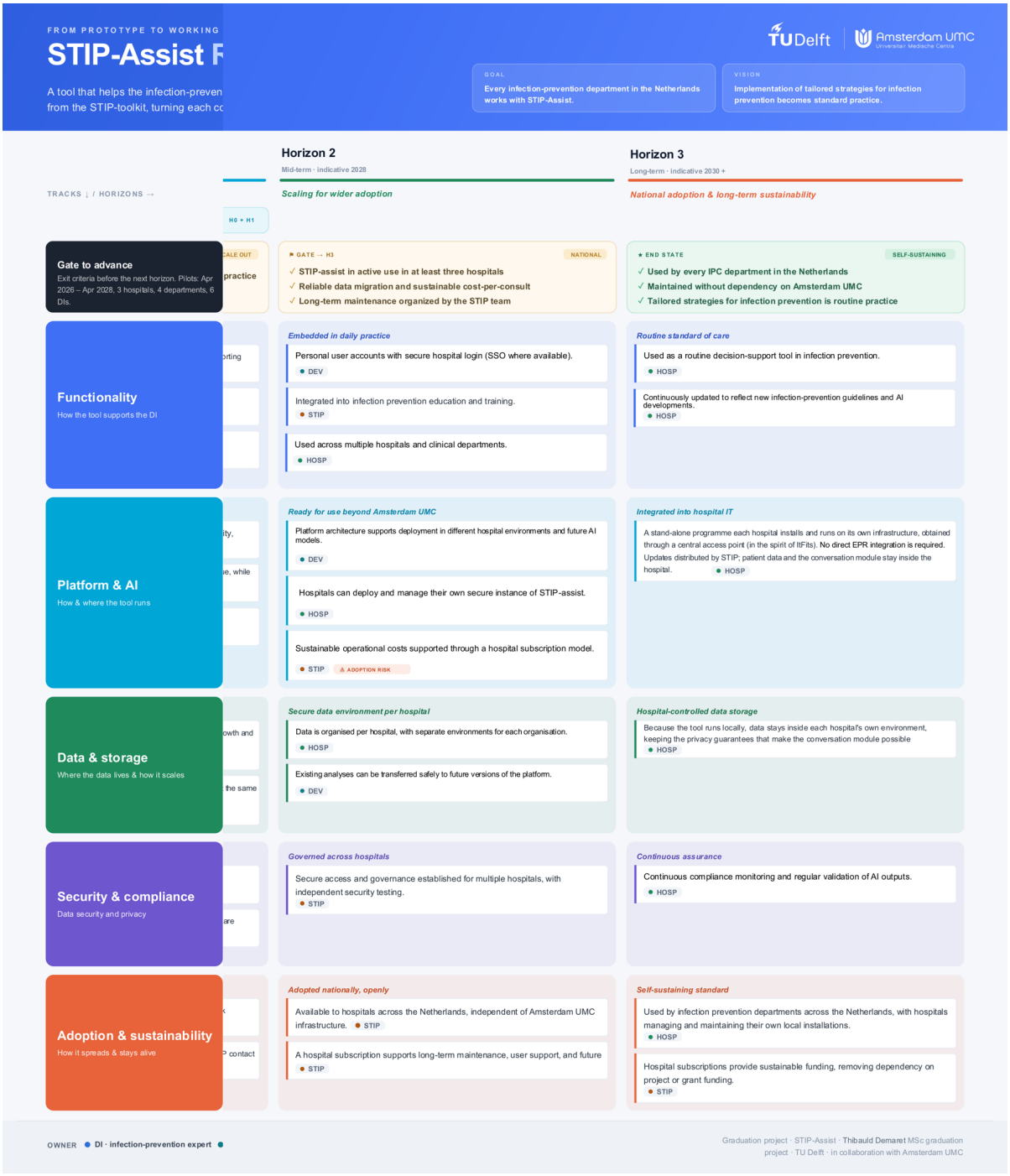


Figure 30. The full STIP-assist roadmap, right half: Horizon 2 and Horizon 3.

8.2 The five tracks of the roadmap

The roadmap considers more than the technical development of STIP-assist. It is organised around five tracks that need to progress together to support successful implementation. Technical development is one track among five. The tool also requires sustainable maintenance, governance, funding, data management, and compliance. By considering these five tracks together, the roadmap highlights the key factors needed to support the long-term adoption and use of STIP-assist.

1 Functionality — what the infection profesionals can do with the tool

The functionality of STIP-assist expands over time. It starts with a guided six-step workflow that supports infection prevention professionals throughout the entire process, from preparation and data collection to barrier analysis, strategy selection, operationalisation, and implementation. The tool uses the COM-B framework to match identified barriers with the 29 strategies from the STIP toolkit and combines results from multiple physicians into a department-level overview.

As the tool develops, additional functionality is introduced. This includes more detailed insights at the department level, an overview of behavioural profiles, and questionnaires that can be sent directly to physicians. In the longer term, STIP-assist supports personal user accounts for infection prevention professionals, becomes part of infection prevention education, and is intended to become a routine part of infection prevention practice across the Netherlands.

2 Platform and AI — how and where the tool runs

This track describes how the technical platform of STIP-assist develops over time. The tool was originally designed as a local prototype using the Nebul API. To support the pilot studies, it was later rebuilt to run securely within the myDRE environment at Amsterdam UMC. The current version is a React and TypeScript web application running on a protected virtual machine. Audio is transcribed using Whisper, while GPT models hosted through Azure OpenAI analyse the conversations. All processing takes place within the secure hospital environment.

In Horizon 2, the platform is prepared for use beyond Amsterdam UMC. A hosted backend and secure user accounts are introduced, together with a flexible AI architecture that allows different AI models to be used without changing the rest of the application.

In Horizon 3, STIP-assist is intended to become a stand-alone application that hospitals can install and manage locally. Access is provided through a central platform, similar to ItFits, while patient data remains within each hospital's own secure IT environment, with no integration into electronic patient record systems.

3 Data and storage — how the data scales

This track describes how the data management of STIP-assist develops over time. The current prototype uses a simple JSON file to store progress, allowing information to remain available even after the daily reset of the virtual machine. Each infection prevention professional works within their own protected virtual machine. This solution is sufficient for the prototype. Long-term use needs more, since the file does not support multiple users, automatic backups, or large amounts of data.

As STIP-assist develops, the data storage is moved to a structured database with version control and data migration support. In the long term, each hospital has its own isolated data environment, with a validated process for transferring data when needed. Because the tool is installed and managed

locally, patient data remains within each hospital's own secure IT environment, helping to meet privacy and security requirements.

4 Security and compliance — how the tool remains secure

Security and compliance are essential throughout the development of STIP-assist. The current prototype already operates within the secure myDRE environment, where each infection prevention professional works in an isolated virtual machine. One important step that still needs to be completed is the formal management of API keys and other sensitive credentials. In addition, the Data Protection Impact Assessment (DPIA) and AVG assessment with the Data Protection Officer (DPO) must be completed before the pilot studies can begin.

As STIP-assist develops, additional security measures are introduced. These include role-based access control, audit logging, and a formal policy for managing API keys. In the longer term, the platform is extended with secure access for multiple hospitals, independent security testing, and continuous monitoring to ensure ongoing compliance with privacy and security requirements.

5 Adoption and sustainability — how the tool is adopted and maintained

This track focuses on the long-term use of STIP-assist. Developing a working tool is only the first step; it also needs to be maintained, supported, and adopted in everyday practice. The first step is to transfer ownership to the STIP team, including the source code, technical documentation, and installation instructions. During the pilot studies, feedback from infection prevention professionals is used to further improve the tool while a maintenance process is established within the STIP team.

As STIP-assist develops, the goal is to make it available to hospitals across the Netherlands as a stand-alone application that no longer depends on the Amsterdam UMC infrastructure. To support long-term maintenance, updates, and user support, the roadmap proposes a sustainable funding model, such as a subscription for participating hospitals. The long-term vision is a self-sustaining system in which hospitals routinely use STIP-assist while ongoing maintenance is supported through a stable organisational and funding structure.

8.3 Strategic value of the roadmap

A successful implementation tool requires more than a working prototype. It also depends on adoption, governance, scalability, long-term maintenance, and integration into everyday clinical practice. The roadmap shows how these elements can develop over time while also identifying the main challenges that need to be addressed at each stage. By recognising these challenges, the roadmap provides a realistic direction for the future development of STIP-assist.

The work presented in this graduation project forms the foundation for this roadmap. It delivers the Horizon 0 prototype and supports the start of the pilot studies in Horizon 1. The later horizons describe a possible pathway for further development, showing how STIP-assist could evolve into a widely adopted implementation tool for infection prevention. (The complete roadmap can be found in Appendix H.)

9 Reflection, recommendations and contribution to STIP

9.1 Reflection

Looking back at the project, the research phase went well. Although my Medisign specialisation had already given me experience with healthcare projects, this one brought a new level of complexity. Infection prevention and implementation science were new to me, as was the strong emphasis on translating theory into practice. In the first weeks I was motivated to build that understanding, and through literature, interviews, and stakeholder interactions I gained insight into both the behavioural side of implementation and the difficulty of applying it in a real hospital setting.

The harder part came after the research phase, when I moved into ideation. I found it difficult to narrow the scope. The problem space is broad and complex, while the actual intervention focuses on a fairly specific interaction between the DI and the physician. Translating that complexity into a focused design direction took more effort than I expected, and it is the part of the process I learned the most from. With hindsight, committing to a direction sooner, even an imperfect one, would have given me more time to iterate. The pressure of delivering a working prototype for the pilot eventually forced that decision, and in a way that pressure helped.

This period also coincided with a more hectic personal phase, including illness and moving house, which affected my focus. Despite that, I was able to keep the project moving in a constructive way and to make meaningful progress.

Overall, this project strengthened my ability to work within complex healthcare systems and to turn broad, research insights into focused, actionable design directions. As a strategic designer, I learned that part of the value I bring is exactly this: holding a complex problem without oversimplifying it, while still committing to a concrete, usable result.

9.2 Limitations

This project has clear limitations. They follow from the choices I made under the constraints of a graduation timeline, and naming them honestly is part of judging what the work does and does not yet show.

The tool has not been tested with users. The most important limitation is that the final version of STIP-assist has not yet been evaluated with infection prevention professionals in practice. The evaluation in this report is analytical: it assesses the tool against the design requirements and draws

on stakeholder feedback from the design phase. A full usability test was designed for this purpose, but its execution falls outside the project window. This means the claims about practicality stay at the level of design intent. Whether DIs can operate the tool effectively, and whether it fits their daily routine, remains to be demonstrated. The protocol in Chapter 7 sets out exactly how that demonstration would be carried out.

The AI analysis has not been validated against a reference standard. The tool uses AI to turn a conversation into factors and to categorise them by COM-B. The quality of that analysis has not been checked against an independent reference, such as the same conversations coded by trained researchers. The design keeps the DI in control by asking them to confirm or reject every factor, which limits the risk of a wrong analysis passing through unnoticed. The accuracy of the AI itself is still an open question, and it is one the pilots are meant to start answering.

The departmental view averages across physicians. When several physicians are analysed together, the tool shows the COM-B picture for the department as an average of the individual scores. An average can hide disagreement. A ward where half the physicians score high and half score low looks the same as a ward where everyone sits in the middle. For the pilot the averaged view is enough to give the DI direction, but a future version should show the spread and flag strong disagreement so that the variation on a ward stays visible.

9.3 Recommendations

Based on the findings of this project, several recommendations can be made for the further development of STIP-assist.

First, the upcoming pilot should focus on validating the tool in daily practice rather than on technical functionality alone. Particular attention should be paid to how infection prevention professionals integrate the tool into their existing workflow, whether the proposed implementation strategies are perceived as relevant and actionable, and how the tool influences conversations with physicians.

Second, future development should prioritise reducing the effort required for tailored implementation. Throughout the project, users consistently emphasised the importance of a simple, intuitive workflow that supports decision making without increasing workload. Maintaining this balance between scientific rigour and ease of use should remain a key design principle.

Third, further development should focus on preparing the platform for wider implementation. This includes completing the transition to a multi-user architecture, strengthening authentication and security, and establishing a sustainable approach to hosting, maintenance, and governance. These technical and organisational aspects are essential for scaling the tool beyond the initial pilot sites.

Finally, future research should evaluate not only the usability of STIP-assist but also its long-term impact on implementation practice. While improvements in infection prevention compliance remain

an important ambition, the primary outcome should be whether the tool enables infection prevention professionals to consistently select and apply tailored implementation strategies with confidence. The strategic roadmap presented in Chapter 8 provides a phased framework for guiding these next steps.

9.4 My contribution to STIP

My contribution to the STIP programme has two sides. First, I brought a different perspective. As a strategic designer, I approached infection prevention not only as a research problem but as a design challenge, focused on the people who have to use the strategies and the context they work in. This perspective helped translate the programme's behavioural research into something practical and user-centred.

Second, I have delivered a working product in some sense. STIP-assist takes the programme's research, its strategies, and the DCE, and turns them into a usable tool that infection prevention professionals can actually work with. It gives the programme a concrete foundation to build on, test further, and develop, rather than a concept on paper. In that sense, the project moves the STIP research one step closer to practice.

References

During this graduation project, I used several AI-based tools. I used Gemini and ChatGPT to generate images. I used Perplexity to occasionally check my spelling, and NotebookLM to organise and query the academic papers I worked with. I used Claude to support the development of the prototype. After using these tools, I thoroughly reviewed and edited the content as needed, taking full responsibility for the final outcome.

Albers, B., Caci, L., Blum, K., Boland, G., Carrara, E., Gentil, P. R., Nguyen, A., Pollard, J., Schechner, V., Sonpar, A., Yanagida, T., Zingg, W., & Clack, L. (2025). Examining tailoring as an implementation strategy for reducing healthcare-associated infections across European acute care hospitals (REVERSE): study protocol for a hybrid type 2 effectiveness-implementation trial. *Trials*, *26*(1), 418. https://doi.org/10.1186/s13063-025-09132-x

Albers, B., Rapley, T., Nilsen, P., & Clack, L. (2023). Editorial: Tailoring in implementation science. *Frontiers in Health Services*, *3*, 1233597. https://doi.org/10.3389/frhs.2023.1233597

Atkins, L., Francis, J., Islam, R., O'Connor, D., Patey, A., Ivers, N., Foy, R., Duncan, E. M., Colquhoun, H., Grimshaw, J. M., Lawton, R., & Michie, S. (2017). A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems. *Implementation Science*, *12*(1), 77. https://doi.org/10.1186/s13012-017-0605-9

Bredin, D., O'Doherty, D., Hannigan, A., & Kingston, L. (2022). Hand hygiene compliance by direct observation in physicians and nurses: a systematic review and meta-analysis. *Journal of Hospital Infection*, *130*, 20–33. https://doi.org/10.1016/j.jhin.2022.08.013

Brooke, J. & Redhatch Consulting Ltd. (1996). SUS - a quick and dirty usability scale. In *Usability Evaluation in Industry* (p. (pp.189-194)). Taylor & Francis. https://www.researchgate.net/profile/John-Brooke-6/publication/319394819_SUS_--_a_quick_and_dirty_usability_scale/links/5ee5c59e458515814a5e6bb8/SUS--a-

quick-and-dirty-usability-scale.pdf?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InB1YmxpY2F0aW9uIiwicGFnZSI6InB1YmxpY2F0aW9uIn19

Bührmann, L., Schuurmans, J., Ruwaard, J., Fleuren, M., Etzelmüller, A., Piera-Jiménez, J., Finch, T., Rapley, T., Potthoff, S., Aouizerate, B., Batterham, P. J., Calear, A., Christensen, H., Pedersen, C. D., Ebert, D. D., Van Der Eycken, E., Fanaj, N., Van Genugten, C., Hanssen, D., . . . Vis, C. (2020). Tailored implementation of internet-based cognitive behavioural therapy in the multinational context of the ImpleMentAll project: a study protocol for a stepped wedge cluster randomized trial. *Trials*, *21*(1), 893. https://doi.org/10.1186/s13063-020-04686-4

Cantrell, D., Shamriz, O., Cohen, M. J., Stern, Z., Block, C., & Brezis, M. (2008). Hand hygiene compliance by physicians: Marked heterogeneity due to local culture? *American Journal of Infection Control*, *37*(4), 301–305. https://doi.org/10.1016/j.ajic.2008.05.001

Erasmus, V., Daha, T. J., Brug, H., Richardus, J. H., Behrendt, M. D., Vos, M. C., & Van Beeck, E. F. (2010). Systematic Review of Studies on Compliance with Hand Hygiene Guidelines in Hospital Care. *Infection Control and Hospital Epidemiology*, *31*(3), 283–294. https://doi.org/10.1086/650451

Finch, T. L., Potthoff, S., May, C. R., Girling, M., Perkins, N., Vis, C., Bührmann, L., Etzelmueller, A., Van Genugten, C. R., Schuurmans, J., Piera-Jiménez, J., Rapley, T., Hoogendoorn, A., O'Connor, A., Whitton, A., Calear, A., Meksi, A., Rømer, A. S., Etzelmüller, A., . . . Sacco, Y. (2024). How is tailored implementation undertaken using a self-guided toolkit? Qualitative study of the ItFits-toolkit in the ImpleMentAll project. *Implementation Science*, *19*(1), 48. https://doi.org/10.1186/s13012-024-01380-w

Flores, A., Wrigley, M., Askew, P., Craig, R., Egan, B., Towey, L., & Shawe, J. (2020). Use of non-sterile gloves in the ward environment: an evaluation of healthcare workers' perception of risk and decision making. *Journal of Infection Prevention*, *21*(3), 108–114. https://doi.org/10.1177/1757177420907687

Fontaine, G., Michie, S., Beidas, R. S., Geng, E., Fahim, C., Powell, B. J., Welch, V., Thomas, J., Chan, J., Abbasgholizadeh-Rahimi, S., Légaré, F., Hastings, J., Lambert, S. D., Presseau, J., Straus, S. E., An, R., Saran, A., & Taylor, N. (2026). AI Methods for Implementation Science (AIM-IS): developing a framework, toolkit, and reporting standard for the responsible use of AI in implementation practice and research. *Implementation Science*, *21*(1). https://doi.org/10.1186/s13012-026-01503-5

Hastings, J., Glauer, M., West, R., Kleinau, A., Thomas, J., Wright, A. J., & Michie, S. (2025). Predicting outcomes of smoking cessation interventions in novel scenarios using ontology-informed, interpretable machine learning. *Wellcome Open Research*, *8*, 503. https://doi.org/10.12688/wellcomeopenres.20012.2

Huis, A., Van Achterberg, T., De Bruin, M., Grol, R., Schoonhoven, L., & Hulscher, M. (2012). A systematic review of hand hygiene improvement strategies: a behavioural approach. *Implementation Science*, *7*(1), 92. https://doi.org/10.1186/1748-5908-7-92

Jackson, C., Lowton, K., & Griffiths, P. (2013). Infection prevention as "a show": A qualitative study of nurses' infection prevention behaviours. *International Journal of Nursing Studies*, *51*(3), 400–408. https://doi.org/10.1016/j.ijnurstu.2013.07.002

Kelly, Y., O'Rourke, N., Hegarty, J., Gannon, J., Flynn, R., & Keyes, L. M. (2024). The co-design of a digitally supported intervention for selecting implementation tools and actions for standards (SITAS). *BMC Health Services Research*, *24*(1), 1582. https://doi.org/10.1186/s12913-024-12075-9

Lancsar, E., & Louviere, J. (2008). Conducting discrete choice experiments to inform healthcare decision making: A user's guide. *Pharmacoeconomics*, 661–677. https://www.researchgate.net/publication/5231028

Lee, K. (2013). Student and infection prevention and control nurses' hand hygiene decision making in simulated clinical scenarios: a qualitative research study of hand washing, gel and glove use choices. *Journal of Infection Prevention*, *14*(3), 96–103. https://doi.org/10.1177/1757177413484784

Loveday, H., Lynam, S., Singleton, J., & Wilson, J. (2013). Clinical glove use: healthcare workers' actions and perceptions. *Journal of Hospital Infection*, *86*(2), 110–116. https://doi.org/10.1016/j.jhin.2013.11.003

Michie, S., Johnston, M., Francis, J., Hardeman, W., & Eccles, M. (2008). From Theory to intervention: Mapping theoretically derived behavioural determinants to behaviour change techniques. *Applied Psychology*, *57*(4), 660–680. https://doi.org/10.1111/j.1464-0597.2008.00341.x

Michie, S., Van Stralen, M. M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, *6*(1), 42. https://doi.org/10.1186/1748-5908-6-42

Nielsen, J. (1994). *Heuristic evaluation* (p. 25). https://dl.acm.org/citation.cfm?id=189200.189209

Nielsen, J., Landauer, T. K., & Bellcore. (1993). *A mathematical model of the finding of usability problems* [Journal-article]. https://course.ccs.neu.edu/is4800/ssl/nielsen93.pdf

Rosenbluth, G., Garritson, S., Green, A. L., Milev, D., Vidyarthi, A. R., Auerbach, A. D., & Baron, R. B. (2015). Achieving hand hygiene success with a partnership between graduate medical education, hospital leadership, and physicians. *American Journal of Medical Quality*, *31*(6), 577–583. https://doi.org/10.1177/1062860615596567

Schutte, M., Dekker, M., Sikkens, J., & Van Mansfeld, R. (2025). Between heuristic and deliberative thinking: a multi-center qualitative study of physicians' decision-making in infection prevention practice. *Antimicrobial Resistance and Infection Control*, *14*(1), 50. https://doi.org/10.1186/s13756-025-01572-z

Schutte, M., Van Mansfeld, R., De Vries, R., & Dekker, M. (2024). Determinants of compliance with infection prevention measures by physicians: a scoping review. *Journal of Hospital Infection*, *153*, 30–38. https://doi.org/10.1016/j.jhin.2024.08.011

Simonse, L. W. L. & Delft University of Technology, Faculty of Industrial Design Engineering. (2024). *Design Roadmapping: Design Roadmapping Guidebook for Future Foresight Techniques* (Jianne Whelton, Ed.). https://orcid.org/0000-0001-8860-6522

Skolarus, T. A., Hawley, S. T., Wittmann, D. A., Forman, J., Metreger, T., Sparks, J. B., Zhu, K., Caram, M. E. V., Hollenbeck, B. K., Makarov, D. V., Leppert, J. T., Shelton, J. B., Shahinian, V., Srinivasaraghavan, S., & Sales, A. E. (2018). De-implementation of low value castration for men with prostate cancer: protocol for a theory-based, mixed methods approach to minimizing low value androgen deprivation therapy (DeADT). *Implementation Science*, *13*(1). https://doi.org/10.1186/s13012-018-0833-7

Smiddy, M. P., Murphy, O. M., Savage, E., & Browne, J. P. (2019). The influence of observational hand hygiene auditing on consultant doctors' hand hygiene behaviors: A qualitative study. *American Journal of Infection Control*, *47*(7), 798-803.e1. https://doi.org/10.1016/j.ajic.2018.12.024

Squires, J. E., Linklater, S., Grimshaw, J. M., Graham, I. D., Sullivan, K., Bruce, N., Gartke, K., Karovitch, A., Roth, V., Stockton, K., Trickett, J., Worthington, J., & Suh, K. N. (2014). Understanding practice: Factors that influence physician hand hygiene compliance. *Infection Control and Hospital Epidemiology*, *35*(12), 1511–1520. https://doi.org/10.1086/678597

Sunbow Productions (Director). (1985). *G.I. Joe: A real American hero*. Marvel Productions.

Tomsic, I., Ebadi, E., Gossé, F., Hartlep, I., Schipper, P., Krauth, C., Schock, B., Chaberny, I. F., & Von Lengerke, T. (2021). Determinants of orthopedic physicians' self-reported compliance with surgical site infection prevention: results of the WACH-trial's pilot survey on COM-B factors in a German university hospital. *Antimicrobial Resistance and Infection Control*, *10*(1), 67. https://doi.org/10.1186/s13756-021-00932-9

Vis, C., Schuurmans, J., Aouizerate, B., Craggs, M. A., Batterham, P., Bührmann, L., Calear, A., Pashoja, A. C., Christensen, H., Dozeman, E., Pedersen, C. D., Ebert, D. D., Etzelmueller, A., Fanaj, N., Finch, T. L., Hanssen, D., Hegerl, U., Hoogendoorn, A., Mathiasen, K., . . . Riper, H. (2022). Effectiveness of self-guided tailored

implementation strategies in integrating and embedding Internet-Based cognitive behavioral therapy in routine mental health care: results of a multicenter Stepped-Wedge cluster randomized trial. *Journal of Medical Internet Research*, *25*, e41532. https://doi.org/10.2196/41532

Von Lengerke, T., Ebadi, E., Schock, B., Krauth, C., Lange, K., Stahmeyer, J. T., & Chaberny, I. F. (2019). Impact of psychologically tailored hand hygiene interventions on nosocomial infections with multidrug-resistant organisms: results of the cluster-randomized controlled trial PSYGIENE. *Antimicrobial Resistance and Infection Control*, *8*(1), 56. https://doi.org/10.1186/s13756-019-0507-5

Von Lengerke, T., Tomsic, I., Krosta, K. M. E., Ebadi, E., Keil, V., Buchta, F., Luz, J. K., Schaumburg, T., Kolbe-Busch, S., & Chaberny, I. F. (2023). Tailoring implementation interventions of different order in infection prevention and control: A cascadic logic model (IPC-CASCADE). *Frontiers in Health Services*, *2*, 960854. https://doi.org/10.3389/frhs.2022.960854

West, R., Bonin, F., Thomas, J., Wright, A. J., Mac Aonghusa, P., Gleize, M., Hou, Y., O'Mara-Eves, A., Hastings, J., Johnston, M., & Michie, S. (2023). Using machine learning to extract information and predict outcomes from reports of randomised trials of smoking cessation interventions in the Human Behaviour-Change Project. *Wellcome Open Research*, *8*, 452. https://doi.org/10.12688/wellcomeopenres.20000.1

White, C. M., Statile, A. M., Conway, P. H., Schoettker, P. J., Solan, L. G., Unaka, N. I., Vidwan, N., Warrick, S. D., Yau, C., & Connelly, B. L. (2012). Utilizing improvement science methods to improve physician compliance with proper hand hygiene. *PEDIATRICS*, *129*(4), e1042–e1050. https://doi.org/10.1542/peds.2011-1864

List of figures

List of figures

Figure 1. Key stakeholder map

Figure 2. Overview of the design process followed during the project

Figure 3. The Behaviour Change Wheel (Michie et al., 2011)

Figure 4. Stakeholder relationship map

Figure 5. Stakeholder rings: core, direct, indirect

Figure 6. Idea generation sheet: the fifteen initial ideas

Figure 7. Clustering of the ideas into six directions

Figure 8. Storyboard of the STIP-assist concept, illustrating the intended six-step flow

Figure 9. Storyboard of the Implementation Co-Creation concept

Figure 10. Dashboard

Figure 11. Conversation module

Figure 12. Transcript screen

Figure 13. Barrier analysis with COM-B profile

Figure 14. Strategy recommendations

Figure 15. From strategy to implementation

Figure 16. The six-phase structure, shown here in the Preparation phase

Figure 17. The simplified ranking that captures physicians' preferences for strategy attributes

Figure 18. The analysis screen, where the AI presents the identified factors, categorised by COM-B, for the DI to validate

Figure 19. The strategy screen, where each recommendation combines the fit with the factors, the local physician preferences, and the research

Figure 20. A collaboration prompt, asking the DI to discuss the result with the team before continuing

Figure 21. The conversation module inside myDRE

Figure 22. The analysis screen running on real AI

Figure 23. The home screen

Figure 24. The do-phase

Figure 25. Horizon 0 in the roadmap: laying the foundation

Figure 26. Horizon 1 in the roadmap: validation in practice through the pilots

Figure 27. Horizon 2 in the roadmap: scaling for wider adoption

Figure 28. Horizon 3 in the roadmap: national adoption and long-term sustainability

Figure 29. The full STIP-assist roadmap, left half: Horizon 0 and Horizon 1

Figure 30. The full STIP-assist roadmap, right half: Horizon 2 and Horizon 3

List of tables

Table 1. Main research findings

Table 2. Interview overview

Table 3. List of design requirements

Table 4. Comparison between Concept 1 and Concept 2

Table 5. Results of the Mentimeter at the VHIG congress

Table 6. Evaluation overview

Appendix

List of appendices

Appendix A. Project Brief Amsterdam UMC

Appendix B. Interviews

Appendix C. Pilot Gelderse Vallei

Appendix D. VHIG Congress

Appendix E. Mentimeter VHIG congress

Appendix F. List of design requirements

Appendix G. Test protocol

Appendix H. Strategic roadmap

Appendix A. Project Brief Amsterdam UMC

Background

Application of infection prevention measures, such as hand hygiene and wearing personal protective equipment, is essential to prevent hospital-acquired infections and the spread of resistant micro-organisms. However, compliance with these measures by healthcare workers is suboptimal. Traditional interventions using strategies such as clinical lessons and campaigns often do not have lasting effects, especially not with physicians. Achieving behaviour change of physicians thus requires a different approach. Context also plays an important role; for example, in the hospital physicians work at different wards and in varying team compositions. This asks for tailored strategies to involve physicians in infection prevention and stimulate adequate compliance.

Aim of this graduation project

The aim is to develop an implementation tool that helps hospitals to select tailored strategies to improve application of infection prevention by physicians. The content of the tool (strategies) are pre-defined; this was based on relevant theories and determined together with the end users. This project focuses on the design of the tool, with consideration for user needs to develop a user-friendly, practical, relevant decision aid to select suitable strategies.

What will you do?

The student has the opportunity to contribute to a practically relevant problem in healthcare. This entails:

• Involving different stakeholders (physicians from different departments, clinical microbiologists, infection prevention practitioners; Dutch language proficiency required)

• Exploring the needs and barriers of stakeholders for the tool

• Designing the tool prototype

• Iterative testing and improving of the tool

Client

This project is offered by the department of Medical Microbiology and Infection Prevention of Amsterdam UMC. You will be part of a young research team with supervision by a PhD student and a Postdoc researcher.

Appendix B Interviews

Interview with infection prevention professional

To gain contextual insight into infection prevention (IP) practice within the hospital setting, a semi-structured interview was conducted with a senior infection prevention professional (Deskundige Infectiepreventie) at Amsterdam UMC. The professional has extensive experience in clinical settings and advises departments on infection prevention policies and implementation. The purpose of this interview was to better understand: • How infection prevention is organised in practice • Which behavioural challenges occur in daily clinical routines • How physicians and nurses respond to infection prevention measures • Which tensions exist between policy, implementation, and practice

Organisational context of infection prevention Infection prevention within the hospital is structured around protocols, audits, hand hygiene observations, and departmental contact persons. Standard measures include dress code policies, hand hygiene protocols, and cleaning procedures for medical equipment. Monitoring primarily occurs through: • Periodic hand hygiene observations • Annual or bi-annual audits • Feedback reports to departments • Informal signalling via departmental contact persons However, feedback on infection rates is not systematically integrated into daily practice. According to the professional, infection outcome data is often fragmented across departments and not consistently linked to behavioural interventions. This limits the ability to directly evaluate the impact of specific implementation strategies.

Behavioural barriers in practice Several recurring barriers were identified:

1. Perceived time pressure Time constraints were frequently mentioned as a reason for non-compliance. For example, proper hand disinfection requires approximately 30 seconds of rubbing until hands are dry. Although objectively short, this duration is perceived as long in high-pressure clinical environments. Importantly, this suggests that the issue is not purely time scarcity, but perceived opportunity cost.

2. Invisible risk A critical insight concerns "invisible contamination." When surfaces are visibly soiled (e.g., blood), staff are more inclined to clean them. However, when contamination is not visible, motivation decreases. This indicates that behavioural compliance is strongly influenced by the perceptibility of risk.

3. Differences between professional groups The professional described a cultural difference between nurses and physicians: • Nurses are generally accustomed to following structured protocols. • Physicians are more likely to question whether a protocol is necessary in a specific situation. Physicians may rely more heavily on their own judgement and professional autonomy when deciding whether to apply infection prevention measures. This highlights a potential mismatch between standardised policy approaches and physician decision-making styles.

Current implementation approaches Traditional implementation strategies primarily include clinical lessons (educational sessions). However, these are often delivered at the end of shifts, when staff are fatigued, limiting their impact. More interactive approaches appear more promising. The professional described the use of: • Black light demonstrations to visualise hand hygiene quality • Markers to test whether equipment was properly cleaned • An infection prevention "escape room" that simulated realistic clinical scenarios The escape room approach was perceived as particularly effective, as it: • Engaged participants actively • Created experiential learning • Increased memorability • Stimulated ongoing peer discussion However, such approaches require significant time and resources, limiting scalability.

Structural challenges in engaging physicians The professional indicated that physicians are more difficult to reach through traditional infection prevention networks. While infection prevention teams often have structured contact with nursing staff, physician engagement depends heavily on local relationships and departmental culture. This suggests that implementation strategies targeting physicians may require different communication channels and framing compared to strategies targeting nurses.

Implications for the toolkit design The interview revealed several implications for the design of the STIP toolkit:

1. Implementation strategies should go beyond information provision and address behavioural mechanisms such as social norms, risk perception, and professional identity.
2. The toolkit should support context-sensitive strategy selection, as barriers vary between departments and professional groups.
3. Strategies that increase visibility of consequences or make behaviour tangible may be particularly effective.
4. Engagement of physicians requires specific attention to autonomy, credibility, and peer influence. Overall, the interview reinforced the need for a tailored, theoretically grounded approach that moves beyond generic educational interventions.

Interview with observator infection prevention

The interview with a hand hygiene observer, responsible for conducting hand hygiene observations within the hospital, provided valuable insights into the operational realities of infection prevention monitoring. This perspective highlighted how formal protocols interact with daily practice, and where structural and behavioural tensions emerge.

Hand hygiene compliance is monitored according to a yearly planning structure in which high-risk departments (e.g., ICU, emergency department, haematology) are observed annually, while lower-risk departments are measured every two years. Observations focus exclusively on the five WHO hand hygiene moments, scored as either performed or not performed. However, in practice, the measurement process is adapted pragmatically. Rather than strictly following the monthly schedule, observations are spread across the year to avoid overexposure of departments and to reduce measurement bias. This indicates that implementation is already being informally tailored at the operational level.

Several contextual differences between departments were identified. For example, on the ICU, computers are considered part of the patient environment, creating additional disinfection moments compared to regular wards. Some departments, such as the NICU, apply stricter local rules than those stated in official guidelines. These micro-contextual differences suggest that infection prevention is not implemented uniformly, but embedded within local cultures and routines.

A key behavioural insight concerns the influence of hierarchy and peer dynamics. During ward rounds, compliance among physicians strongly depended on the behaviour of the leading consultant. If the senior physician performed hand hygiene, others followed; if not, compliance dropped significantly. This demonstrates the importance of social norms and modelling behaviour in physician groups.

The observer also described recurrent behavioural challenges. The use of gloves often creates a false sense of security, leading staff to move between patient rooms and the corridor without appropriate disinfection. Additionally, mobile phone use after patient contact frequently results in missed hygiene moments. These behaviours appear to be driven less by knowledge deficits and more by habit, risk perception, and workflow interruptions.

Another important finding relates to the perception of monitoring. Staff responses to observations were mixed. While some departments responded positively, others experienced the measurements as control rather than support. The presence of observers visibly altered behaviour, indicating a strong Hawthorne effect. Although this may temporarily improve compliance, it also raises questions about the validity of measurements and the sustainability of behavioural change. Finally, significant frustrations were reported regarding the digital registration system (K2). Compared to the previously used system, the current tool is experienced as time-consuming, unintuitive, and lacking feedback functionality. Observers do not receive insight into aggregated results, and departments do

not have direct access to their own performance data. As a result, monitoring remains disconnected from immediate learning or motivation.

Overall, this interview revealed that hand hygiene compliance is shaped by contextual variation, hierarchical influence, perceived control dynamics, workflow interruptions, and limited feedback loops. These findings reinforce the need for implementation strategies that move beyond generic educational approaches and instead address social norms, real-time feedback, and local adaptation mechanisms.

To gain insight into how infection prevention policies are monitored in daily hospital practice, a semi-structured interview was conducted with a hand hygiene observer, responsible for performing hand hygiene observations on behalf of the infection prevention department at Amsterdam UMC. This role focuses on conducting structured compliance measurements across multiple hospital departments according to a predefined annual plan. The purpose of this interview was to better understand: • How hand hygiene observations are organised in practice • How monitoring interacts with daily clinical workflows • Which behavioural patterns are repeatedly observed • Which tensions arise during measurement and feedback processes • How monitoring is perceived by healthcare professionals

Organisational context of hand hygiene monitoring Hand hygiene compliance is monitored according to an annual planning structure. High-risk departments (e.g., ICU, emergency department, haematology, NICU) are measured annually, while lower-risk departments are observed once every two years. Each department must reach a minimum number of measurements per year. Observations are based exclusively on the five WHO hand hygiene moments, which are scored dichotomously (performed / not performed). Observers conduct the measurements, while infection prevention specialists analyse the aggregated results. In practice, however, the formal planning structure is applied flexibly. Rather than strictly following monthly allocations, observations are spread throughout the year to: • Reduce measurement bias • Avoid repeated monitoring within short timeframes • Minimise resistance from departments This illustrates that even monitoring processes require contextual adaptation. Furthermore, important contextual variation exists between departments. For example: • On ICU wards, computers are considered part of the patient environment, creating additional hand hygiene moments. • Some departments (e.g., NICU) apply stricter internal rules than official guidelines prescribe. These micro-contextual differences indicate that infection prevention practices are not uniform, but embedded in local routines and interpretations.

Behavioural patterns observed in practice The observer identified several recurring behavioural patterns.

1. Glove-related risk misperception Staff frequently leave patient rooms wearing gloves and touch objects in the corridor without performing hand hygiene. Gloves appear to create a false sense of safety, reducing perceived urgency for disinfection. This suggests that compliance failures are often linked to risk perception rather than lack of knowledge.
2. Mobile phone use Healthcare professionals sometimes handle their phone immediately after patient contact without prior disinfection. Since phones are rarely cleaned between patients, this behaviour creates potential contamination loops. This reflects habitual behaviour combined with workflow interruptions.

3. Hierarchical modelling during ward rounds During physician ward rounds, compliance strongly depends on the behaviour of the senior consultant. If the leading physician performs hand hygiene, others follow. If not, compliance decreases. This demonstrates the importance of social modelling and peer influence in physician groups. Importantly, the observer did not note major structural differences in compliance between nurses and physicians. However, nurses encounter significantly more hand hygiene moments due to the nature of their tasks.

Perception of monitoring and behavioural reactivity Monitoring is not experienced uniformly across departments. While some departments respond positively, others perceive the presence of observers as a form of control. Reactions vary depending on departmental culture and familiarity with infection prevention staff. A strong Hawthorne effect was reported: healthcare professionals visibly adjust their behaviour when observers are present. Staff sometimes explicitly warn colleagues ("infection prevention is here"), resulting in temporary behavioural correction. Although this may increase compliance during measurement, it raises questions about whether observations reflect habitual behaviour. This dynamic illustrates a tension between monitoring as control and monitoring as support. When monitoring is perceived as evaluative rather than developmental, it may trigger defensive reactions rather than intrinsic motivation.

Structural and digital barriers Significant frustration was expressed regarding the current digital data-entry system (K2). Compared to the previously used system, K2 is experienced as: • Time-consuming • Complex • Repetitive in data entry • Lacking real-time feedback Observers must often record measurements manually and re-enter them later, increasing workload and reducing usability. Furthermore, observers do not have access to aggregated performance data, and departments do not receive direct real-time insight into their own results. As a result, monitoring remains disconnected from immediate learning or behavioural reinforcement.

Educational and structural tensions An additional tension concerns perceived educational gaps. According to feedback heard on departments, hand hygiene is reportedly no longer consistently emphasised in professional education programmes. This leads to a mismatch where senior nurses feel responsible for teaching basic infection prevention to trainees, despite limited time and capacity. This reflects a broader systemic issue: responsibility for infection prevention is distributed across education, clinical practice, and organisational policy, but not always aligned.

Implications for the toolkit design The interview highlights several implications for the STIP toolkit:

1. Monitoring alone does not ensure sustainable behaviour change; feedback loops and motivational components are crucial.
2. Implementation strategies must account for hierarchical dynamics and peer modelling effects.
3. Risk perception (e.g., glove use, invisible contamination) should be explicitly addressed.

4. Strategies should aim to reduce the perception of control and increase perceived support.
5. Digital systems that support implementation must prioritise usability and meaningful feedback.
6. Local contextual variation between departments should be incorporated into strategy selection. Overall, the interview demonstrates that hand hygiene compliance is shaped by behavioural norms, hierarchical influence, workflow interruptions, digital infrastructure, and perceived control dynamics. These findings further support the need for a tailored, theoretically grounded implementation approach that moves beyond generic monitoring or educational interventions.

Interview with an infection prevention professional

To gain further insight into how infection prevention is monitored and managed within the hospital, a semi-structured interview was conducted with an infection prevention professional (Deskundige Infectiepreventie) at Amsterdam UMC. This professional works closely with clinical departments and is involved in monitoring infection prevention practices, analysing compliance data, and coordinating improvement initiatives. The purpose of this interview was to understand: • how infection prevention performance is monitored within the hospital • which data sources are used to evaluate infection prevention practices • how monitoring results are communicated and interpreted • how these insights are currently used to guide improvement efforts

Monitoring of infection prevention practices Infection prevention performance at Amsterdam UMC is monitored through multiple data sources and evaluation mechanisms. One important monitoring system consists of hand hygiene observations, which are collected and visualised through internal dashboards developed in Power BI. These dashboards provide an overview of hand hygiene compliance across the organisation, including: • overall compliance percentages • compliance trends over time • differences between departments and divisions • compliance per professional group • compliance per hand hygiene indication (e.g. before patient contact, after contact with bodily fluids) For example, the dashboards show that overall hand hygiene compliance within the hospital fluctuates around approximately 55–60%. However, considerable variation exists between departments and professional groups. In addition to hand hygiene observations, the infection prevention team conducts regular audits on hospital wards. These audits evaluate broader infection prevention practices, such as environmental hygiene, availability of disinfectants, waste management, and adherence to protocols. The results of these audits are documented in the K2 tracer system, which assesses compliance with infection prevention standards through structured checklists.

Interpretation and use of monitoring data The collected monitoring data are primarily used to identify areas where infection prevention practices can be improved. Infection prevention professionals analyse the dashboards and tracer results to detect patterns, such as departments with lower compliance rates or specific infection prevention measures that are not consistently followed. These insights are typically discussed with clinical departments during feedback sessions or quality improvement meetings. In this process, infection prevention professionals act as advisors who interpret monitoring results and support departments in identifying possible improvement actions. However, while the monitoring systems provide extensive data about infection prevention performance, translating these insights into concrete implementation strategies remains a complex task. Infection prevention professionals must interpret the data, identify potential causes of non-compliance, and determine appropriate interventions for the specific clinical context.

Organisational context of infection prevention The professional also provided insight into how infection prevention is organised within Amsterdam UMC. The infection prevention department

operates within the broader structure of the Department of Medical Microbiology and Infection Prevention. Within the hospital, infection prevention professionals collaborate with various stakeholders, including: • physicians • nurses • contact persons for infection prevention within departments • observers conducting hand hygiene measurements • microbiologists and researchers These stakeholders play different roles in monitoring, implementing, and evaluating infection prevention practices. Because infection prevention professionals interact with many actors across the organisation, their role often involves coordinating between monitoring data, clinical practice, and improvement initiatives.

Implications for the toolkit design The interview highlighted the central role of monitoring data in infection prevention work. While the hospital already has extensive dashboards and audit systems that provide insight into compliance levels, these tools mainly focus on reporting and monitoring performance. The process of translating monitoring results into concrete implementation strategies still relies heavily on the expertise and interpretation of infection prevention professionals. These insights suggest that a toolkit supporting infection prevention professionals in analysing implementation challenges and selecting appropriate improvement strategies could help bridge the gap between monitoring data and practical implementation actions.

Interview with creator of ItFits toolkit

Interview with an implementation-toolkit developer To gain insight into the development and practical use of implementation toolkits, a semi-structured interview was conducted with a senior researcher at Amsterdam UMC involved in developing the ItFits toolkit. The purpose of this interview was to better understand how implementation theory can be translated into a usable decision-support tool and to identify lessons relevant for the design of the STIP toolkit. The interview focused on four main aspects: • the purpose and functioning of the ItFits toolkit, • the design and development process, • challenges encountered during toolkit creation, • and user experiences during implementation.

Purpose and functioning of the ItFits toolkit The ItFits toolkit was developed to support healthcare organisations in implementing eHealth interventions into routine care. Rather than providing predefined solutions, the toolkit guides users through a structured process to identify local implementation barriers and select appropriate implementation strategies. The toolkit follows a stepwise implementation logic in which users:

1. identify implementation challenges,
2. prioritise relevant barriers,
3. match these barriers to suitable strategies,
4. develop an actionable implementation plan. An important insight discussed during the interview is that the toolkit primarily supports structured reflection and decision-making, instead of acting as a prescriptive intervention tool.

Design process and development challenges According to the developer, one of the main challenges during development was translating complex implementation science frameworks into an interface that healthcare professionals could realistically use in practice. While implementation theories provide systematic guidance, they are often perceived as abstract or difficult to apply by practitioners. During development, the team discovered that real-world users rarely follow a strictly linear implementation process. Many professionals already believe they understand the problem and prefer to move directly towards solutions. As a result, the toolkit had to balance methodological structure with flexibility. This led to several key design decisions: • allowing users to enter the process at different stages, • simplifying theoretical language, • supporting guided decision-making without forcing users through rigid steps.

User interaction and practical use The interview highlighted that toolkits such as ItFits are often not used continuously. Instead, users tend to engage with the toolkit during specific implementation phases, after which the learned approach is applied independently in practice. Users particularly struggled with: • formulating implementation problems explicitly, • selecting priorities among multiple barriers, • and translating insights into concrete actions. Facilitation or external guidance was therefore sometimes necessary, especially during early use.

Key lessons for toolkit design Several insights from the ItFits experience are directly relevant for the development of the STIP toolkit: • Implementation support tools should allow flexibility rather than enforce a fixed workflow. • Successful adoption depends on alignment with existing professional routines. • Toolkits should focus on supporting decision-making instead of prescribing behaviour. • Stakeholder involvement plays a crucial role in implementation success. • Translating theory into

practical actions requires careful simplification without losing credibility. Overall, the interview emphasised that implementation challenges are highly context-dependent and that effective toolkits must support users in adapting strategies to their local environment rather than providing universal solutions.

Interview PhD researcher AI in Healthcare

To explore the technical feasibility of integrating AI into the STIP toolkit, an interview was conducted with a PhD researcher at Amsterdam UMC working on AI applications in healthcare. The purpose of this interview was to assess whether the concept of an AI-supported decision tool could realistically be developed within the Amsterdam UMC infrastructure, and to identify technical and organisational constraints relevant for the design. The interview focused on four main aspects: • the feasibility of AI-supported analysis of conversations, • the available infrastructure within Amsterdam UMC, • the requirements for developing and deploying such a system, • and potential constraints related to privacy and data governance.

Feasibility of AI-supported concept The discussed concept involves using large language models to support the analysis of conversations between infection prevention professionals and physicians. This includes speech-to-text transcription, classification of barriers, and the suggestion of tailored implementation strategies. According to the researcher, such a system is technically feasible in principle. Current developments in natural language processing allow for the extraction of structured insights from unstructured text data, including conversational transcripts. This suggests that the core functionality of the concept could be realised from a technical perspective. However, feasibility strongly depends on how the system is implemented and the type of input data that is used. Fully automated analysis of raw conversation data introduces additional complexity compared to more structured input formats.

Infrastructure within Amsterdam UMC The researcher indicated that Amsterdam UMC already provides several environments that could support the development of AI applications. In particular, platforms such as MyDRE enable secure data storage and processing within the hospital's internal infrastructure. These environments are designed to ensure that sensitive data remains within institutional boundaries, which is essential when working with healthcare-related data. In addition, access to computational resources, such as GPU clusters, may be available for more advanced model development. This suggests that the concept could potentially be developed within an internal, secure infrastructure, rather than relying on external services.

Development considerations The interview highlighted that building such a system would require more than just technical implementation. Several components would need to be developed and aligned, including: • a clearly defined structure for identifying and categorising barriers, • a database linking barriers to implementation strategies, • a decision logic or model for matching input to strategies, • and an interface that supports user interaction and interpretation of results. Additionally, scalability across multiple hospitals would require careful consideration of interoperability and data governance across institutions.

Privacy and data governance A key point raised during the interview is the importance of addressing privacy and data governance from the start. The use of conversational data, especially audio recordings, introduces significant challenges in terms of data sensitivity and compliance. The researcher therefore advised exploring privacy constraints further before continuing development. Based on this recommendation, a follow-up meeting with a privacy specialist was scheduled to better understand the limitations and requirements.

Key lessons for toolkit design Several insights from this interview are directly relevant for the development of the STIP toolkit: • AI-supported analysis of implementation challenges is technically feasible, but highly dependent on infrastructure and governance. • Internal platforms such as

MyDRE offer opportunities for secure data processing within a hospital setting. • Privacy considerations are a major constraint, especially when working with conversational data. • Starting with structured or semi-structured input may be a more feasible first step than fully automated speech analysis. • The integration of AI should support decision-making rather than replace human judgement. Overall, the interview emphasised that while AI offers promising opportunities to support implementation processes, its application must be carefully aligned with technical, organisational, and ethical constraints.

Meeting with data protection officer

To explore the privacy and data governance implications of the proposed AI-supported concept, a semi-structured interview was conducted with a privacy expert at Amsterdam UMC, on April 13th. This meeting followed the earlier discussion with the AI-in-healthcare researcher and aimed to better understand the feasibility of the concept from a privacy perspective. The interview focused on four main aspects: • the conditions under which data can be used and processed, • the sensitivity and origin of the data, • the requirements for compliance within a hospital setting, • and the overall feasibility of implementing the concept within these constraints.

Privacy considerations and data sensitivity During the meeting, it became clear that the feasibility of the concept strongly depends on the specific design of the data collection and processing. Key questions raised included: • Where does the data originate from? • What type of information is included in the data? • To what extent is the data traceable to individuals or departments? Particularly when working with conversational data, such as interviews or recordings, the level of sensitivity increases significantly. This requires careful consideration of how data is captured, stored, and processed.

Conditions for implementation The officer emphasised that there is no single straightforward answer to whether the concept is allowed. Instead, feasibility depends on multiple factors, including: • the structure of the data (structured vs. unstructured), • the presence of personal or sensitive information, • the way data is stored and processed, • and the intended use of the system (research vs. operational tool). This indicates that privacy considerations are not a fixed barrier, but rather a set of conditions that must be addressed through design decisions.

Feasibility within Amsterdam UMC Despite the complexity, the officer indicated that the concept is in principle feasible within the Amsterdam UMC context, provided that it is designed in accordance with privacy regulations and institutional guidelines. This suggests that the concept does not need to be excluded, but that privacy must be integrated as a core design constraint from the early stages of development.

Implications for concept development The meeting highlighted that privacy is not only a technical requirement, but also a design challenge. Decisions regarding input formats, data processing, and system architecture directly influence feasibility. In particular, using fully unstructured conversational data (such as audio recordings) introduces more complexity than working with structured or summarised input. This suggests that a more controlled form of data input may be a more realistic starting point.

Key lessons for toolkit design Several insights from this meeting are directly relevant for the development of the STIP toolkit: • The feasibility of AI-supported solutions depends on how data is collected, structured, and processed. • Privacy constraints are highly context-dependent and must be considered early in the design process. • Working with structured or semi-structured input can significantly reduce complexity and risk. • The distinction between research use and operational deployment affects what is allowed. • Privacy should be treated as a design parameter rather than a limitation. Overall, the interview emphasised that while the concept is potentially feasible, its implementation requires careful alignment with privacy regulations, making design decisions around data handling critical for success.

Meeting with an AI architect

To further explore the technical implementation and scalability of the AI-supported concept, an interview was conducted with an AI architect on April 20th. This person is a data scientist at Amsterdam UMC, specialising in artificial intelligence and its application in healthcare, with experience in AI system design and high-performance computing. The purpose of this meeting was to assess how the concept could be technically realised and deployed, both within and beyond Amsterdam UMC, as the toolkit is intended to be tested across multiple hospitals. The interview focused on four main aspects: • the structure and functionality of the proposed concept, • the technical requirements for development and deployment, • the possibilities for external accessibility, • and available infrastructures within Amsterdam UMC to support development.

Structure of the proposed concept During the meeting, the concept was discussed in detail. The toolkit is designed to support infection prevention professionals in selecting tailored implementation strategies based on identified local barriers. The proposed workflow includes:

1. A conversation between the infection prevention professional and a physician,
2. Automatic transcription of this conversation using speech-to-text,
3. Analysis of the transcript to identify potential barriers,
4. User validation of identified barriers (accepting or rejecting them),
5. Matching selected barriers to appropriate implementation strategies,
6. Providing structured recommendations to support decision-making. This concept builds on existing research within the STIP project and aims to translate theoretical knowledge into practical support.

Technical implementation and infrastructure The architect explained several technical pathways for implementing such a system. Within Amsterdam UMC, there are environments available for developing and testing AI applications, including internal AI experimentation environments and high-performance computing resources. The architect indicated that it may be possible to develop the concept within an internal "AI playground," where access to models, tools, and computational resources is facilitated. This environment would allow for controlled experimentation and prototyping while maintaining alignment with institutional requirements.

External accessibility and scalability A key topic of discussion was the requirement to make the toolkit accessible beyond Amsterdam UMC, as it will be tested in multiple hospitals. The architect mentioned that there may be possibilities to use shared endpoints or APIs that enable controlled external access. For example, secure API connections could allow other hospitals to interact with the system without directly accessing internal infrastructure. However, this introduces additional considerations regarding security, authentication, and data governance.

Development considerations The discussion highlighted that implementing such a system requires careful architectural design. Key considerations include: • how data is processed and where it is stored, • how users interact with the system, • how different components (transcription, analysis, strategy matching) are integrated, • and how access is managed across different institutions. In addition, the architect suggested that the project could potentially benefit from participation in internal innovation environments, which could provide access to resources and support for further development.

Key lessons for toolkit design Several insights from this meeting are directly relevant for the development of the STIP toolkit: • The concept is technically feasible, but requires careful system architecture and integration. • Internal AI environments provide opportunities for safe prototyping and experimentation. • Scalability across multiple hospitals introduces additional technical and organisational complexity. • External accessibility requires secure interfaces such as APIs. • Design decisions around infrastructure and accessibility are closely linked to privacy and governance considerations. Overall, the meeting emphasised that while the concept can be realised technically, its success depends on aligning technical feasibility with organisational and cross-institutional requirements.

Appendix C

Pilot particpating hospital

To gain insight into the practical implementation context of infection prevention interventions, an observational visit was conducted during a pilot session on April 9th. The purpose of this visit was to observe real-world interactions between infection prevention professionals and physicians, and to understand how implementation challenges manifest in daily clinical practice.

The pilot visit focused on four main aspects:
• the interaction between infection prevention professionals and physicians,
• the practical challenges of conducting implementation activities,
• the reception of new ideas and innovations,
• and early reflections on the applicability of the STIP toolkit concept.

Interaction between stakeholders

During the pilot, direct interaction between an infection prevention professional and a medical specialist was observed. The specialist indicated that he was already aware of infection prevention practices and considered himself engaged with the topic. He also expressed openness towards new ideas and innovations in this area.

This interaction highlighted that physicians may not always perceive themselves as part of the problem, but rather as already contributing positively. At the same time, openness to innovation suggests potential opportunities for engagement, provided that interventions are aligned with their perspective and daily practice.

Practical implementation challenges

An important challenge discussed during the pilot was related to the execution of baseline measurements. Infection prevention professionals indicated that conducting these measurements in practice can be difficult, for example due to limited time, access, or alignment with clinical workflows.

During the conversation, the physician suggested that infection prevention professionals could benefit from directly experiencing clinical practice, for example by shadowing physicians on the ward. This would provide better contextual understanding and could improve the accuracy and relevance of observations.

Observations on context and behaviour

The pilot emphasised that implementation takes place in a complex and context-dependent environment. Behaviour is influenced not only by knowledge, but also by routines, professional identity, and situational constraints.

The observed interaction also illustrated the importance of mutual understanding between infection prevention professionals and physicians. Bridging this gap appears to be a key factor in successful implementation.

Relevance for concept development

The pilot provided early opportunities to present the concept to infection prevention professionals and gather initial reactions. While this feedback was informal, it helped to validate the relevance of focusing on practical usability and stakeholder interaction within the toolkit.

Key lessons for toolkit design

Several insights from the pilot are directly relevant for the development of the STIP toolkit:

• Physicians may perceive themselves as already engaged, which requires careful framing of interventions.
• Successful implementation depends on alignment with daily clinical practice and routines.
• Infection prevention professionals may benefit from deeper contextual understanding of clinical workflows.
• Implementation activities such as measurements can be difficult to execute in practice and require pragmatic solutions.
• Facilitating interaction and mutual understanding between stakeholders is crucial for effective implementation.

Overall, the pilot highlighted that implementation is not only a technical or procedural challenge, but also a social and contextual one, requiring tools that support interaction, understanding, and adaptation.

Appendix D

VHIG Congress

To gather feedback from the primary stakeholder group and validate early concept directions, a presentation was given at the VHIG congress on April 16th, together with Miriam. The audience consisted of approximately 76 infection prevention professionals, representing a highly relevant user group for the STIP toolkit.

The purpose of this session was twofold:
• to present the current state of the STIP project and the developed concept,
• to collect structured and unstructured feedback from professionals in the field.

The session included a presentation of the research background (by Miriam) and the introduction of the concept (by me), followed by an interactive Mentimeter and an open discussion.

Mentimeter and user preferences

During the presentation, an interactive Mentimeter was used to collect input from the audience. The questions focused on:

• preferred ways of working on implementation projects (individual vs. collaborative),
• preferred formats for support tools (digital, physical, or hybrid),
• the importance of efficiency versus thoroughness in implementation,
• and preferred ways of collecting input from physicians.

The results provided quantitative insights into user preferences and highlighted variation within the stakeholder group. While clear tendencies were visible, the responses confirmed that there is no single preferred approach that fits all users.

Stakeholder reactions to the concept

The discussion following the presentation revealed diverse perspectives on the proposed AI-supported concept. Some participants expressed strong interest and saw clear potential in using such a tool to support their work. They indicated that a structured and intelligent system could help in navigating complex implementation challenges.

At the same time, other participants were more critical. Concerns were raised about increased reliance on digital tools, the risk of spending more time behind a computer, and the potential complexity of integrating such a system into daily practice.

This contrast in perspectives highlights differences in attitudes towards digitalisation, as well as variations in experience, preferences, and working styles among professionals.

Implications for concept development

The VHIG session provided valuable validation of the relevance of the problem and the potential of the concept, while also revealing important design challenges.

The diversity in responses confirms that a “one-size-fits-all” solution is unlikely to be effective. Instead, the toolkit should be flexible and adaptable to different user preferences and contexts.

Furthermore, the concerns raised emphasise the importance of designing a tool that is efficient, intuitive, and clearly adds value to existing workflows, rather than increasing workload.

Key lessons for toolkit design

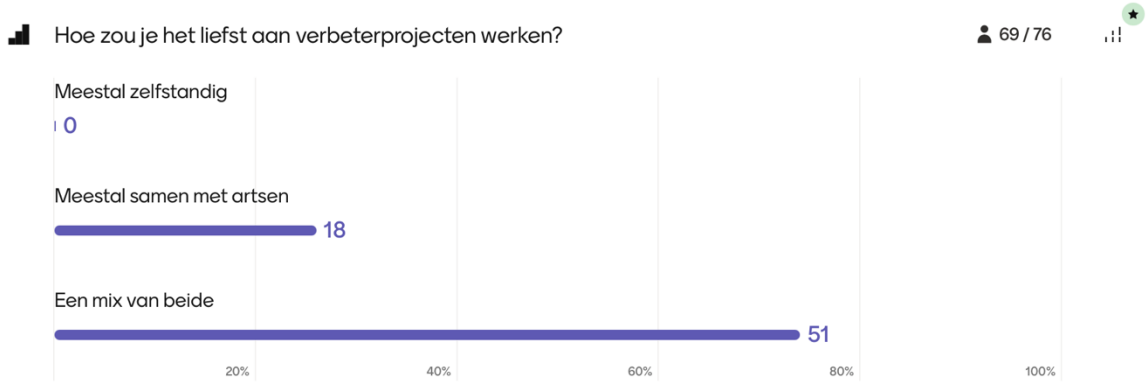
Several insights from the VHIG congress are directly relevant for the development of the STIP toolkit:

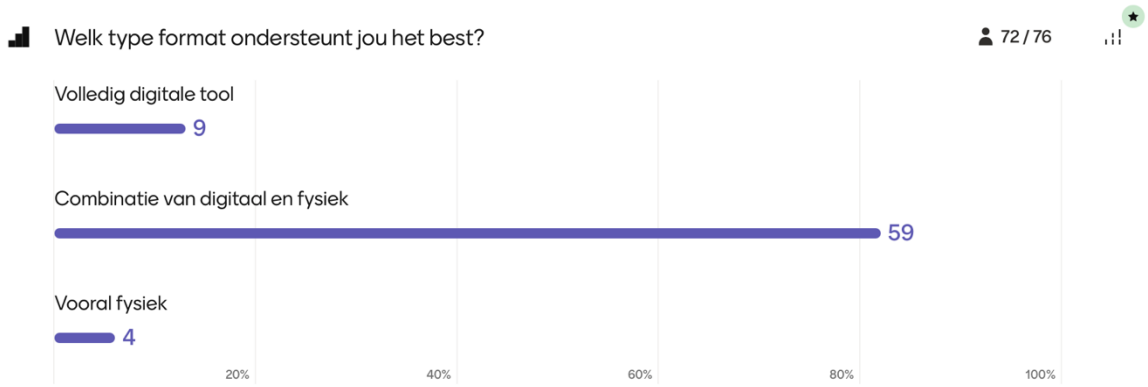
- User preferences vary significantly, requiring flexibility in both format and use.
- There is no universal approach to implementation support; context matters.
- Digital solutions are not universally accepted and should be carefully integrated into workflows.
- Efficiency is a key requirement for adoption in practice.
- Stakeholders are open to innovation, but critical of added complexity.

Overall, the session emphasised the importance of designing a toolkit that balances structure with flexibility and ensures usability in diverse real-world contexts.

Appendix E

Mentimeter VHIG congres





Wat is in verbeterprojecten belangrijker?

71 / 76

Untitled

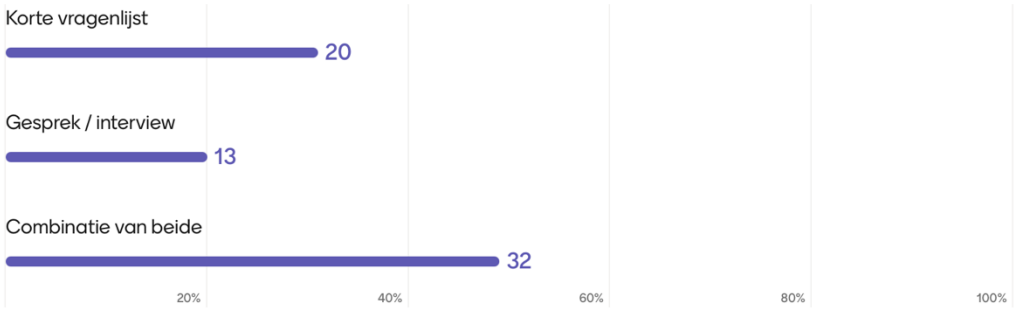
5.1

Snel en efficiënt

Grondig en diepgaand

Hoe verzamel je het liefst input van artsen?

65 / 76



Appendix F List of design requirements

DR	Design Requirement	Main Source
DR 1	Fits a busy clinic	Interview (time pressure); stakeholder (little time of physicians)
DR 2	Adapts to the ward	Interview (department variation); literature (Section 2.1)
DR 3	Guides strategy selection	Interview (data-to-action gap)
DR 4	Structures reflection	Interview (ItFits); literature (Section 2.5)
DR 5	Surfaces determinants	Literature (COM-B, TDF, Ch. 2); interview (data-to-action gap)
DR 6	Matches determinants to strategies	Literature (Theory and Techniques Tool, Ch. 2); interview
DR 7	Makes consequences visible	Interview (invisible risk, gloves); literature (Section 2.3)
DR 8	Works iteratively	Interview (ItFits); literature (Section 2.5)
DR 9	Supports collaboration	Interview (DI as linking pin); stakeholder map (Section 3.3)
DR 10	Needs no training	Interview (ItFits); stakeholder (varied DI backgrounds)
DR 11	Translates theory	Interview (ItFits); stakeholder (basic vs. deeper layer)
DR 12	Is engaging to use	Interview (escape room); stakeholder (DIs like creative work)

DR 13	Tailors physician engagement	Interview (hierarchy, autonomy); literature (Sections 2.2, 2.6)

Appendix G Test protocol

Usability Test Protocol

STIP-assist: an AI-supported decision tool for Infection Prevention professionals

Author: Thibauld Demaret

Programme: MSc Strategic Product Design, TU Delft

In collaboration with: Amsterdam UMC

Version: 1.0 | Date: 16 June 2026 |

Contents

Contents *111*

1. Purpose and scope *112*

1.1 Research questions 112

1.2 Out of scope 112

2. Method and rationale *112*

3. Participants and sampling *113*

4. Test tasks *114*

5. Outcome measures *115*

6. Procedure *115*

6.1 Facilitator conduct 116

7. Data analysis *116*

7.1 Qualitative analysis 116

7.2 Quantitative analysis 116

7.3 Triangulation 116

8. Ethical and data-protection considerations *116*

9. Anticipated limitations *117*

Appendix A. Session checklist (facilitator) *118*

Appendix B. References *118*

1. Purpose and scope

This protocol describes a formal usability evaluation of the STIP-assist prototype. The prototype supports Infection Prevention professionals (Deskundigen Infectiepreventie, hereafter DIs) in analysing ward conversations with physicians, surfacing COM-B behavioural factors, and matching those factors to evidence-based strategies from the STIP toolkit.

Desirability and viability were addressed in earlier phases of the project, including the klankbordgroep session and the expert pilot. This study therefore concentrates on usability: whether DIs can operate the tool effectively, efficiently, and with acceptable satisfaction across the full six-step workflow.

The protocol is written before data collection. Tasks, success criteria, and outcome measures are defined in advance. This pre-specification limits researcher bias during analysis and supports reproducibility, in line with the methodological standards expected of the graduation report.

1.1 Research questions

1. **RQ1.** Can DIs complete each of the six workflow steps without critical errors or facilitator intervention?
2. **RQ2.** Where in the workflow do usability problems occur, and how severe are they?
3. **RQ3.** How do DIs rate the overall usability of the tool on a standardised instrument?
4. **RQ4.** Does the tool support the DI workflow as it is carried out in practice, or does it impose steps that do not fit existing routines?

1.2 Out of scope

This study does not assess clinical effectiveness, the accuracy of the underlying AI analysis against a gold standard, or long-term adoption. Those questions require separate designs and are noted in the roadmap (H8) as later validation milestones.

2. Method and rationale

The study uses a mixed-methods, within-subjects design. Each participant completes the same set of tasks, so each acts as their own reference point. Two complementary data streams are collected.

Stream	Instrument	What it captures
Qualitative	Concurrent think-aloud + structured observation	Where and why problems occur; reasoning, confusion, expectations

Stream	Instrument	What it captures
Quantitative	Task metrics + System Usability Scale (SUS)	Task completion, time, errors, interventions; a comparable overall usability score

The think-aloud method is chosen because it exposes the reasoning behind actions, not only the actions themselves. A participant who hesitates reveals a usability problem that a completion metric alone would miss.

The **System Usability Scale** (Brooke, 1996) is used as the standardised post-task measure. It is a ten-item questionnaire that yields a single score from 0 to 100. It is selected because it is brief, validated, widely benchmarked, and interpretable even with a small sample. A SUS score can be compared against the established adjective-anchored benchmark (a score near 68 is average; above 80 is good).

A small sample cannot support statistical generalisation. The quantitative measures are therefore used descriptively, to characterise the sample and to triangulate the qualitative findings, not to make inferential claims. Stating this limit explicitly is itself part of the methodological rigour.

3. Participants and sampling

Participants are DIs or DI trainees recruited through Amsterdam UMC contacts. Purposive sampling is used: the target population is the intended end-user, so participants are selected for their role rather than at random.

Parameter	Specification
Sample size	5 to 8 participants. Five is the established threshold for surfacing the majority of usability problems in a single round (Nielsen & Landauer, 1993).
Inclusion	Working as a DI or in DI training; involved in ward conversations with physicians.
Exclusion	Prior hands-on use of the STIP-assist prototype (to avoid learning effects).
Recruitment	Via Miriam Schutte and the Amsterdam UMC IPC department.

If recruitment yields fewer than five participants, the study is reported as an exploratory qualitative evaluation and the framing of the findings is adjusted accordingly.

4. Test tasks

Participants work through the full six-step workflow using a prepared, realistic scenario. Completing the workflow end to end gives ecological validity: the steps are experienced in the order and combination intended for practice. Each task has a pre-defined success criterion. The facilitator records the outcome against that criterion.

Tasks are written as goals, not as click-by-click instructions. This forces participants to find their own route and exposes navigation problems. The scenario and any sample conversation are prepared in advance and kept identical across participants.

#	Workflow step	Task goal	Success criterion
1	**Voorbereiding**	Set up a new case for a named ward and physician, ready to record a conversation.	A case is created with the correct ward and context fields completed.
2	**Gespreksmodule**	Capture or import a ward conversation and confirm the transcript is usable.	A transcript is produced and the participant confirms it reflects the conversation.
3	**Factoren-Analyse**	Review the surfaced COM-B factors and judge whether they match the conversation.	The participant locates the factors, interprets them, and can state agree or disagree with reasoning.
4	**Strategieën**	Inspect the recommended strategies and select one to take forward.	The participant understands why a strategy is recommended and selects one deliberately.
5	**Operationalisatie**	Turn the chosen strategy into a concrete action for the ward.	A specific, actionable plan is produced from the selected strategy.
6	**Do-fase**	Locate the output and explain how it would be used after the session.	The participant finds the result and describes its use in their own words.

Observation focus: steps flagged as friction points in the earlier pilot are observed in greater depth. Insert the specific steps and observations from the 15 June pilot here once analysed.

5. Outcome measures

Measures are defined in advance so that scoring is consistent across participants and sessions.

Measure	Definition and scoring
Task completion	Coded per task: completed unaided / completed with a hint / not completed. Only unaided completion counts as full success.
Time on task	Seconds from task start to task end. Recorded as a descriptive indicator of effort, not a pass or fail.
Errors	Count of actions that move away from the goal (wrong screen, wrong control, misread output).
Facilitator interventions	Count of times the facilitator must prompt to let the participant continue.
Problem severity	Each observed problem rated on a four-point scale: 0 cosmetic, 1 minor, 2 major, 3 catastrophic (Nielsen severity rating).
SUS score	Ten-item questionnaire completed after all tasks. Scored 0 to 100 per the standard SUS scoring procedure.

6. Procedure

Each session lasts approximately 45 to 60 minutes and follows a fixed sequence so that conditions are comparable.

1. Welcome and briefing (5 min). Explain the purpose, that the tool is being tested and not the participant, and that there are no wrong answers.
2. Informed consent (5 min). Obtain written consent for participation and recording. Confirm the right to stop at any time.
3. Think-aloud warm-up (3 min). A short neutral task to practise speaking thoughts aloud.
4. Task block (25 to 30 min). Participant works through tasks 1 to 6 while thinking aloud. The facilitator observes and records measures without guiding, intervening only when the participant is fully blocked.
5. SUS questionnaire (5 min). Completed independently, immediately after the tasks.
6. Semi-structured debrief (10 min). Open questions on overall impression, hardest and easiest steps, and fit with daily practice.

6.1 Facilitator conduct

The facilitator follows a script for the briefing and debrief to keep wording constant. During tasks the facilitator stays neutral. Standard non-leading prompts are: “What are you trying to do now?”, “What did you expect to happen?”, and “What would you do next?”. The facilitator does not explain the interface or confirm whether an action was correct.

7. Data analysis

Sessions are recorded (screen and audio) and the relevant segments transcribed. Analysis combines the two data streams.

7.1 Qualitative analysis

Think-aloud and observation data are analysed with a deductive thematic analysis. A coding frame is defined in advance, based on Nielsen’s usability heuristics and the six workflow steps, so that problems can be located and compared systematically. New codes are added inductively if data fall outside the frame. Each identified problem is logged with its step, a description, the number of participants affected, and its severity rating.

7.2 Quantitative analysis

Task metrics and SUS scores are summarised descriptively (counts, medians, range). The mean SUS score is compared against the published benchmark for interpretation only. With this sample size no inferential statistics are reported.

7.3 Triangulation

The two streams are read together. A low SUS score is interpreted through the observed problems behind it. A step with many interventions is cross-checked against the think-aloud reasoning for that step. Convergent findings across streams are treated as the most robust.

8. Ethical and data-protection considerations

- Informed consent is obtained in writing before each session, covering participation, recording, and data use.
- Participation is voluntary and can be stopped at any moment without consequence.
- Recordings and transcripts are pseudonymised. Participants are referred to by code (P1, P2, and so on).
- No real patient data are used. The test scenario and any sample conversation are fictional or fully anonymised.
- Data are stored and processed in line with the GDPR and Amsterdam UMC data-protection requirements. Storage location and retention follow the project privacy impact assessment.

- Ethical approval or exemption is confirmed with the relevant TU Delft and Amsterdam UMC bodies before recruitment.

9. Anticipated limitations

The small purposive sample supports problem discovery, not generalisation. The test uses a prepared scenario rather than a live ward conversation, which controls conditions but reduces realism. A think-aloud protocol can itself alter behaviour, since speaking thoughts aloud may slow or change how a participant works. These limitations are stated in the report and the claims are scoped to match them.

Appendix A. Session checklist (facilitator)

Before the session

- Prototype running and reset to a clean state.
- Test scenario and sample conversation loaded.
- Recording set up and tested.
- Consent form and SUS questionnaire printed or ready.
- Observation sheet with the six tasks and measure columns ready.

During the session

- Read the briefing script verbatim.
- Confirm consent and start recording.
- Run the warm-up, then tasks 1 to 6.
- Record completion, time, errors, and interventions per task.
- Administer the SUS, then the debrief.

After the session

- Stop and save the recording under the participant code.
- Note immediate impressions while fresh.
- Reset the prototype for the next participant.

Appendix B. References

Brooke, J. (1996). SUS: A quick and dirty usability scale. In Usability Evaluation in Industry. Taylor & Francis.

Nielsen, J., & Landauer, T. K. (1993). A mathematical model of the finding of usability problems. Proceedings of INTERCHI ’93.

Nielsen, J. (1994). Heuristic evaluation. In Usability Inspection Methods. John Wiley & Sons.

Appendix H Strategic roadmap

FROM PROTOTYPE TO WORKING IMPLEMENTATION PLATFORM

STIP-Assist Roadmap

A tool that helps the infection-prevention expert (DI) analyse a conversation with a physician, extract behavioural factors via the COM-B model, and match them to st
from the STIP-toolkit, turning each consult into an evidence-based implementation plan.

TRACKS ↓ / HORIZONS →

Horizon 0 · WE ARE HERE

Starting point · june 2026

Laying the foundation

Horizon 1 · PILOTS

In myDRE · starting now

Validation in practice

Pilots start now and run inside myDRE at Amsterdam UMC — Horizon 0 & 1 share the same secure environment.

Gate to advance

Exit criteria before the next horizon. Pilots: Apr 2026 – Apr 2028, 3 hospitals, 4 departments, 6 DIs.

GATE → H1 — START PILOTS

- ✓ End-to-end workflow validated and running reliably within myDRE
- ✓ Infection prevention professionals recognise the identified barriers and suggested strategies as relevant.
- ✓ At least one department enrolled; secure access configured

GATE → H2

- ✓ Pilot studies completed with a positive qualitative evalua
- ✓ Infection prevention professionals use the tool confiden
- ✓ Users consider the tool valuable and relevant
- ✓ SUS ≥ 68 (supporting evidence)

Functionality

How the tool supports the DI

Complete guided workflow

- 6-step workflow: preparation → conversation / questionnaire → analysis → strategies → operationalisation → execution — DI
- COM-B analysis matched to the 29 STIP-toolkit strategies — DI
- Multiple physicians per department, merged into a department view — DI

Sharper insight per department

- Continue previous implementation plans; Export analyses for resea — DEV
- Department profile overview (% per profile) + top strategies per profile — DEV
- Collect physician input through digital questionnaires — DEV · ⚠ Depends on privacy approval and external integration

Platform & AI

How & where the tool runs

Runs securely inside myDRE

THE PIVOT · 2026

~~Planned: a local prototype calling the Nebul API~~

↳ **Now rebuilt to run on myDRE at Amsterdam UMC and ready for the pilots**

- React / TypeScript web app on a protected myDRE VM; transcription (Whisper) + analysis (GPT) via Azure OpenAI inside myDRE — DEV
- All processing takes place within the secure hospital environment, allowing sensitive conversations to be analysed without leaving the organisation. — DEV

Refined within myDRE during the pilot studies

- The myDRE platform is refined based on pilot feedback, impro reliability, and workflow support. — DEV
- **AI-model governance:** AI outputs are validated against the STIP stra model and prompt versions are managed to ensure consistent and reli — DEV
- Barrier identification and strategy matching improved through real-wor — DEV

Data & storage

Where the data lives & how it scales

Data survives VM reset

- Progress written to a JSON file that is retained between sessions; one protected VM per DI — DEV
- **Known limit:** flat JSON, single-writer. **Prototype storage suitable for a single user, but not yet designed for collaboration or long-term scalability.** ⚠ SCALING RISK

From prototype storage to structured database

- Migrate from prototype storage to a structured database that supports maintenance. — DEV
- Backup and recovery procedures, with support for multiple users wor time. — DEV · HOSP

Security & compliance

Data security and privacy

Contained by design

- All data remains within the secure myDRE environment, with an isolated workspace for each infection prevention professional. — DEV
- **Open issue:** Formal management of API credentials — ⚠ CRITICAL

Formal compliance for pilots

- DPIA and AVG assessment completed with the DPO before the pilot — STIP · HOSP
- Only authorised users have access, all activity is logged, and sensitiv managed securely. — DEV

Adoption & sustainability

How it spreads & stays alive

Ownership handed over

- Prepare transfer to the STIP team, including source code, technical documentation and deployment instructions. — DEV · STIP

Validated in the pilots

- Pilot studies completed with infection prevention professionals; structu collected to improve the tool. — DI · STIP
- Maintenance process established, with developer support and a de — STIP

OWNER ● DI · infection-prevention expert ● DEV · developer / UMC tech ● STIP team ● HOSP · adopting hospital | ⚑ Gate / exit criteria ⚠ Risk / dependency

strategies

GOAL

Every infection-prevention department in the Netherlands works with STIP-Assist.

VISION

Implementation of tailored strategies for infection prevention becomes standard practice.

H0 + H1

SCALE OUT

luation
ently in daily practice

search and reporting

file group

proving usability,

trategy catalogue, while
reliable results.

world use.

rts long-term growth and

working safely at the same

lot studies.

tive credentials are

ctured feedback

dedicated STIP contact

Horizon 2

Mid-term · indicative 2028

Scaling for wider adoption

⚑ GATE → H3 — NATIONAL

- ✓ **STIP-assist in active use in at least three hospitals**
- ✓ **Reliable data migration and sustainable cost-per-consult**
- ✓ **Long-term maintenance organized by the STIP team**

Embedded in daily practice

Personal user accounts with secure hospital login (SSO where available).
● DEV

Integrated into infection prevention education and training.
● STIP

Used across multiple hospitals and clinical departments.
● HOSP

Ready for use beyond Amsterdam UMC

Platform architecture supports deployment in different hospital environments and future AI models.
● DEV

Hospitals can deploy and manage their own secure instance of STIP-assist.
● HOSP

Sustainable operational costs supported through a hospital subscription model.
● STIP ⚠ ADOPTION RISK

Secure data environment per hospital

Data is organised per hospital, with separate environments for each organisation.
● HOSP

Existing analyses can be transferred safely to future versions of the platform.
● DEV

Governed across hospitals

Secure access and governance established for multiple hospitals, with independent security testing.
● STIP

Adopted nationally, openly

Available to hospitals across the Netherlands, independent of Amsterdam UMC infrastructure. ● STIP

A hospital subscription supports long-term maintenance, user support, and future
● STIP

Horizon 3

Long-term · indicative 2030 +

National adoption & long-term sustainability

★ END STATE — SELF-SUSTAINING

- ✓ **Used by every IPC department in the Netherlands**
- ✓ **Maintained without dependency on Amsterdam UMC**
- ✓ **Tailored strategies for infection prevention is routine practice**

Routine standard of care

Used as a routine decision-support tool in infection prevention.
● HOSP

Continuously updated to reflect new infection-prevention guidelines and AI developments.
● HOSP

Integrated into hospital IT

A stand-alone programme each hospital installs and runs on its own infrastructure, obtained through a central access point (in the spirit of ItFits). No direct EPR integration is required. Updates distributed by STIP; patient data and the conversation module stay inside the hospital. ● HOSP

Hospital-controlled data storage

Because the tool runs locally, data stays inside each hospital's own environment, keeping the privacy guarantees that make the conversation module possible
● HOSP

Continuous assurance

Continuous compliance monitoring and regular validation of AI outputs.
● HOSP

Self-sustaining standard

Used by infection prevention departments across the Netherlands, with hospitals managing and maintaining their own local installations.
● HOSP

Hospital subscriptions provide sustainable funding, removing dependency on project or grant funding.
● STIP