Exploring habit and context stability in the maintenance of adherence to medication for people living with long-term conditions

Louisa Robinson

School of Health and Related Research (ScHARR)

Faculty of Medicine, Dentistry & Health (MDH)

The University of Sheffield

A thesis submitted for the degree of Doctor of Philosophy

March 2023

Acknowledgements

Abstract

Adherence to prescribed medications is a global challenge, with average adherence to long-term medications estimated at 50%. The active ingredients of successful interventions to improve medication adherence remain elusive. Behavioural science research has identified habit as a mechanism of interest for the maintenance of health behaviours, including medication adherence. A habit is defined as a mental process which is automatically elicited by a stable contextual cue. Existing methods to measure habits are methodologically limited: many require participants to self-report on implicit mental associations and are subject to recall bias; others require time- and resource-intensive procedures to capture data and most fail to capture contextual stability of the behaviour. This is important because elicitation of behaviour in response to a stable contextual cue is a key feature in distinguishing habits from other forms of automatic behaviour.

The overarching aim of this thesis is to explore the role, conceptualisation and measurement of habit and context stability in adherence to medication for people with long-term conditions.

The thesis presents a series of critiques and research studies which explore habit, habitual behaviour, context stability and methods of measuring context stability in people taking medications for long-term conditions. A machine learning assisted review found that habit has not been uniformly conceptualised in line with contemporary habit theory in the medication adherence intervention literature, that behaviour change techniques used to form habits are multiple and heterogeneous and that habit formation is seldom measured. The review also highlighted that medical devices that collect time- and date-stamped adherence data are widely available in medication adherence research and this data could be used to objectively measure habit using the 'behaviour frequency x context stability' approach. This measure has been in use since 2005 but in self-report form. A subsequent critique of the 'behaviour frequency x context stability' measure of habit describes the need to develop an objective metric of context-stability which is demonstrably delineated from frequency of behaviour. The following research studies developed and evaluated the utility of an objective metric of context stability to meet this need. This was conducted using a sample of 607 people with Cystic Fibrosis whose data described the time of day at which they took prescribed nebulised medications over a 12-month follow-up period. This data was captured on electronic medical devices. The chosen metric of context stability was derived from a combination of simple summary statistics and non-linear methods which capture complex patterns in longitudinal data. Albeit with small effect sizes, the metric explained statistically significant variance in ongoing (change in adjusted R^2 =10.3%, p<.001, after controlling for other variables; a 10% increase in

context stability predicted a 7.3% increase in concurrent behaviour), future (change in adjusted R^2=8.2%, p<.001 after controlling for other variables; a 10% increase in context stability predicted a 7.8% increase in future behaviour) and the degree of maintained behaviour (people with higher context stability had 5% (p<.001) less probability of failing to maintain medication adherence). These findings are consistent with performance of self-report metrics of habit and context stability used elsewhere in the literature.

Future work can benefit from the review's explicit clarification of common misconceptions about what a habit is and how it can be targeted, as well as theoretical critiques of common practice in habit research. The remaining chapters provide a starting point for future research to continue the development of the objective metric of context stability by examining the scope of its generalisability in other samples, population groups, study designs and behaviours.

Table of Contents

List of Tables

List of Figures

Abbreviations

ACtiF	Development and evaluation of an intervention to support Adherence to treatment in adults with Cystic Fibrosis
BCT	Behaviour Change Technique
BFCS	Behaviour Frequency x Context Stability
CF	Cystic fibrosis
CFHH	CFHealthHub
CFTR	Cystic Fibrosis Transmembrane conductance Regulator
CI	Confidence interval
COM-B	Capability Opportunity Motivation- Behaviour
DET	Determinism
DIFF_AVG	average of the differences in time of day between consecutive inhalations
DIFF_ENT	entropy in the differences in time of day between consecutive inhalations
DIFF_SD	standard deviation of the differences in time of day between consecutive inhalation
DTW	Dynamic time warping
ENT	Entropy
FEV$_1$%	Forced expiratory volume in first second (per cent)
IQR	Inter-quartile range
IV	Intravenous antibiotics
LAM	Laminarity
LASSO	Least Absolute Shrinkage and Selection Operator algorithm
LTC	Long-term condition
LQ	Lower quartile
MeSH	Medical Subject Heading
PICOTS	Population, Intervention type, Comparison, Outcome, Timeline and Study

RAD	Recurrence radius
RCT	Randomised controlled trial
RoB	Risk of Bias
RP	Recurrence plot
RQA	Recurrence Quantification Analysis
RR	Recurrence Rate
SCT	Social Cognitive Theory
SD	Standard deviation
SE	Standard error
SRBAI	Self-Reported Behavioural Automaticity Index
SRHI	Self-Reported Habit Index
TiDIER	Template for intervention description and replication
TS1	Time series 1
TS2	Time series 2
TT	Trapping time
UQ	Upper quartile
VENT	Entropy in the vertical lines
VMAX	Longest vertical line length
W	Average white vertical line length
WENT	Entropy in the white vertical line length
WHO	World Health Organisation
WP	Work-package

Chapter 1. Introduction to medication adherence, habit theory and issues with the measurement of habits

This thesis converges knowledge, challenges and opportunities from three key areas of health and behavioural research; medication adherence in long-term conditions, habit theory and the objective measurement of psychological constructs. Chapter 1 reviews the literature, highlighting key challenges and opportunities arising from these three areas before drawing them together in the rationale for this thesis; that is, the exploration of habit, habitual behaviour and context stability in maintenance of adherence to medication for people living with long-term conditions.

1.1 Medication adherence

Medication adherence is a complex, evolving and context-variable construct (Stirratt et al., 2018; Vrijens et al., 2012). It is most popularly described as the extent to which a person's behaviour follows that which is prescribed by, and agreed with, their healthcare provider (Osterberg & Blaschke, 2005; World Health Organisation, 2003). Departing from the prior established concept of 'compliance' which implied little or no patient input in the specification of medication regimen, both the degree of inappropriate drug use and the need for consumers of care to be included in the discussion about their prescription are central to the concept of adherence (Vrijens et al., 2012). Importantly, this term promotes an understanding of the significance of patient self-management and engagement in care (Dimatteo et al., 1994).

1.2 Incidence of non-adherence in long-term conditions

Adherence to medication has been studied in relation to both acute and long-term conditions (LTCs; Haynes et al., 2002, 2008; Jackevicius et al., 2002; Osterberg & Blaschke, 2005). The World Health Organization (WHO) describes LTCs as a disease with at least one of the following characteristics: "they are permanent, leave residual disability, are caused by non-reversible pathological alteration, require special training of the patient for rehabilitation, or may be expected to require a long period of supervision, observation or care" (World Health Organisation, 2003, pg. 4).

According to WHO, adherence in LTCs is approximately 50%, meaning one in every two prescribed medications is missed (World Health Organisation, 2003). Estimates across the field echo this figure, frequently reporting non-adherence between 40-60% (Briesacher et al., 2008; Stirratt et al., 2018;

Yeaw et al., 2009). Variation in treatment types and methods of measurement affect estimates of non-adherence (Briesacher et al., 2008; DiMatteo, 2004; Vink et al., 2009; Yeaw et al., 2009), but for people living with LTCs, adherence is consistently reported below the minimally accepted level of 80% (Brown & Bussell, 2011).

Lower adherence is associated with higher healthcare costs (Cutler et al., 2018; Lee et al., 2006; Tan et al., 2011; Tappenden et al., 2017; Vrijens et al., 2017), and has been described as the leading barrier to successful pharmacological therapy in ambulatory patients (Osterberg & Blaschke, 2005). Non-adherence has been consistently associated with poorer health outcomes and mortality in coronary disease (Ho et al., 2008), respiratory disease (Mäkelä et al., 2013; van Boven et al., 2014; Weiner et al., 2008) and diabetes (Asche et al., 2011), among many other long term conditions (World Health Organisation, 2003).

In 2001, WHO launched the Adherence to Long-Term Therapies Project (World Health Organisation, 2001, 2003). This followed the publication of a systematic review of 13 randomised controlled trials (RCTs) of interventions to improve adherence (Haynes et al., 1996). WHO concluded from this review that "increasing the effectiveness of adherence interventions may have a far greater impact on the health of the population than any improvement in specific medical treatments" (World Health Organisation, 2003, p.2). Since this statement, a significant amount of research has been published that is dedicated to developing interventions to improve adherence.

1.3 Adherence interventions

WHO's (2003) report stipulated the need for patient-tailored, effective interventions, as no intervention strategy at the time could evidence universal effectiveness. Between Haynes and colleagues' (1996) systematic review of interventions to improve medication adherence and their Cochrane updates in 2008 (Haynes et al., 2008), and 2014 (Nieuwlaat et al., 2014), there is little evidence that this objective has been achieved.

Haynes and colleagues' studies were intended to bring together the broad spectrum of studies examining the effectiveness of interventions delivered in RCTs, across an array of disease areas. The corpus of studies detailing the delivery and effectiveness of adherence interventions is vast, exemplified by the inclusion of 182 studies in the most recent publication. Reviewers concluded that the breadth and complexity of existing interventions described in the literature was too heterogeneous to synthesise in a meta-analysis (Nieuwlaat et al., 2014). Elsewhere, a large body of syntheses exist which attempt to understand medication adherence intervention effectiveness within subgroups of intervention types (Conn, Ruppar, Enriquez, et al., 2016; Conn & Ruppar, 2017; Demonceau et al., 2013; Edmondson et al., 2018; Kini & Michael Ho, 2018). The breadth of the

literature and lack of consistency between syntheses in their categorisation of intervention types is perhaps one reason that there remains a lack of consensus as to the necessary 'ingredients' of an effective adherence intervention.

1.4 Measuring adherence: capturing and quantifying adherence

Another source of heterogeneity is an abundance of methods to capture and process adherence data. The tools and information used to calculate an adherence summary measure can be broadly categorised into self-report, pharmacy refill, pill count, physiological and electronic data-capture methods (Williams et al., 2013; Zullig et al., 2017), each with an array of associated measurement tools (Jeffery et al., 2014). Innovations in electronic data capture have led some to describe this as the 'gold standard' (Blaschke et al., 2012) particularly when contrasted with self-report measures of adherence. Historically, self-report measures were most frequently used (Jeffery et al., 2014; Stirratt et al., 2015), but are often unreliable and overestimate adherence (Bangsberg et al., 2001; Daniels et al., 2011). In a recent review, high concordance between self-report and electronic measures of adherence was evident in only 17% of comparisons (Garber et al., 2004). Electronic devices pose the potential to overcome recall and social desirability biases associated with self-report measures and offer a powerful window into the observation of adherence behaviours over time.

Once raw data are collected, it must be combined with the expected or prescribed behaviours to obtain a ratio of performed to expected behaviours to give the adherence estimate. Most frequently, this ratio is obtained at one time-point and operationalised as a binary 'adherence' or 'non-adherence' construct (Jeffery et al., 2014; Vrijens et al., 2012). However, this is one of many ways to summarise this ratio. Vrijens et al., (2012) summarised six, commonly used methods to quantify the degree of prescription implementation: proportion of prescribed doses taken; proportion of days with correct number of doses taken; proportion of doses taken on time; distribution of inter-dose intervals; number of drug holidays; longest interval between doses. Each of these could be further diversified by operationalisation as continuous or binary estimates.

1.5 Approaches to the problem of adherence: identifying targets for intervention

Two major schools of thought have dominated perspectives on adherence to medication, largely falling under the umbrellas of the biomedical perspective and the psychosocial and behaviour change perspective.

The biomedical perspective assumes a mechanistic approach in which patients are prescribed treatment and are passive in the management of their illness (Munro et al., 2007; World Health

Organisation, 2003). Patient motivation, cognitions and appraisal of perceived necessity and concerns about taking prescribed medications are all largely disregarded from this perspective.

The behaviour change perspective also focuses on the identification of predictors of human behaviour but the key difference is in the application of theories which can explain human behaviour in terms of individual, social, environmental and contextual features of any given situation (Hagger et al., 2020; Kwasnicka et al., 2016). The benefit of adopting a behaviour change perspective on health-related challenges such as medication adherence is the specification of *modifiable* aspects of the social, psychological and physical contexts which determine human behaviour.

Historically, theories of behaviour change have largely taken the perspective that beliefs, attitudes and intentions are the proximal determinants of behaviour. This view assumes that behaviour is guided by conscious and deliberate decisions to act. Theories of reasoned behaviour including Theory of Planned Behaviour (Ajzen, 1991, 2002), Health Belief Model (Rosenstock et al., 1988) and Social Cognitive Theory (Bandura, 1998) have been popular over the past half-century, and applied to a range of health behaviours (Armitage & Conner, 2000, 2001; Conner & Norman, 2005; McEachan et al., 2011), including medication adherence (Amico et al., 2018; Conn, Enriquez, et al., 2016). Common to these models is the preposition that once an intention is formed, behaviour will manifest and be repeated in the future.

Reflective models have latterly been criticised for failing to address the difference between factors determining behaviour initiation, and behaviour maintenance (Kwasnicka et al., 2016; Rothman, 2000; Rothman et al., 2009). This is important to behaviour change research as typically, change must be lasting to observe the benefits of behaviour. Acknowledgement of the determinants of behaviour *maintenance* is therefore critical in intervention development.

One of the key constructs identified as a mechanism of maintenance is habit (Kwasnicka et al., 2016; Rothman et al., 2009). Habits have been hypothesised to facilitate maintenance of behaviour by bypassing reflective decision-making processes when self-control is diminished, enabling people to engage in behaviours as they have in the past, with minimal cognitive expense (Gardner et al., 2020). This prospect has drawn significant research attention to habits in recent years. Increasingly, researchers have sought to extend existing models of behaviour to include habit as a construct (e.g. see Phillips et al., 2016), develop frameworks to understand the formation of habit and its determinants (Gardner et al., 2020; Gardner & Lally, 2018) and integrate habit formation into behaviour change interventions (Gardner & Rebar, 2019).

1.6 What is a habit?

Definitions of habit converge on two core features: habits are (i) elicited *automatically* by a learned cue-behaviour association and are (ii) elicited in response to a *stable contextual cue* (Gardner, 2015; Mazar & Wood, 2019; Orbell & Verplanken, 2020; Wood & Rünger, 2016). Habits are acquired slowly over time, through repetition of the behaviour in a stable context (Lally et al., 2010; Mazar & Wood, 2019), and this process facilitates the formation of cue-behaviour associations in memory (Wood & Rünger, 2016).

Contemporary definitions distinguish habit from habitual behaviour: habit is defined as a process which can explain mechanisms of behaviour within models of health behaviour; habitual behaviour is the outcome of that process (Gardner, 2015; Verplanken, 2006).

1.7 Defining features of habit: automaticity and context stability

1.7.1 Automaticity

Early behaviourist work paved the foundation for theory which identifies automaticity as a fundamental feature of habit. With the use of devaluation studies in animal models, it was demonstrated that behaviours could be performed efficiently and persist despite devaluation of rewards, after an extensive training period in which cue-response associations were formed in reward-based contingencies (Adams & Dickinson, 1981; Dickinson et al., 1983). Evidence from devaluation studies gave rise to the conceptualisation of habits as a form of goal-independent automaticity because behaviour could persist despite a change in value of the original goal.

Reward devaluation studies in people have also successfully demonstrated persistence of behaviour when reward-outcome contingencies are withdrawn (Tricomi et al., 2009). For example, Neal et al., (2011) gave people either stale or fresh popcorn in either a cinema or conference room. Despite expressing dislike for the stale popcorn, habitual popcorn eaters ate the same amount of stale and fresh popcorn when contextually cued by the cinema setting, but this effect was not observed in non-habitual popcorn eaters.

Tests of implicit attentional biases and cognitive accessibility in response to habit cues have demonstrated that habitual behaviours have stronger, automatic associations with cues in memory (Danner et al., 2008; Neal et al., 2012; Orbell & Verplanken, 2010; Verplanken & Orbell, 2003). Habits have been shown to facilitate the cognitive accessibility of behaviours in response to cues as well as biasing attention toward those cues (Wood & Rünger, 2016).

Goal-independence of behaviour has also been evidenced in the measurement of 'action slips' (e.g. Orbell and Verplanken, 2010, Study 2) and in the moderated effect of intentions on behaviour for different levels of habit strength (Gardner, 2015; Gardner et al., 2011; Rebar et al., 2016).

Historically, the intention-behaviour interaction hypothesis was thought to show that the ability of intentions to predict behaviour was dependent on habit strength, such that intentions do not predict behaviour in people with strong habits. The moderation effect has been recently clarified with evidence to suggest that direction and strength of intentions and habits as well as self-control must all be considered to accurately interpret the findings of studies examining this hypothesis (Gardner et al., 2020); specifically, when self-control is diminished, habits facilitate behaviour regardless of intention strength or direction and when self-control is not diminished, habits can enable enactment of aligned but weakened intentions. These two properties of habits are focal to the rationale for pursuit of habit formation in health behaviour change interventions; it proposes that even on days on which people are stressed or tired (i.e., with low self-control), habits that are aligned with favourable intentions can enable behaviour to be enacted.

1.7.2 Context stability of habits

Habits are formed through repetition of behaviour in response to a contextual cue; over time the association between the cue and behavioural response is formed in memory (Gardner et al., 2020; Gardner & Lally, 2018; Lally et al., 2010; Orbell & Verplanken, 2020).

This has been evidenced in empirical studies: Danner et al., (2008, Study 3) demonstrated that habitual automaticity, indicated by cognitive accessibility of a habitual behaviour, was dependent on the presence of an associated cue; Neal et al., (2012, Study 1) demonstrated reduced response latencies to contextual cues for habitual runners that was not observed in non-habitual runners.

Some of the most compelling evidence for the importance of contextual cues has been demonstrated in habit discontinuity studies; the habit discontinuity hypothesis states that a change in environment, such that contextual cues are no longer encountered, will lead to discontinuation of habitually elicited behavioural responses. In a study of university students' travel mode choice by Verplanken et al., (2008), people who had recently moved house and indicated concern for the environment were less likely to use the car to get work, compared to those with environmental concerns but had not moved house. Similar results were presented more recently by Thomas et al., (2016) in a larger and more geographically diverse sample and both Wood et al., (2005) and Danner et al., (2008, Studies 1 and 2) have also presented discontinuity study findings for exercise behaviours, reading the newspaper and watching television.

1.8 Challenges in measuring habit

Habit is conceptualised as a mental process; this is distinct from the outcome of this process known as habitual behaviour (Gardner, 2015). It follows that habit itself cannot be directly measured. Historically, measures of 'habit' have mostly focused on the assessment of observable behaviour (e.g. frequency of behaviour or perceived automaticity of that behaviour) and inferring the presence, absence or strength of habit on the basis of the hypothesised effect of habit on the behaviour (e.g. decreased response latency to a cue associated with a habitual behaviour). Available measures of habit largely fall into the categories of self-report scales, implicit association tests and 'behavioural frequency x context stability' measures.

1.8.1 Self-report habit scales

Self-report habit scales constitute the most frequently used measure in the habit literature; 88% (n=119) of studies included in a 2015 review examining habit in relation to behaviour used either the Self-Report Habit Index (SRHI; Verplanken & Orbell, 2003) or subscale, including the Self-Report Behavioural Automaticity Index (SRBAI; Gardner et al., 2012; Gardner, 2015).

The SRHI is a 12-item scale which asks participants to reflect on their behaviour (i.e. each question begins 'Behaviour X is something..'), in terms of perceived automaticity, lack of control, lack of awareness, frequency, mental efficiency, and the extent to which the behaviour is perceived to integrate with the participant's identity. The SRHI reliably predicts future behaviour (Gardner et al., 2011; Verplanken & Orbell, 2003), demonstrates strong internal consistency and appears to represent a single factor structure, indicating that together, the 12 items measure a single construct (Verplanken et al., 2005). However, from a conceptual perspective, the SRHI has been criticised for lack of parsimony (Gardner et al., 2012), with particular concern surrounding the inclusion of items assessing behavioural frequency (e.g. 'Behaviour X is something I do frequently') and identity ('Behaviour X is something that is typically me').

The issue with measuring frequency within the SRHI when predicting future behaviour is that researchers tend to use self-report questions to measure future behaviour, too. This could lead to inflated covariances between the independent and dependent variables because of measurement errors caused by biases in the participant's answers to both past and future behaviour response items, which are then wrongly interpreted as explained variance attributable to causal effects of past on future behaviour (Ajzen, 2002). Furthermore, the specific format in which participants are asked to self-report on the frequency of a past behaviour (e.g. 'Behaviour X is something I do frequently') is problematic. The wording of this question is relative because it does not contextualise actual frequency with expected frequency; to do something 'frequently' to one individual may not

carry the same objective meaning to another. This can lead to between-person variance in interpretation and responses and has been shown to lead respondents to give midpoint answers on response scales when unsure on how to answer this question (Gardner & Tang, 2014).

In considering these issues, Gardner and colleagues later developed the Self-Report Behavioural Automaticity Index (SRBAI; Gardner et al., 2012), a four-item automaticity subscale of the SRHI which, among other excluded items, does not include a question about frequency of past behaviour. In their rationale for the SRBAI, Gardner and colleagues posit that automaticity, alone, reflects the 'active ingredient' of habit (Gardner, 2012). The brevity and simplicity of the SRBAI has likely underpinned its popularity as a metric of habit in the proceeding literature.

However, it remains that both the SRHI and SRBAI are self-report scales and this has been the key cause for concern for a number of commentaries on the measurement of habit. First, the assumption that people can self-report on an inherently automatic mental process has been strongly challenged. Hagger et al., (2015) contend that people are not able to accurately report on the automaticity of habit or the causes of habitual behaviour; habits are mental associations activated automatically and implicitly. Psychological theories of human behaviour have described and evidenced the lack of insight that people have into their own implicit processes; for example Bem (1972) on self-perception theory, wrote:

> Individuals come to know their own attitudes, emotions and other internal states by partially inferring them from observations of their own overt behavior and/or the circumstances in which this behavior occurs. Thus, to the extent that internal cues are weak, ambiguous, or uninterpretable the individual is functionally in the same position as an outside observer which must necessarily rely upon those same cues to infer the individual's inner states. (p. 2)

In their review, Nisbett and Wilson (1977) synthesised the existing literature of introspective accounts of higher order mental processes and drew the conclusion that people demonstrate little evidence of accurate introspection on implicit cognitive processes. This conclusion has been experimentally supported in relation to habits, for example: Neal et al., (2012) study showed that despite cognitive associations only being demonstrated for context-specific cues rather than goal-specific cues, people introspected that their behaviour was goal-driven; Ji and Wood (2007) found that people with stronger habits reported stronger certainty of their intentions to perform habitual action across a range of behaviours; Adriaanse et al., (2018) demonstrated that people systematically confabulate reasons for their non-consciously guided behaviours when probed for explanations.

This feature of human behaviour is particularly problematic for self-report measures of habit when intentions and habit align. In contingencies in which an old habit must be suppressed in favour of a new behaviour- e.g. snacking on unhealthy foods being replaced with healthful diet, and an action slip is observed- e.g. an unhealthy snack is eaten- the distinction between intention and habit are conceptually clear and should be easier to self-report upon (although see Adriaanse et al., 2016, 2018). However, in contingencies in which behaviour and intention are aligned, such as adhering to prescribed medication, the line between intention and habit is more difficult to draw. In addition, both empirical evidence and theoretical commentaries on habit formation suggest that the transition from intentional to automatic behaviour occurs along a continuum, as automaticity is gradually acquired over time (Lally et al., 2010). This theoretical and empirical evidence base poses significant challenges to the premise of being able to self-report on habitual automaticity, at all.

Another challenge for the suitability of self-report habit scales is that they do not ask people to report on their implicit cognitive processes, but to reflect on the 'symptoms' of habit (Orbell & Verplanken, 2015). In exploring how people interact with the specific scale items which probe experiential responses from the SRHI and SRBAI, Gardner and Tang (2014) conducted a think aloud study with 20 participants. A range of questionnaire items relating to the experience of automaticity were found to be incorrectly interpreted by participants. Fundamentally, this undermines the justification used to defend the SRHI and SRBAI as valid measures of the 'experience' of automaticity in habit (Orbell & Verplanken, 2015; Rebar et al., 2018).

1.8.2 Implicit measures of habit

In pursuit of a gold-standard measure of habit, some commentators have suggested using implicit measures to tap the underlying associations which form the basis of habit in memory (Gardner, 2015; Labrecque & Wood, 2015; Mazar & Wood, 2019). To do this, researchers have measured reaction times as a proxy of cue-response accessibility, or attentional bias in Stroop tasks (Mazar & Wood, 2019). Whilst primarily used to demonstrate empirical evidence of automaticity in habits, it has been suggested they could also be used to measure presence, absence or strength of habit in a similar way to self-report measures, whilst circumventing the issues related to introspection of automatic processes (Rebar et al., 2018). Their validity for this purpose and the extent to which they tap the same or different aspects of habit remains largely unresearched and this has caused some to caution surrounding their use as it stands (see Rebar et al., 2018). For researchers that wish to examine the role of habit in health behaviour maintenance, these measures only capture behaviours indicative of automaticity (e.g. reduced response latency to a cue) rather than the health behaviour of interest itself (e.g. attendances at the gym), and thus may not be appropriate for this type of research question.

For the present time at least, their practicality for use is perhaps a more fundamental issue; typically, implicit association tests require controlled experimental conditions, many data points collected for each participant in each session and specially designed equipment. Some authors defend the use and practicality of implicit measures in real-world studies however have cited extremely short studies that required a large time commitment from the participant at each measurement sitting (e.g. Marhe et al., 2013 asked participants to complete Stroop task requiring 99 trials per session, administered four times a day alongside other self-report and implicit measures. This study was conducted over a one-week period). As will be discussed, the trend for studies of behaviour and in particular, behaviour change interventions, is to move toward designs which allow longitudinal data collection and analysis using many within-person assessments to identify and understand idiosyncrasies in behaviour change over time. Requirements to complete lengthy measurement tasks over prolonged periods of time could lead to significant participant burden and loss of data, rendering existing implicit association tests of habit impractical.

1.8.3 Measuring context-stability in habits

Despite ubiquitous agreement that context stability is integral to the definition and mechanism of habits, the aforementioned habit measures do little to measure presence or absence of cues in the performance of behaviour, in order to classify them as habitual.

Neither the SRHI nor the SRBAI explicitly assess the dependence of behaviours on a stable contextual cue and thus are unable to dissociate habits from other forms of automatic behaviour (e.g. see discussion on goal-directed automaticity in Aarts, 2007). Some studies have amended the wording of the SRHI or SRBAI stem to include a contextual cue (e.g. 'Behaviour X in context Y is something …'; Gardner et al., 2012; Orbell & Verplanken, 2015) and explicit comparison of context-specific with context-free versions of the SRBAI have demonstrated enhanced sensitivity for demonstrating habit acquisition and maintenance (Diefenbacher et al., 2022).

However, incorporation of context in these metrics has not historically been the norm. Aside from the previously described issues with self-report measures, the lack of acknowledgement of the importance of contextual cues in discerning presence, absence or strength of habit has stood as a major criticism for these measures (Sniehotta & Presseau, 2012).

More importantly, articulation of the specific cue in these measures may be problematic. It remains unclear what properties of cues are conducive to stronger or more quickly acquired habits. As noted by Phillips (2020) there is a tendency to assume that physical aspects of the environment serve as cues to action in habits however, it has been elsewhere suggested that cues could take the form of moods (Ji & Wood, 2007), social contexts (Wood et al., 2005) or complex but recurring interactions

between internal states and the environment (see Botvinick & Plaut, 2004). Wood (2017) posits the following example to capture this problem: "a breakfast routine might involve getting a cup of coffee and sitting down at the kitchen counter. These, in conjunction with morning bleariness, might cue a behaviour of skimming news reports", (p.392). In this example, any one of the coffee, the fact it is morning, sitting at the kitchen counter, or the feeling of bleariness, or all four features in combination, could conceivably cue the news-reading behaviour.

Lack of clarity surrounding what constitutes a contextual cue has implications for the presentation of cues in empirical studies of habit and for the measurement of habits. With respect to implicit measures, these necessarily require that a cue is used to induce responses under experimental procedures: for example, Neal et al., (2012) presented participants with a lexical decision task that required that runners were primed with a cue which was associated with their usual running location (e.g. "park"), in order to record response latencies following their presentation and similarly for the evidence presented by Diefenbacher et al., (2022). Such experimental designs require pre-identification of contextual cues to elicit the hypothesised effects which can be problematic for researchers as the precise nature of contextual cues which elicit behaviours remains elusive. This is compounded with evidence which suggests people are poor at identifying cues for their own habitual behaviour. For example, participants have retrospectively attributed habitual behaviours such as eating (Adriaanse et al., 2016), or smoking (Shiffman et al., 1997), to negative emotions in the absence of evidence to support this.

1.8.4 The 'Behaviour Frequency x Context Stability' (BFCS) measure

To date, only one habit measure has been explicitly designed to capture the context stability of habitual behaviours. This is known as the 'Behaviour Frequency x Context Stability' (BFCS) measure (Ouellette & Wood, 1998; Wood et al., 2005). Scores are calculated on the BFCS by multiplying the frequency of a behaviour by the degree to which it was performed in the same context (i.e., context stability). The frequency with which behaviour X is performed in context Y is assumed to indicate the likelihood that behaviour X is habitual because high scores on the BFCS indicate frequent performance in stable circumstances. In their analyses, Wood et al., (2005) demonstrated that capturing the stability of the behaviour optimised prediction of behaviour over and above past frequency of behaviour alone.

However, to date, the BFCS scale has been almost universally implemented as a self-report measure. Therefore, the BFCS can be subjected to some of the same criticisms as the SRHI and implicit measures presented above when both context and frequency of past behaviour are self-reported.

1.9 Demand for longitudinal and data-driven approaches to explore habit in maintenance of health behaviour

The preceding critique discussed the range of available and commonly used habit measures. Each was found to have strengths and limitations, but none which can adequately or validly capture all fundamental concepts of habit.

Rebar et al., (2018) have noted that the selection and use of any measure of habit in current research will necessarily depend on the aims of the research question. Gardner (2015) highlighted that the majority of the habit literature had sought to examine correlational relationships between existing habits and behaviour, or study the formation of simple habitual behaviours such as button-pressing (e.g. Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977; Webb et al., 2009). However, with increasing attention on the hypothesis that habits may facilitate maintenance of real-world health behaviour, the utility of habit as a target for behaviour change interventions is gaining increasing research attention (Kwasnicka et al., 2016; Phillips, 2011; Phillips et al., 2016; Rothman et al., 2009). Maintenance of health behaviour is the end goal of many behaviour change interventions, and yet most intervention studies have not been able to demonstrate maintenance of long-term change (Curioni & Lourenço, 2005; Nilsen et al., 2010; Ory et al., 2010; Wiecek et al., 2019). Habit is theoretically well-placed to serve as a mechanism of long-term behaviour change and habit formation research in behaviour change interventions is therefore timely. Measures which capture the key constituent components of habit with theoretical validity, and which are suited to the design of these studies are therefore in demand.

In terms of study design, the use of RCTs to evaluate behaviour change interventions is becoming more frequently challenged (Kwasnicka et al., 2018). Whilst RCTs test the effectiveness of interventions for the 'average' person over a given follow-up period, behaviour change science is turning its attention to the use of repeated assessments of a battery of measures within individuals, commonly known as 'N-of-1' designs (Kwasnicka et al., 2017, 2018; Kwasnicka & Naughton, 2020). This idiographic, person-based approach enables researchers to capture how an intervention functions within individuals, and to look for commonalities and differences in how this may change over time, and in response to events encountered over the period of assessment. Whilst, to date, N-of-1 studies are few, they have been identified as the key methodological design which will enable behavioural science to produce research and evidence which goes beyond intervention design for the average individual, to the development of interventions which are applicable to specific individuals (Kwasnicka & Naughton, 2020; Nielsen et al., 2018).

Increasingly, studies are seeking to examine habit formation using longitudinal, repeated measures of habit within individuals (Fournier et al., 2017; Fournier et al., 2017; Kaushal & Rhodes, 2015; Keller et al., 2021; Lally et al., 2010). In their study, Lally et al., (2010) asked participants to perform a healthful behaviour (eating a piece of fruit/drinking water/ exercising), in the same context, followed by the completion of a subscale of the SRHI, every day for 84 days. The key findings of this study have informed some of the core knowledge underpinning contemporary habit theory. According to a Google Scholar search in 2023, Lally and colleagues' study has been cited over 2,200 times, demonstrating the importance and utility of this design in understanding the nature and parameters of habit formation in development of new health behaviour maintenance.

This case study exemplifies how idiographic designs necessarily require many measurements of the same variables to be conducted over time. Gardner et al., (2012) defended the use of the SRBAI for studies of behaviour prediction and habit formation; it is shorter than the SRHI from which it was derived, and Gardner and colleagues argue it is sufficient to demonstrate the development of behavioural automaticity required for this type of research question. However, even with only 4 items administered , use of self-report scales can incur a large participant burden, particularly when considering it is common for N-of-1 studies to administer a battery of assessments, sometimes multiple times daily (e.g. Kwasnicka et al., 2017; Fisher et al., 2016). As such, unobtrusive measures which can be collected without the need for active participant input could serve to facilitate the implementation of this type of research in the real-world.

1.10 Utilising automatic data-capture to innovate the BFCS approach

Leveraging technological innovation with the BFCS approach may provide a solution to meet this demand. Technologies such as smartphones, wearable devices and adaptation of other devices used in a range of health behaviours are becoming increasingly available for research use. This prospect has important implications for the study of habits; such devices could enable unobtrusive capture of both behaviour frequency and context (e.g., time, date and location of behavioural enactment). Not only does this overcome practical issues in designs requiring intensive data collection, this approach circumvents the afore-discussed issues relating to self-reporting behavioural frequency and context.

In addition, in the interest of behaviour change intervention design, use of a 'frequency in context' approach with automatic data capture offers the opportunity to understand the degree to which context stability of the behaviour is important for maintenance of health behaviours in a way that measurement of the automaticity of behaviour (i.e. SRBAI) alone, cannot address. The degree to which context stability of behaviour over time is predictive of maintenance of behaviour in intervention studies is yet to be explored.

Finally, for habit-based behaviour change interventions, the primary goal is to facilitate maintenance of behaviour. Whilst it has been argued that observation of sustained, contextually stable repetition is not necessarily automatic, and thereby habitual, it is necessarily a form of maintained behaviour. In addition, whilst criticised for not distinguishing pre-requisites of habit formation from behaviours observed following successful habit formation, frameworks describing this process assume a natural transition between these stages (e.g. see Gardner and Lally's 2018 framework, p.211); that is, if a behaviour is consistently repeated in a stable context, habitual cue-response associations will form (Wood & Neal, 2007). From the perspective of intervention development, when the goal of the intervention is to reach consistent and maintained performance of a behaviour, distinction of maintained repetition of behaviour as pre- and post-habit formation is arguably less important than understanding the process by which consistently repeated and maintained behaviour was achieved. As such, the development of objective measures of frequency and context stability could be integral to the future to intervention developers to track progress toward and maintenance of a behavioural goal.

1.11 Interim summary of literature review findings

Medication adherence is consistently low across long-term health conditions (LTCs) and is associated with poor health outcomes and increased healthcare costs. Interventions to improve medication adherence are both abundant and diverse, yet few demonstrate maintained effectiveness. A range of measures have been used to examine medication adherence. With the advent of technology to enable automatic data capture, electronic devices have enabled more accurate and objective measurement of adherence.

Medication adherence is a health behaviour, and in the context of LTCs, this behaviour must often be maintained over the lifetime of an individual. Behaviour change perspectives on mechanisms of maintained behaviour are increasingly turning attention to habits.

Habits are defined as learned associations formed in memory which are automatically triggered in response to contextual cues. Habits are mental processes and are distinct from the behaviour which manifests from the elicitation of a habit. It follows that habits cannot be directly measured. Most existing measures of habit rely on a participant's ability to self-report on the experience of habit or require time-consuming data collection in carefully controlled experimental paradigms; these measures each have methodological and practical issues.

An alternative way think about the problem of measuring habit is to capture its core, observable features. Habitual behaviours are stimulated in response to contextual cues and some researchers measure habit by quantifying the extent to which a behaviour is performed in the same context:

highly frequent behaviours in highly stable contexts indicate that the behaviour is likely elicited by habit. However, measurement of frequency and context has to date relied on participants to self-report how often the behaviour has been performed, and how frequently in the same context. As with other self-report measures of habit, this is liable to recall and self-presentation biases.

In attempts to identify successful intervention components for specific individuals, behavioural science is increasingly favouring evidence from N-of-1 designs over randomised controlled trials. This methodology requires an intensive schedule of longitudinal data collection and participant reported outcome measures can be impractical, particularly for lengthy follow-up periods. There is demand for methods which enable unobtrusive and automatic capture of some variables of interest.

Together, these findings indicate that it is timely to explore the development of new methods to measure habit which capture the context-stability of behaviour without the need for participant recall, introspection of mental processes or undertaking of time-consuming data collection procedures.

Medication adherence is a useful behaviour in which to explore habit using a BFCS approach owing to the availability of electronic medical devices which enable objective and automatic capture of the frequency of doses administered (i.e. information about the frequency of behaviour) and the time and date that they were administered (i.e. information of the context of the behaviour). Reciprocally, exploration of the role of habit using behaviour change theory and insights applied to the problem of medication adherence may offer new insights into limitations of the existing research as well as potential target mechanisms of maintenance for medication adherence intervention.

1.12 Habit and medication adherence

1.12.1 Challenges around conceptualisation of habit in the medication adherence literature

Research examining the relationship between 'habit' and medication adherence is limited but the available evidence suggests that the relationship between habit (or rather habitual behaviour) and medication adherence aligns with that in other health behaviours. A systematic review published in 2020 identified 11 studies measuring the association between 'habit' and adherence, with 91% demonstrating that self-reported habit (i.e. habitual behaviour or perceived automaticity of behaviour) positively correlates with adherence in both cross-sectional and longitudinal designs (Badawy et al., 2020). Among the evidence, Badaway et al., (2020) concluded that 'habit' was more predictive of adherence than other belief-related constructs.

However, one notable disparity is highlighted by this review. Despite the correlational evidence, Badaway and colleagues identified an apparent lack of literature on habit-based medication adherence *interventions*. Similarly, Gardner and Rebar (2019) conducted a systematic review of habit-based interventions, but no medication adherence interventions were included in their review. This is in stark contrast with the findings of two large reviews of medication adherence interventions which have identified 'habit-based' interventions as most effective in moderator analyses of intervention content (Conn, Ruppar, & Chase, 2016; Conn & Ruppar, 2017). The discrepancy in findings between the habit literature as described by behavioural science, and habit in the context of the medication adherence literature perhaps reveals a disparity in the ways in which habit and habitual behaviour have been conceptualised and targeted by intervention designers from these different fields.

This is an important and timely discrepancy to address with the rising profile of habit as a potential mechanism to facilitate maintained medication adherence.

1.12.2 Other challenges for the BFCS approach applied to medication adherence

Whilst devices which automatically capture adherence data are becoming increasingly available, most only tend to capture information on date and time. In terms of capturing contextual cues that elicit habits and habitual behaviour, this could be problematic. A number of studies and commentators have suggested that cues to action may not be times of day per se, but perhaps more salient cues such as the completion of a previous routine (Gardner & Lally, 2018; Judah et al., 2018; Lally & Gardner, 2013; also see evidence from script elicitation studies, e.g. Judah et al., 2018; Mohideen et al., 2023). Expecting participants to show context stability by taking medications at the time of day on, for example, weekdays and weekends implies that the assumed underlying cue is the time of day, rather than events that occur at similar times of day, depending on what day of the week it is. Even allowing for variability by weekday and weekend may not be sensitive enough to identify idiosyncrasies in routine cycles between participants. Consider the following example:

> Participant X habitually takes their medication after brushing their teeth. Putting the toothbrush down cues the initiation of the medication routine. Participant X works Monday to Thursday and so medication taking reliably occurs at around 7:20am, within a 10-minute window either side. On Friday to Sunday participant X does not work and so their morning routine starts later at 8:30 and medication taking is usually around 8:50am, with a window of 10 minutes either side of this.

Arbitrary assignment of an expected cycle from Monday to Friday in this case would falsely indicate that inhalations on Fridays demonstrated contextual inconsistencies, merely because of inadequately pre-specified routine cycles.

Any metric of context stability derived from automatic data-capture of the date and time that behaviours occur may be limited in its utility to only those whose routines do not change in real-time throughout the week if this is not accounted for.

1.13 Summary of challenges

Two clear challenges emerged from the literature review findings: first, it is unclear whether 'habits' have been conceptualised in line with contemporary habit theory in the medication adherence literature; second, habit is a process which cannot be directly measured. Objective measurement of frequency and context collected on electronic medical devices may overcome issues relating to response bias and self-perceived cues for habitually elicited behaviours. However, theory and evidence indicates that cues may likely be related to other daily routines therefore any metric of context stability derived only from the time of day must be able to account for cyclical changes in these routines which repeat over longer time-horizons than a single day.

These findings and identified challenges motivated the development of the thesis aims and objectives, which are presented in Chapter 2.

Chapter 2. Thesis aims, objectives and outline of the remaining chapters

2.1 Thesis aims and objectives

The aim of this thesis is to explore the role, conceptualisation and measurement of habit, habitual behaviour and context stability in adherence to medication for people living with long-term conditions.

2.1.1 Aim 1

To examine the conceptualisation of 'habit' and the implementation and utility of 'habit formation' in interventions to support the maintenance of medication adherence.

2.1.1.1 Objectives

1. To describe how habit has been conceptualised, and whether theory has been applied (i.e., evidence of the requirement for the behaviour to be repeated in the same context to acquire associations between cues and behaviour), in interventions designed to improve medication adherence in long-term conditions;

2. To describe specific behaviour change techniques (BCTs) that have been used to target habit in interventions to improve adherence to medication in long-term conditions;

3. To examine the effectiveness of interventions identified as using context-dependent repetition in improving adherence to medication in people with long-term conditions.

2.1.2 Aim 2

To explore whether the time of day at which people take their medications can be used to derive an objective metric of context stability.

2.1.2.1 Objectives

1. To critique previous implementations of the 'behaviour frequency x context stability' measure to understand how this approach might best be implemented with objective data.

2. To present a systematic approach to identifying and selecting a metric of context stability, derived from objective data on the times of day at which people initiate medication adherence behaviour.

   2.1) To describe a range of methods that could be used to produce objective variables that describe context stability, with theoretical improvements on previous measures;

   2.2) To verify the interpretability of these variables in terms of their theoretical relationship with concurrent medication adherence behaviour;

   2.3) To identify which of these variables explain the most variance in concurrent medication adherence behaviour and use this information to define an objective metric of context stability.

3. To evidence the utility of the defined objective metric of context stability in explaining variance in maintained and future behaviour, and in related constructs in evidence of predictive, convergent and divergent validity.

   3.1) To examine the amount of variance explained in concurrent behaviour by the objective metric of context stability;

   3.2) To examine the amount of variance explained in participants' self-reported habitual automaticity for taking their medication by the objective metric of context stability;

   3.3) To examine the amount of variance explained in perceived effort participants required to take their medication by the objective metric of context stability;

   3.4) To examine whether distinct variance in concurrent behaviour is explained by reflective motivational constructs and the objective metric of context stability;

   3.5) To examine the amount of variance in future adherence behaviour explained by the objective metric of context stability;

   3.6) To examine the moderating effect of the objective metric of context stability in the relationship between past and future behaviour;

3.7) To examine the mediating effect of the objective metric of context stability in the relationship between past and future behaviour.

4. To demonstrate the application of the objective metric of context stability in applied clinical research questions.

2.2 Outline of the following chapters

Chapter 3 addresses Aim 1 with a machine-learning assisted review of the use of habit formation in medication adherence interventions for long-term conditions. The review was published in 2022 (see Robinson et al., 2022) and presented on invitation at the European Health Psychology Society Conference 2022 Habit Special Interest Group meeting in Bratislava, Slovakia.

Chapter 4 addresses Aim 2, Objective 1 with a detailed critique of the numerical properties of Behaviour Frequency x Context Stability measures as implemented in the existing literature. The critique highlights a series of issues with the use of a multiplicative score derived from behavioural frequency and context stability metrics and concludes with specific rationale for the trajectory of the following thesis chapters.

Chapter 5 introduces The ACtiF Programme and randomised controlled trial of the CFHealthHub intervention, from which the dataset used in the empirical analyses throughout the thesis was derived. The participant group was sampled from the adult cystic fibrosis (CF) patient population.

Chapter 6 addresses Aim 2, Objective 2.1 and 2.2. It presents novel methods to derive variables that can objectively describe context stability from temporal medication adherence data. The novel methods are presented alongside other metrics used in the literature and followed by a preliminary exploration of their relationship with adherence behaviour.

Chapter 7 addresses Aim 2, Objective 2.3 and Aim 3, Objectives 3.1 to 3.4. The chapter presents a systematic approach to identifying a metric of context stability with the variables derived in Chapter 6, followed by statistical analyses which test hypotheses relating to construct validity of the objective context stability metric.

Chapter 8 addresses Aim 2, Objectives 3.5 to 3.7. The chapter follows closely from Chapter 7 by presenting a series of analyses that examine the role of the objective metric of context stability in future behaviour.

Chapter 9 addresses thesis Aim 2, Objective 4 by demonstrating the utility of the objective metric of context stability in applied clinical research questions.

Finally, Chapter 10 presents a discussion of the findings with respect to the overarching aim which was to explore the role, conceptualisation and measurement of habit and context stability in adherence to medication for people with long term conditions. The strengths and limitations of the thesis are presented with a series of reflections and recommendations for future research.

The thesis aims, objectives and respective chapters are summarised in Table 1.

Table 1

Summary of thesis aims, objectives and respective chapters.

Aim	Objective	Chapter	Chapter summary
Aim 1: To examine the conceptualisation of 'habit' and the implementation and utility of 'habit formation' in interventions to support the maintenance of medication adherence.	1. To describe how habit has been conceptualised, and whether theory has been applied (i.e., evidence of the requirement for the behaviour to be repeated in the same context to acquire associations between cues and behaviour), in interventions designed to improve medication adherence in long-term conditions 2. To describe specific behaviour change techniques (BCTs) that have been used to target habit in interventions to improve adherence to medication in long-term conditions 3. To examine the effectiveness of interventions identified as using context-dependent repetition in improving adherence to medication in people with long-term conditions	3	A machine-learning assisted review of the use of habit formation in medication adherence interventions for long-term conditions.
Aim 2: To explore whether the time of day at which people take their medications can be used to derive an objective metric of context stability.	1. To critique previous implementations of the 'behaviour frequency x context stability' measure to understand how this approach might best be implemented with objective data.	4	A critique of the numerical properties of Behaviour Frequency x Context Stability measures as implemented in the existing literature.
	2. To present a systematic approach to identifying and selecting a metric of context stability, derived from objective data on the times of day at which people initiate medication adherence behaviour. 2.1) To present a systematic approach to identifying and selecting a metric of context stability, derived from objective data on the times of day at which people initiate medication adherence behaviour. 2.2) To verify the interpretability of these variables in terms of their theoretical relationship with concurrent medication adherence behaviour.	6	Presents novel methods to derive variables that can objectively describe context stability from temporal medication adherence data.
	2.3) To identify which of these variables explain the most variance in concurrent medication adherence behaviour and use this information to define an objective metric of context stability. 3. To evidence the utility of the defined objective metric of context stability in explaining variance in maintained and future behaviour, and in related constructs in evidence of predictive, convergent and divergent validity. 3.1) To examine the amount of variance explained in concurrent behaviour by the objective metric of context stability; 3.2) To examine the amount of variance explained in participants' self-reported habitual automaticity for taking their medication by the objective metric of context stability; 3.3) To examine the amount of variance explained in perceived effort participants required to take their medication by the objective metric of context stability; 3.4) To examine whether distinct variance in concurrent behaviour is explained by reflective motivational constructs and the objective metric of context stability;	7	Presents a systematic approach to identifying a metric of context stability with the variables derived in Chapter 6, followed by statistical analyses which test hypotheses relating to construct validity of the derived objective context stability metric.

Table 1

Continued

Aim	Objective	Chapter	Chapter summary
Aim 2: To explore whether the time of day at which people take their medications can be used to derive an objective metric of context stability.	3.5) To examine the amount of variance in future adherence behaviour explained by the objective metric of context stability; 3.6) To examine the moderating effect of the objective metric of context stability in the relationship between past and future behaviour; 3.7) To examine the mediating effect of the objective metric of context stability in the relationship between past and future behaviour.	8	Presents a series of analyses that examine the role of the derived objective metric of context stability in future behaviour.
	4.To demonstrate the application of the objective metric of context stability in applied clinical research questions.	9	Demonstrates the utility of the objective metric of context stability in applied clinical research questions

Chapter 3: A machine-learning assisted review of the use of habit formation in medication adherence interventions for long-term conditions

3.1 Chapter overview

A key finding of the literature review presented in Chapter 1 was a discrepancy between the ways in which habit has been conceptualised and targeted by intervention designers in the behavioural science habit literature and that in the medication adherence literature. This is an important and timely discrepancy to address with increasing attention on habit as a potential mechanism to facilitate maintained medication adherence. As such, Aim 1 of the thesis was to examine the conceptualisation of 'habit' and the implementation and utility of 'habit formation' in interventions to support the maintenance of medication adherence.

Chapter 3 presents the publication of a machine-learning assisted review that was conducted to address this aim.

3.1.1 Publications associated with Chapter 3

Robinson, L., Arden, M.A., Dawson, S., Walters, S.J., Wildman, M.J. & Stevenson, M. (2022). A machine-learning assisted review of the use of habit formation in medication adherence interventions for long-term conditions, Health Psychology Review, DOI:10.1080/17437199.2022.2034516

3.1.2 Presentations associated with Chapter 3

Robinson, L., & Arden, M.A. (2022). Applying BCT taxonomies to habit formation interventions: Reflections on a review of medication adherence interventions [Presentation]. EHPS 2022 (Habit Special Interest Group Meeting), Bratislava, Slovakia. https://2022.ehps.net/program/

3.2 Abstract

Adherence to medication in long-term conditions is around 50%. The key components of successful interventions to improve medication adherence remain unclear, particularly when examined over prolonged follow-up periods. Behaviour change theories are increasingly interested in the utility of habit formation for the maintenance of health behaviour change, but there is no documentation on how habit has been conceptualised in the medication adherence intervention literature, or what

effect the key technique identified in habit formation theory (context dependent repetition) has in these studies. To examine this, a machine-learning assisted review was conducted. Searches of MEDLINE, EMBASE and PSYCInfo and the reference list of a comprehensive systematic review of medication adherence interventions yielded 5973 articles. Machine learning-assisted title and abstract screening identified 15 independent RCTs published between 1976 and 2021, including 18 intervention comparisons of interest. Key findings indicate that conceptualisations of habit in the medication adherence literature are varied and intervention behaviour change techniques were diverse among these studies. Behaviour change technique coding identified only six studies which explicitly described using habit formation. Conclusions indicate that despite the potential utility of habits as a technique to support maintenance in medication adherence, randomised controlled trials of habit formation interventions are few. Future work should aim to develop this evidence base, drawing on contemporary habit theory and with explicit demonstration of what techniques have been used to promote habit formation.

3.3 Background

Adherence to medication is a cornerstone of medical prescribing in long-term conditions but is frequently estimated to be in the region of 50% (Osterberg & Blaschke, 2005; World Health Organisation, 2003). Adherence to medication describes the degree to which a person takes medication as agreed with a healthcare provider (Osterberg & Blaschke, 2005). There are a wide range of methods to measure and quantify adherence (Lehmann et al., 2013; Vrijens et al., 2012) but evidence from large studies of people with long-term conditions demonstrates that adherence is consistently below the minimally accepted standard of 80% (Briesacher et al., 2008; World Health Organisation, 2003; Yeaw et al., 2009).

Non-adherence incurs huge costs to both economic and health domains. The economic impact of non-adherence was recently estimated to cost between $949 and $44,190 per person, per year (Cutler et al., 2018). The implications of non-adherence on health outcomes are numerous, including worsening of disease and increased hospitalisations and mortality, and this is evidenced across a wide-range of long-term conditions (Asche et al., 2011; Bangsberg et al., 2001; Chowdhury et al., 2013; Mäkelä et al., 2013). There is widespread consensus on the negative impact of non-adherence to medication, and this was acknowledged by the World Health Organisation with the publication of the 'Adherence to Long Term Therapies' reports (World Health Organisation, 2001, 2003).

The development of interventions to support people with a range of chronic health conditions to improve their adherence to medication stands as one of the most important factors in improving

outcomes to prescribed medications (Brown et al., 2016; Saag et al., 2018; World Health Organisation, 2003). Despite this, the 'active ingredients' of effective medication adherence interventions remain unclear (Anderson et al., 2020; Nieuwlaat et al., 2014). In a Cochrane review of medication adherence interventions, authors were prevented from drawing conclusions as to the effectiveness of adherence interventions due to the range of different conceptualisations and measurements of adherence, studied in an array of disease areas, and using a multitude of intervention types (Nieuwlaat et al., 2014). Even within systematic reviews focusing on specific intervention types (e.g. Conn & Ruppar, 2017; Conn, Ruppar, Enriquez, & Cooper, 2016; Demonceau et al., 2013; Edmondson et al., 2018), there appears to be a diverse range of methods to implement seemingly similar intervention types (e.g. digital interventions, reminder-based interventions), with varied outcomes. The study of medication adherence interventions could benefit from the use of standardised taxonomies of intervention types and behaviour change techniques (BCTs; Michie et al., 2013) to examine intervention effectiveness.

In addition to the limited understanding of the 'active ingredients' of medication adherence interventions, even less is known about how to maintain adherence to medication in the long-term. Evidence has shown reduced effectiveness of medication adherence interventions with longer follow-up periods (Wiecek et al., 2019). This effect is not unique to medication adherence; failure to maintain health behaviour change effects in the long-term has been identified as a key research priority (Nilsen et al., 2010; Ory et al., 2010) and is gaining increasing attention in the context of social cognitive models of health behaviour (Norman & Connor, 2015). In a systematic review of over 100 theories describing maintenance of behaviour change, Kwasnicka and colleagues (2016) identified five key themes that theorists place at the core of maintenance theory, including maintenance motives, self-regulation, psychological and physical resources, the physical and social environment, and habit.

Habits have been hypothesised to facilitate maintenance of behaviour by bypassing reflective decision-making processes and enabling people to engage in behaviours as they have in the past, with minimal cognitive expense (Gardner & Lally, 2018; Gardner, Rebar, & Lally, 2020). Whilst the exact definition of habit is widely disputed, contemporary habit theory definitions, including modern, interdisciplinary accounts of habit (see Fleetwood, 2019) converge on the idea that habitual behaviour (e.g. brushing teeth) is initiated by a stable contextual cue (e.g. getting out of the shower) which is associated with that behaviour (Gardner, 2015; Gardner et al., 2021).The acquisition of a habit for a specific behaviour relies on repetition of the behaviour in response to a stable contextual cue, as this facilitates learning of cue-response associations over time (Lally et al., 2010).

Medication adherence can encompass different behaviours which may vary in complexity, for example, swallowing a pill, to preparing several medications to be administered via a medical device (e.g. a nebuliser). Contemporary habit theory places emphasis on the importance of the habitual instigation of the behaviour, whereby a contextual cue can habitually trigger the initiation of the first sub-action, among a potential sequence of sub-actions which constitute a behaviour (Gardner et al., 2016). Through this lens, techniques required to support the habitual instigation of behaviour should theoretically translate across the range of medication adherence behaviours.

Identifying which BCTs are required to target specific 'mechanisms of action' (MoAs) in behaviour change interventions is challenging (Hagger, Moyers, et al., 2020), but there is a growing evidence base to support researchers in the context of habit formation. Recent research has emphasised that context-dependent repetition (equivalent to the Michie et al., 2013 'Habit formation' BCT) is the principal technique by which habits can form (Gardner et al., 2020; Gardner & Lally, 2018) and thereby any intervention reporting to utilise contemporary habit theory in its design should at least demonstrate use of this technique. However, whilst context dependent repetition is minimally required, growing consensus indicates that self-regulatory BCTs (e.g. action planning, prompts and cues, self-monitoring) may support the individual through the period of effortful repetition in the path to habit formation (Gardner et al., 2020; see also Carey et al., 2018; Connell et al., 2018). In a recent review, Gardner & Rebar (2019) searched for studies utilising context-dependent repetition and identified which *other* BCTs were used in combination to facilitate the formation of health behaviour habits. All included interventions combined context-dependent repetition with some other motivation or action control techniques, such as 'Action planning', 'Goal setting (behaviour)', 'Prompts and cues' and 'Problem solving' (see Michie et al, 2013 BCT taxonomy). This is unsurprising as both theoretical models of habit formation (Gardner & Lally, 2018; Gardner et al., 2020) and lay understandings of how habits are acquired (Brown et al., 2019) implicate a period during which the behaviour must be effortfully repeated before it becomes habitual.

Notably however, no studies included in Gardner and Rebar's review examined habit formation in a medication adherence intervention, yet in the largest synthesis of medication adherence interventions to date, Conn & Ruppar (2017) concluded that 'habit formation' interventions were likely to hold the most promise for facilitating changes in medication adherence. This disparity in findings indicates there may be differences in conceptualisations of what a habit is, and ideas about how it can be targeted, between the medication adherence and contemporary habit theory literatures. The extent to which the existing medication adherence intervention literature aligns with recommendations from contemporary habit theory has not yet been documented. An exploration, through the lens of contemporary habit theory, of how habit has been conceptualised, whether

context-dependent repetition was used and which BCTs have been used alongside it could usefully draw attention to discrepancies between these two literatures. Furthermore, examination of the effectiveness of 'context dependent repetition' as a BCT in the context of these medication adherence interventions could provide a useful addition to the growing evidence base for the utility of this technique for supporting long-term behaviour change.

A machine-learning assisted review of randomised controlled trials (RCTs) examining the effectiveness of 'habit' interventions in maintaining adherence was conducted to address these issues.

The specific research objectives of this review were:

 1) To describe how habit has been conceptualised, and whether theory has been applied (i.e. evidence of the requirement for the behaviour to be repeated in the same context to acquire associations between cues and behaviour), in interventions designed to improve medication adherence in long-term conditions;

 2) To describe behaviour change techniques (BCTs) that have been used to target habit in interventions to improve adherence to medication in long-term conditions;

3) To examine the effectiveness of interventions identified as using context-dependent repetition in improving adherence to medication in people with long-term conditions.

3.4 Materials and methods

The review was preregistered on PROSPERO, available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020187890

The scope of the review was extensive and inclusive across Population, Intervention type, Comparison, Outcome, Timeline and Study (PICOTS) domains. The scope covered individual and cluster randomised controlled trials (S) of habit-related adherence interventions (I) compared to usual care or control interventions (C). These had to be conducted in adults with long term conditions (P) and measure effects on adherence(O) at a minimum of 6-months post-randomisation follow-up (T).

3.4.1 Search strategy

The search strategy was implemented in two stages.

3.4.1.1 Stage 1: Identification of studies published prior to 2015

Conn and Ruppar (2017) published a large-scale systematic review of interventions designed to improve medication adherence. The final searches conducted for their review, as reported by Conn and Ruppar , were in 2015. The strategy used was assessed and it was understood that the search conducted by Conn and Ruppar (2017) would be both comprehensive and inclusive enough for the present review. A list of the included studies was retrieved and screened for eligible studies against the inclusion criteria for the present review.

3.4.1.2 Stage 2: Identification of studies published after 2015

Electronic database searches were conducted on 13th May 2020: MEDLINE (1946 to May 12, 2020), PSYCInfo (1806 to May Week 2 2020) and EMBASE (1974 to 2020 May 13) were searched. Key search terms included medication adherence (MeSH), randomised controlled trials, habit and behavioural or complex interventions. The full MEDLINE, PSYCInfo and EMBASE search strategies are available in Appendix A.

Reference lists of included studies were also screened for eligible studies.

3.4.2 Inclusion criteria

Texts reporting RCTs of interventions designed to improve medication adherence in adults living with long-term conditions were the focus of this review. Eligible studies had to use the term 'habit', or 'automaticity' with reference to the SRBAI, in either the introduction, methods or associated documentation describing the intervention in more detail. Studies examining intervention effects in health conditions for which there is no cure and which is managed with a regimen of medication were eligible for this review. This specific criterion, used by the department of health (Department of Health, 2012), was chosen because: i) this definition excludes disease groups such as tuberculosis, which is sometimes considered a long-term condition but can be cured with a 6-month prescription of antibiotics. The present review is concerned with long term (>6month) effects of interventions and therefore inclusion of diseases of this nature would be inappropriate; ii) this definition excludes conditions such as sleep apnoea; whilst sleep apnoea is also sometimes considered to be a long-term condition, it is primarily treated using mechanical devices worn overnight, rather than medicines taken in discrete episodes.

The focus of the review is to examine how the development of habits in the context of medication adherence facilitates long-term maintenance of medication-taking behaviour. This extends beyond

initiation of the behaviour and requires evidence over prolonged periods. In line with the Transtheoretical model's conceptualisation of maintenance (Prochaska & Di Clemente, 1982), studies with a minimum follow-up of 6-months only were eligible for inclusion in this review.

All studies reporting full-scale RCTs published in peer reviewed journals were eligible for inclusion. Any number or type of comparison intervention was acceptable. Studies reporting adherence as a primary or secondary outcome variable were included.

3.4.3 Exclusion criteria

Studies of intervention effects on people who were in prison or people with substance abuse and/or psychological disorders were excluded. This is in line with Conn and Ruppar's (2017) exclusion criteria and avoids inclusion of populations requiring specialist intervention design to support additional complexities in the self-management of disease.

Studies published as conference abstracts and journal supplements were excluded due to insufficient detail to code BCTs. Pilot and feasibility studies and studies not available in English or in full-text were excluded.

3.4.4 Study selection and data extraction

3.4.4.1 Title and abstract screening

Title and abstract screening was conducted using an active learning approach (see O'Mara-Eves, Thomas, McNaught, Miwa, & Ananiadou, 2015), in which a portion of studies were manually labelled (include/exclude) and used to train a machine learning text-classification algorithm. The algorithm outputs the remaining, unlabelled studies in rank order of likelihood of inclusion, which the reviewer then continues to manually label. This process is iterated until the likelihood of identification of additional studies for inclusion is deemed low enough to cease screening (see O'Mara-Eves, Thomas, McNaught, Miwa, & Ananiadou, 2015). Application of machine learning techniques for text classification is gaining increasing popularity in health science reviews (e.g. Currie et al., 2019; Shemilt, Anneliese, et al., 2021; Shemilt, Noel-Storr, et al., 2021) and a detailed description of the methods applied here are available in Appendix B (see also Marshall & Wallace, 2019; O'Mara-Eves et al., 2015).

3.4.4.2 Full-text screening

All abstracts identified in the title and abstract screening process were manually screened. Due to the nature of the interventions of interest, a large number of articles were expected to be eligible for the full-text screening stage. In anticipation of this, a two-stage full-text screening process was planned and implemented.

In the first stage, full-texts were searched for keywords in either the introduction, methods or associated documentation (referenced in the methods) which detailed intervention content. Key words were 'habit(s/ual)' and /or 'automatic(ity)', used specifically in relation to medication adherence behaviours. The protocol states that the keyword 'Routine(s)' would also be searched at this stage but further experience from piloting this stage of the screening revealed that whilst some studies were using the keyword routine, it was unclear if they were using this in reference to habit formation and therefore this was excluded from the search process at this stage (see *Discussion* for more detail on this decision). If identified as potentially relevant in this stage, texts progressed to the second stage in which they were screened against the remaining inclusion and exclusion criteria.

Conference abstracts and journal supplement abstracts reporting potentially relevant studies were followed up to identify any journal articles which reported the study in full. The authors of protocols detailing relevant interventions were emailed to request trial results if available[1], and full-texts of articles that could not be accessed by other means were requested from authors.

3.4.5 Data extraction

Study characteristics (design, recruitment, retention, population details, sample characteristics, study setting, follow-up), intervention features (including detailed coding of BCTs - see below), description of comparison groups, and adherence and habit outcomes were extracted using a data extraction form. Conceptualisations of habit were derived by reading descriptions around the location of keywords ('habit(s/ual)' and /or 'automatic(ity)') in the full-texts or (associated documentation which were referenced in the methods) and detailed intervention content.

3.4.5.1 Intervention BCT coding

Two reviewers [LR, SD] independently coded intervention BCTs using Michie and colleagues' taxonomy of 93 BCTs v1 (Michie et al., 2013). Intervention descriptions were extracted; if protocols or methodology papers were available, these were also examined for additional details on intervention content. Differences were then discussed and a third expert reviewer [MA] was included to resolve any outstanding discrepancies. Some studies described some or all intervention features using Michie et al.'s (2013) taxonomy (indicated in Table 3 and Appendix C, Table C1). In these instances, BCTs were extracted directly and the remaining text was checked for any additional BCTs that could be coded.

---

[1] This led papers being included that were published after the end date of the searches but for which results were provided prior to completion of data extraction (Barankay et al., 2020; Wildman et al, 2021).

3.4.5.2 Adherence data extraction

The majority of studies included more than one summary of adherence, and/or at more than one timepoint. 'Primary' outcomes for the purposes of principle comparisons in this review were selected in the following order of priority: outcomes nearest 6 months; stated as the primary outcome/ power calculations based on expected effect size of this outcome; objective measurement tool; taking adherence (i.e. the proportion of medications taken relative to those prescribed). This was favoured over scheduling adherence (i.e. the proportion of doses taken within a specified time period relative to those prescribed), as taking adherence is most similar to other primarily extracted outcome measures, compared to scheduling adherence which is more stringent). Adherence outcomes summarised at time-points <6 months, for within-group analyses or for sub-groups of participants were not extracted.

3.4.6 Quality assessment

The Cochrane revised risk of bias tool and variant for cluster RCTs (Higgins et al., 2019) was used to conduct quality appraisals. One reviewer [LR]conducted quality assessments. The primary extracted outcome was assessed for each study. Protocols and/or trial registry records were used to support assessment, when available. All five domains were assessed but items relating to blinding of participants and interventionists to treatment allocation were not assessed, due to the nature of behaviour change interventions being inevitably unblinded. To enable algorithm calculation, these items were scored as low risk for all studies. The cluster randomised trial tool includes two assessments for domain 1. In order to compare assessment of bias across individual and cluster randomised studies, assessments were combined on this domain: studies assessed as low risk on both items of this domain were given an overall summary of low; studies with at least one assessment of 'some concerns' or 'high risk' were summarised as 'some concerns' or 'high risk' overall, respectively. This is in line with overall summaries as described in the Cochrane RoB handbook (Higgins et al., 2022).

3.4.7 Data analysis

An assessment was conducted to determine the appropriateness of meta-analysis of findings from studies eligible for inclusion. The homogeneity of populations, intervention types, intervention content, adherence measurement tools and summary measures and timepoints at which participants were followed-up across all studies was considered. Owing primarily to between-study variation in conceptualisations and use of 'habit' as a construct in this review (*see Results*), and due to insufficient data available, a meta-analysis to determine the effectiveness of habit-based interventions would not be appropriate or achievable with this pool of studies. A planned, narrative

review of conceptualisations of habit and the intervention BCTs was conducted. Following this, studies which explicitly used 'context-dependent repetition' (i.e. the principal BCT required to form a habit) were carried forward for examination of intervention effects. Effect sizes were calculated for adherence outcomes for each RCT based on the raw observed data; if this was not available this was calculated on the adjusted difference. Adherence outcomes were synthesised following the Cochrane Handbook 'Vote counting based on the direction of effect' method, as planned (McKenzie & Brennan, 2019). Probability of observing the overall direction of effect was assessed using a binomial probability test and confidence intervals calculated using the Wilson interval method (Brown, Cai, & Dasgupta, 2001).

3.5 Results

Searches yielded a total of 5973 studies of which, 17 articles were included (Figure 1). Two of these were protocols for studies for which the main articles were also included, thus screening identified 15 independent studies, including 18 intervention comparisons.

3.5.1 Overview of included studies

Characteristics of the 15 included studies are given in Table 2.

Interventions were designed for a variety of target populations, including people with cardiovascular disease and/or high cholesterol (n=5 studies), human immunodeficiency virus (HIV; n=3), respiratory disease (n=3), epilepsy (n=2), diabetes (n=1) and for people prescribed immunosuppressants following kidney transplantation (n=1).

Few details on the regimen prescribed for the included samples were given but likely ranged from requiring one opening of a pill bottle per day (e.g. Barankay et al., 2020) to a mean of 4.39 (SD= 1.8) doses of medication prescribed per day (Milam et al., 2005).
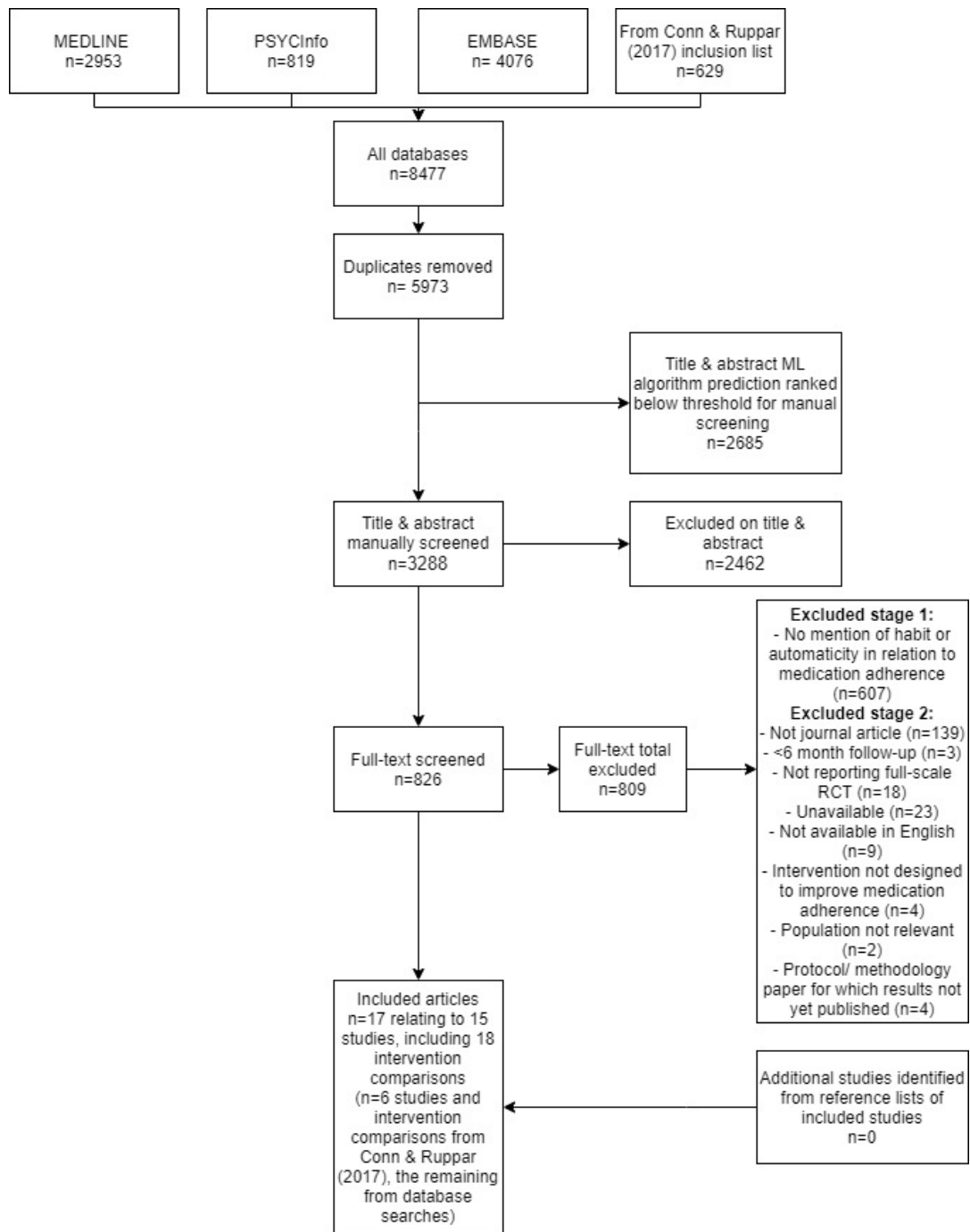
Figure 1
Screening flow-diagram

Table 2

Study characteristics

First author (Year)	Disease group and characteristics	Description of medication and dosage	Participant characteristics	n per group randomized	Intervention and comparator	Maximum follow-up period
Barankay (2020)	Statin prescription; self-reported non-adherence; LDL-C level >100mg/dL and diagnosed ASCVD /10-year CVD risk score >7.5% OR LDL-C >190mg/dL and no other risk factors OR both	Pills; One pill bottle opening per day expected	Mean age (SD) =58.5 (10.3); female=519 (64.5%)	Simple daily sweepstakes: n=199; Deadline sweepstakes: n=204; Sweepstakes plus deposit contract: n=201; Control: n=201	Simple daily sweepstakes; Deadline sweepstakes; Sweepstakes plus deposit contract; vs Control	6 months
De Bruin (2010)	Diagnosis of HIV; ≥6 months on HAART; ≥18 years old	Pills; Majority of participants (90%+) on polytherapy	Mean age (SD): intervention = 47.3 (9.8); control= 48.7(9.8); Male %: intervention= 92%, control= 90%	Intervention n=66; control n=67	Electronic Monitoring-Based Counselling vs usual care	8-9 months
Farmer (2016)	Diagnosis of type-2 diabetes for ≥3 months; ≥18 years old; no known CVD events	Pills; One medication session daily	Mean age (SD): Action planning= 61.5(11.1), usual care =63.9(12); Male%: Action planning=54, usual care= 59	Intervention n clusters = 30, n participants=265; Control: n clusters= 29, n participants=378	Brief action planning intervention vs usual care	12 months
Gregoriano (2019)	Diagnosis of asthma and/or COPD; prescribed daily inhaled medication and at least one exacerbation in the previous 12 months	Puff inhaled/dry powders; Intervention mean medications prescribed daily (SD)= 1.9 (0.8); Control mean(SD)=2.0 (0.8)	Mean age (SD): Intervention= 64.7(12.4); Control= 69.0 (8.8); Male(%): Intervention= 61%; Control(%)= 69%	Total n randomised (n)= 169; Total entering baseline visit (n)= 165: Intervention=84; Control n=81	Daily alarm clock and support phone calls vs control	6 months

Table 2

Continued

First author (Year)	Disease group and characteristics	Description of medication and dosage	Participant characteristics	n per group randomized	Intervention and comparator	Maximum follow-up period
Haynes (1976)	Cardiovascular medication pill-count<80% and high blood pressure (≥90 mm Hg); Dominion Foundries employees	Pills; dosage not stated	Mean age (SD) not reported. Male %: intervention = 100%, control=100%	Intervention n=20; control n=19	Behavioural intervention vs usual care	6 months
Lin (2017)	People with CAD that had undergone CABG; responsible for self-administration of medications; not meeting criteria for other pre-specified comorbid diseases or health problems	Pills; Prescriptions across the sample: >90% aspirin, 80%+ beta blockers, 60%+ ACE inhibitors, 70%+ lipid lowering drugs	Mean age (SD): Intervention=74.32 (5.26), control= 75.23 (5.82); Male%: intervention= 67.4% , control= 65.3%	Intervention: n centres=6, n participants=144; Control: n=6 centres, n participants=144	Multifaceted intervention including motivational interviewing vs usual care	18 months
Milam (2005)	Diagnosis of HIV for ≥3 months; attending participating HIV clinic; sexually active in past 3 months; ≥18 years old	Pills; Mean number of pills per day (SD)= 4.39 (1.8)	Mean age (SD): adherence intervention= 39(8.1); safer sex group=39(7.8); Male %: adherence intervention= 87.9%; safer sex arm=89.2%	Adherence arm: n clusters= 2 , n participants=149; safe sex intervention: n clusters= 4, n participants = 288	Brief adherence intervention vs safe sex intervention (no adherence component)	17-18 months
O'Dwyer (2020)	People prescribed salmeterol/fluticasone Diskus inhaler for asthma, COPD, or other (3-6% unknown diagnosis); filled ≥3 prescriptions in past 6 months	Inhaled medications; dosage not stated for this sample	Mean age (SD): Biofeedback group= 54(15), demonstration group = 53(15); control = 55(13); Male %: Biofeedback group= 42%; demonstration group = 57%; control = 50%	Biofeedback group: n clusters= 27, n participants = 74; Demonstration group: n clusters= 37, n participants = 56; Control group: n clusters =10, n participants =22	Biofeedback or demonstration vs usual care	6 months

Table 2

Continued

First author (Year)	Disease group and characteristics	Description of medication and dosage	Participant characteristics	n per group randomized	Intervention and comparator	Maximum follow-up period
Pakpour (2015)	Diagnosis of epilepsy; ≥18 years old; independently responsible for medication taking; prescribed AEDs	Pills; Majority of participants (~70%) on monotherapy[a]	Mean age (SD): intervention = 41.37 (16.25), comparator = 39.86(15.01); Female %: intervention= 32.8%; control = 35.5%	Intervention n=137; control n=138	Multimodal behavioural intervention vs usual care	6 months
Reddy (2017)	Veterans with diagnosis of CAD; 30-75 years old; <80% adherence at entry to study	Pills; 1 session daily	Mean age (SD): Individual feedback = 65.6(4.1), partner feedback= 64.9(6.2), usual care=64.1(6.6); Male %: Individual feedback = 100%, partner feedback = 96.3%, control=91.7%	Individual feedback n= 36; partner feedback n= 54; control n=36	Individual feedback or partner feedback vs usual care	26 weeks
Russell (2020)	People prescribed ≥1 twice-daily immunosuppressive medication following kidney transplant; ≥18 years old; functioning kidney transplant; no other life-shortening diagnosis	Pills; At least 1 medication, administered twice daily	Mean age (SD): SystemCHANGE= 53(11.2), Attention control=50.7(9.7); Male%: SystemCHANGE= 66.7%; attention control = 50%	Intervention n=45; control n=44	SystemCHANGE intervention vs attention control	12 months
Stacy (2009)	Dyslipidaemia; ≥21 years; enrolled in pharmacy plan and placed claim for a statin prescription	Pills; 57.8% prescribed 3+ medications	Mean age: intervention =54.6, control = 54.2; Females: intervention= 62.1%; control= 62.7%	Intervention: n= 298; Enhanced care control n=280	Interactive voice response technology intervention vs enhanced care control	6 months

Table 2

Continued

First author (Year)	Disease group and characteristics	Description of medication and dosage	Participant characteristics	n per group randomized	Intervention and comparator	Maximum follow-up period
Tang (2014)	Diagnosis of epilepsy; ≥16 years old; AEDs prescribed ≥6 months; non-adherence occurred in the last 6 months	Pills; Not stated	Mean age (SD): Education and behavioural intervention=30.8(11.6), Education only=31.6(13); male%: Education and behavioural intervention = 58.9%; education only= 49.1%	Education and behavioural intervention n= 65, Education only n= 59,	Education plus behavioural intervention vs education only	6 months
Tuldrà (2000)	Attending HIV-outpatient clinic; initiating first or second line HAART	Pills; Mean (SD) doses per day =3(2)	Mean age (SD): intervention=39(10), control = 38(7); male%: intervention= 73%; control = 79%	Intervention: n=55; Control n=61	Psychoeducative intervention vs usual care	48 weeks
Wildman (2021)	Diagnosis of CF; ≥16 years old	Nebulised; dosage not stated for this sample	Mean age (SD): Intervention=31.1(10.6), control= 30.3(10.8); Female%: intervention= 51.3% , control= 50.8%	Intervention n=305; control n= 303	CFHealthHub intervention vs usual care	12 months

CAD = Coronary artery disease; AED= anti-epileptic drug; ASCVD=. Atherosclerotic cardiovascular disease; LDL-C= Low density lipoprotein cholesterol; HIV= Human immunodeficiency virus; HAART= Highly active anti-retroviral therapy; COPD= Chronic obstructive pulmonary disease; CVD= cardiovascular disease; CABG= coronary artery bypass graft; CF= cystic fibrosis; SD= standard deviation
[a]based on n in table 1 of (Pakpour et al., 2015)

3.5.2 Conceptualisations of 'habit', intervention characteristics and 'habit formation' interventions

3.5.2.1 Conceptualisation of habit

Conceptualisations of what a habit is, references to key habit literature and provision of detail on how habits were targeted with specific intervention content varied greatly between studies.

Descriptions of habit in two studies closely aligned with contemporary habit theory, indicating both that habits arise from context-dependent repetition and that the development of habits reduces the need for effortful self-regulation of medication-taking behaviour (De Bruin et al., 2010; Wildman et al., 2021). Seminal literature on modern habit theory was referenced in both studies (e.g. Gardner & Rebar, 2019; Ouellette & Wood, 1998).

Barankay et al., (2020) dissociate three perspectives on habit from the psychological, economic and management literatures and apply these ideas to each of three intervention arms respectively, and examine their effects in maintaining the behaviour after withdrawal of incentives. In all three arms, including the control arm, a daily reminder was sent out. Whilst this could potentially act as a cue for the behaviour, this does not appear to be the active ingredient of their habit intervention arms, and whilst participants could choose the time at which this was received, the participants were not explicitly asked to link medication with the reminder. In the second arm, the authors state they draw on management theory in that 'habits arise as a consequence of newly established routines'. To encourage routines, participants in this arm were told they will only be entered into a sweepstake if they take their medications before they receive their reminder. The potential for reward was removed if behaviour was performed after the reminder each day. The idea of building medication-taking routines aligns more closely with contemporary habit theory in that existing routines could be used as a consistent cue to trigger performance of the new behaviour. However, there was no indication that the participant was supported to identify a cue or daily routine to enable repeated performance of the behaviour prior to their deadline, just that the behaviour itself was to become a routine. Other than specifying that the routine had to occur before the alarm, it is unclear what strategies were used to support this.

Reddy and colleagues (2017) designed their study to last 13 weeks, referencing key habit literature which demonstrated the average time for habits to form is 66 days (Lally et al., 2010). Reddy et al., (2017) stated that their aim was to create a '3-step habit loop' (referencing a book by Duhigg, 2012) but habit was here defined as medication-taking behaviour that is demonstrated by 'persistent adherence once the intervention is complete'. Context-dependent repetition was not indicated as a feature of habit formation in this study. Similarly, O'Dwyer et al., (2020) described habit as a behaviour which occurs regularly or routinely, but did not state that performance of the behaviour in a consistent context is also key to habit formation. Habitual performance of medication-taking in this intervention referred to both frequency

with which the behaviour occurred but also in the way the medication was administered. This distinction, regarding habitual 'instigation' vs habitual 'execution' has been the topic of discussion in the wider habit literature (e.g. Gardner et al., 2016).

Six studies refer to habit as a behaviour that can be achieved by linking or performing medication-taking behaviours with other routines or 'habits' (Farmer et al., 2016; Gregoriano et al., 2019; Haynes et al., 1976; Milam et al., 2005; Russell et al., 2020; Tang et al., 2014). Whilst the need to link medication-taking behaviour with an existing routine was clearly indicated in the intervention description as well as in the background and rationale for most of these studies, the action planning intervention described by Farmer et al., (2016) did not explicitly mention that it was a routine (i.e. repeated behaviour or cue) which was to be linked with medication-taking. This was important for coding habit formation as a BCT (see below).

Conceptualisation and implementation of habit remains unclear in two studies (Stacy et al., 2009; Tuldrà et al., 2000). Both studies explicitly link medication adherence to 'developing a habit', 'habit formation' or supporting participants in the 'acquisition of habits', but provide no elaboration on how this was achieved, and made no reference to other publications, studies or theoretical descriptions of habit.

Two interventions conceptualised habit as automaticity, using the Self-Report Behavioural Automaticity Index (SRBAI; Gardner et al., 2012), a widely used measure of habit and behavioural automaticity (Lin et al., 2017; Pakpour et al., 2015). Pakpour et al., (2015) first mention behavioural automaticity in the methods when describing the SRBAI and is only discussed thereafter in terms of change in behavioural automaticity; no explicit explanation as to the purpose or value in increasing automaticity (or habit) is given. Lin et al., (2017) state the rationale for measurement of behavioural automaticity: "Behavioral automaticity reflects whether a patient engages in a behavior (e.g. taking medication) relatively automatically; that is, quickly, easily, and without the need for conscious thought". Neither intervention states that a key mechanism to achieve behavioural automaticity is through context-dependent repetition.

3.5.2.2 Intervention content and techniques

Behaviour change techniques are presented in Appendix C. The most frequently coded technique was 'Feedback on behaviour' (n=11 interventions), followed by 'Prompts/cues' (n=10), 'Problem solving' and 'Action planning' (both n=8) and 'Behavioural practice/rehearsal (n=6)'. Thereafter, all other identified intervention techniques (n=34) were coded for five or fewer interventions

Only two studies used a BCT taxonomy to describe intervention content; both used Michie et al.'s (2013) taxonomy. The number of intervention techniques coded within a study varied from two to 15. Whilst for some studies this is likely an accurate reflection of the BCTs actually implemented in the interventions, some BCTs were precluded from being coded in some studies due to lack of detail.

'Context-dependent repetition' was coded for six interventions (see below).

3.5.3 Interventions utilising 'Context-dependent repetition' and effects on adherence

Context-dependent repetition was identifiable in six of the 18 interventions (33%; Table 3), despite this being the primary mechanism by which theory predicts new habits will form (Gardner & Lally, 2018; Gardner et al., 2020). One study self-coded habit formation using Michie and colleagues' (2013) taxonomy (Wildman et al., 2021). For some studies, absence of this BCT could be due to a lack of clarity in intervention descriptions leading to inability to definitively code it. For example, some studies described enough to code action planning, but context-dependent repetition could not be coded because there was not a clear indication that the details of the action plan involved *repetition* of the action plan in a specific context (e.g., Farmer et al., 2016). Stacy and colleagues (2009) included a figure to describe their intervention content which only contained the phrase 'Developing a habit' but with no reference to a taxonomy or with any additional description to contextualise exactly what was meant by this, and therefore context-dependent repetition could not be definitively coded for this study. Generally, context-dependent repetition was coded when authors described linking medication-taking behaviour with a routine or cue which occurred on a *daily* basis, therefore implying repetition.

Habit formation was never used as a standalone technique. Coding of Michie et al.'s (2013) 'Habit formation' BCT necessarily requires coding of 'Behavioural practice'. Five of the six interventions also used 'Prompts/cues', 'Action planning' and 'Problem solving'.

All six interventions indicated that an interventionist helped the participant to identify a consistent cue or daily routine which was to be linked with the behaviour. One study utilised an alarm reminder and allowed the participant to choose the time of the alarm to coincide with their daily habits. Unlike the four other interventions, the alarm reminder itself was the intended cue in this intervention rather than the daily routine. Some interventions also described that the cue and behaviour were explicitly linked by writing them down on a dedicated action planning form which was designed to help with this task (Milam et al., 2005; Tang et al., 2014; Wildman et al., 2021).

3.5.3.1 Intervention effects on adherence

The present review aimed to examine the effectiveness of habit-based interventions. As identified by Michie and Prestwich (2010) referencing habit theories does not equate to an intervention being theory-based. For the present review, intervention effects were summarised for all included interventions (Appendix D), but more utility can be gained from focusing on only the studies which used context-dependent repetition. Use of context-dependent repetition as a BCT was minimally required for an intervention to be described as 'habit-based' in line with contemporary habit theory and therefore detailed examination of intervention effects has been limited to these studies.

Findings from all six of these interventions showed positive effects on adherence (100%, 95%CI=[61-100%], p=0.03), but effect sizes ranged from small (RR= 1.03, Tang et al., 2014) to very large (adjusted d=1.71, Wildman et al., 2021; see Table 3). Interventions were implemented in diverse study contexts, used a range of outcome measures and were conducted across a number of disease groups (cardiovascular disease, COPD, epilepsy, HIV, kidney transplant and cystic fibrosis); some delivered multiple, relatively frequent face-to-face sessions, whilst others appeared to deliver the intervention in a single clinic visit. The small effects found by Tang and colleagues (2014) related to 'improvements in adherence'; this was a dichotomous outcome (improvement/no improvement) indicated by the score on a self-report measure of adherence. Comparatively, more objective measures of adherence were used by Gregoriano et al., (2019), Haynes et al., (1976), Wildman et al., (2021) and Russell et al., (2020); these studies found medium to very large effects. The largest effect (adjusted d=1.71 ; Wildman et al., 2021) was also the study with the most BCTs coded, although this is likely attributable to the authors coding the intervention BCTs themselves, using Michie and colleagues' (2013) taxonomy. In addition, some of these interventions ranged from simplistic interventions (e.g. completion of an action planning form; Tang et al., 2014) to complex interventions (e.g. Russell et al., 2020; Wildman et al., 2021) designed to improve adherence by targeting issues across Capability, Opportunity and Motivation (Michie et al., 2013).

Table 3

'Habit formation' intervention content and findings.

First author (Year)	Intervention name	BCTs coded[a]	Intervention duration to assessment timepoint details	Data collection	Adherence summary and timepoint measured	Adjustment to estimates	n (I,C) analysed	Effect on adherence (nearest 6 months)	Direction of effect and effect size (nearest 6 months)[b]	Effect on adherence >6 months	Effect on habit
Gregoriano (2019)	Daily alarm clock and support phone calls vs control	1.2 Problem solving; 2.2 Feedback on behaviour; 7.1 Prompts/cues; 8.1 Behavioural practice/rehearsal; 8.3 Habit formation	Intervention delivered over 6 months	Electronic device or POEMs where appropriate for dry powder medications	Mean number of days (max 200) that taking-adherence was in target adherence range (80-100%)	n/a	I (Puff inhaler)= 57, I (Dry powders) = 41 C(Puff inhaler)= 60, C (Dry powders) = 49	Significant difference for puff inhalers and dry powder capsules: Puff inhalers, Intervention mean(SD)= 81.6(14.2), control mean(SD) =60.1(30.3), p<.001; Dry powders Intervention mean(SD)= 89.6(9.8), Control mean(SD)= 80.2(21.3), p=.01	Favours intervention **Puff inhalers: Unadjusted d (95%CI)= 0.90 (0.52, 1.28) Dry powders: Unadjusted d (95%CI)= 0.55 (0.13, 0.97)**	n/a	Not measured
Haynes (1976)	Behavioural intervention vs usual care	1.2 Problem solving; 1.3 Goal setting (outcome); 1.4 Action planning; 1.6 Discrepancy between current behaviour and goal; 2.2 Feedback on behaviour; 2.3 Self-monitoring behaviour; 2.4 Self-monitoring outcome of behaviour; 2.6 Biofeedback; 3.1 Social support (unspecified); 7.1 Prompts/cues; 8.1 Behavioural practice/ rehearsal; 8.3 Habit formation; 10.4 Social reward;10.10 Reward (outcome); 12.1 restructuring the physical environment	Intervention delivered over full 6 months	Unused pill count	Change in adherence at 6-months, adherence calculated as proportion of pills prescribed that are taken in the month of follow-up	n/a	I=20 C= 18	Significant difference: Intervention mean (SE)=65.8(8.2) vs control mean(SE) = 43.2 (10.1), Baseline adjusted difference = 22.8, p=.025	Favours intervention **Unadjusted Hedges g (95%CI)= 0.56 (-0.09, 1.21)**	n/a	Not measured

Table 3

Continued

First author (Year)	Intervention name	BCTs coded[a]	Intervention duration to assessment timepoint details	Data collection	Adherence summary and timepoint measured	Adjustments to estimates	n (I,C) analysed	Effect on adherence (nearest 6 months)	Direction of effect and effect size (nearest 6 months)[b]	Effect on adherence >6 months	Effect on habit
Milam (2005)	Brief adherence intervention vs safe sex intervention (no adherence component)	1.2 Problem solving; 1.4 Action planning; 7.1 Prompts/cues; 8.1 Behavioural practice/ rehearsal; 8.3 Habit formation	Intervention delivered until month 10-11	Self-report of number of pills taken over past 7 days	Percent of patients whose 7-day adherence >95% measured at 17-18 months	Income, ethnicity, employment status, AIDS diagnosis, HAART regimen (vs. non-HAART), and number of pills per day, adherence at baseline and clustering	I=149 C=288	No significant effect: Intervention n(%)=128 (85.9%) vs control n(%)=201(69.8%). Adjusted OR = 2.05, [95% CI: 0.92 to 4.56], p=.077	Favours intervention **(Unadjusted) RR (95% CI)=1.23 (1.11, 1.36)**	(Outcome measured at 17-18 months)	Not measured
Russell (2020)	SystemChange™ vs attention control	1.1 Goal setting (behaviour); 1.4 Action planning; 1.5 Review behaviour goals; 1.6 Discrepancy between current behaviour and goal; 2.2 Feedback on behaviour; 4.4 Behavioural experiments; 15.3 Focus on past success	Intervention delivered over 6 months	MEMS-cap	Adherence rate defined as doses taken on time/total doses at 6 months	Ethnicity, marital status, perceived health score, and perceived social support	I=45 C= 44	Significant difference: Intervention mean(SD)=0.81(0.25) vs control mean(SD)= 0.64(0.24), p<.001. Adjusted difference B= 0.2 (95% CI= 0.12 to 0.27; SE = 0.039, p<.001)	Favours intervention **Unadjusted d (95%CI)= 0.69(0.27, 1.12)**	Significant difference at 12 months: Intervention mean(SD)=0.65(0.37) vs control mean(SD)= 0.53(0.29), p=.004. Adjusted B=0.16 (95% CI= 0.06 to 0.26); SE = 0.05) p<.001. Favours intervention **Unadjusted d (95%CI)= 0.36 (-0.06, 0.78)**	Not measured

Table 3

Continued

First author (Year)	Intervention name	BCTs coded[a]	Intervention duration to assessment timepoint details	Data collection	Adherence summary and timepoint measured	Adjustments to estimates	n (I,C) analysed	Effect on adherence (nearest 6 months)	Direction of effect and effect size (nearest 6 months)[b]	Effect on adherence >6 months	Effect on habit
Tang (2014)	Education plus behavioural intervention vs education only	1.4 Action planning; 7.1 Prompts/cues; 8.1 Behavioural practice/ rehearsal; 8.3 Habit formation; 9.1 Credible source	Intervention delivered over full 6 months	Morisky Medication Adherence Scale (MMAS-4)	Number of individuals whose self-report adherence improved at 6-months.	n/a	I=56 C=53	Favours intervention. No significant effect: Intervention n improved (%)= 36(64.3%) vs Control n improved (%)=33 (62.3%), p=0.827	Favours intervention **RR(95% CI)= 1.03 (0.78, 1.37)**	n/a	Not measured

Table 3
Continued

First author (Year)	Intervention name	BCTs coded[a]	Intervention duration to assessment timepoint details	Data collection	Adherence summary and timepoint measured	Adjustments to estimates	n (I,C) analysed	Effect on adherence (nearest 6 months)	Direction of effect and effect size (nearest 6 months)[b]	Effect on adherence >6 months	Effect on habit
Wildman (2021)	CFHealthHub intervention vs usual care	**1.1 Goal setting (behaviour); 1.2 Problem solving; 1.4 Action planning; 1.5 Review behavioural goals; 1.6 Discrepancy between current behaviour and goal; 2.2 Feedback on behaviour; 2.3 Self-monitoring of behaviour; 3.2 Social support (practical); 4.1 Instruction on how to perform the behaviour; 5.1 Information about health consequences; 5.2 Salience of consequences; 6.1 Demonstration of the behaviour; 7.1 Prompts/cues; 8.1 Behavioural practice/ rehearsal; 8.3 Habit formation; 8.7 Graded tasks; 9.1 Credible source; 10.4 Social reward; 12.1 Restructure the physical environment; 12.5 Adding objects to the environment; 15.3 Focus on past success; 15.4 Self-talk; 16.3 Vicarious consequences**	Intervention delivered over 12-month period	PARI eTrack™ electronic device	Normative adherence (adjustments made for ideal treatment for effectiveness) at 12 months	Treatment arm, time in weeks, baseline adherence (first two weeks), and past-year IV days	I=295 C=293	Favours intervention. Significant effect: intervention mean(SD)= 52.9%(31.4) vs control mean(SD)= 34.9%(31.7), adjusted mean difference= 9.5% (95%CI, 8.6-10.4)	Favours intervention **Adjusted d (95%CI)= 1.71 (1.52, 1.90)**	(Outcome measured at 12 months)	Favours intervention

I= Intervention; C= Control; MEMS= Medication event monitoring system; POEMs= Polymedication Electronic Monitoring System; HAART= Highly active anti-retroviral therapy; SE= standard error; SD= standard deviation; CI= confidence interval

[a] BCTs in bold represent BCTs coded by the study authors directly, using Michie et al.'s (2013) taxonomy

[b] The effect size was estimated for each RCT based on the raw observed data; if this was not available on the adjusted difference.

3.5.3.2 Intervention effects on habit strength

Three studies (Lin et al., 2017; Pakpour et al., 2015; Wildman et al., 2021) measured change in habit strength and all used the SRBAI (Gardner, Abraham, Lally, & de Bruijn, 2012), but Wildman and colleagues' (2021) study was the only intervention of the three in which context-dependent repetition' was identified as a BCT. A small effect of the intervention on habitual automaticity was observed in this study, with an adjusted difference of 1.2 points on the SRBAI (scale range 4 to 20, 95% CI= 0.5 to 1.8, adjusted d= 0.31).

Both Pakpour et al., (2015) and Lin et al., (2017) rescaled the SRBAI to give an outcome measure scale of 1 to 5. A significantly greater increase in habit strength was observed by Pakpour et al., (2015) in the intervention group at 6-months compared to usual care (Intervention mean (SD) = 1.64(0.56); control mean (SD)= 1.35(0.49); Beta(SE)= 0.49(0.09) p<0.001) with a medium effect size (unadjusted d=0.55). Similar effects were observed by Lin et al (2017) at six-months (B(SE)=0.57(0.05), p<.01, unadjusted d=0.59) and were maintained at 12 months (B(SE)=0.47(0.05), p<.01, unadjusted d=0.55), and 18 months (B(SE)=0.5(0.06), p<.01, unadjusted d=0.53).

3.5.4 Risk of Bias

All included studies were assessed to have at least some quality concerns. Most studies scored high in risk of bias overall (n=12), mostly triggered by high risk of bias in domains relating to the randomisation process, missing outcome data and risk of bias in measurement of the outcome when participants self-reported adherence and were aware of their intervention assignment. Least concerns were expressed with risk of bias caused by intervention assignment, although all studies were coded low by default on items assessing blinding of allocation and blinding of interventionists in this domain. Four of the six interventions that were coded as including context-dependent repetition as a BCT, and for which intervention effects on adherence has been discussed, were coded as having high risk of bias and two with some concerns. Two of three interventions which measured change in habit were coded as having high risk of bias and one with some concerns. Risk of bias assessments are summarised in Table 4.

Table 4
Quality assessment. Domain 1: risk of bias arising from the randomisation process; Domain 2: risk of bias due to deviations from the intended intervention (effect of assignment to intervention); Domain 3: risk of bias due to missing outcome data; Domain 4: risk of bias in measurement of an outcome; Domain 5: risk of bias in selection of reported result. 'High' refers to high risk of bias and 'low' refers to low risk of bias in the respective domain.

Study	Domain 1	Domain 2	Domain 3	Domain 4	Domain 5	Overall summary
Barankay (2020)	Low	High	High	Low	Some concerns	High
De Bruin (2010)	Low	Low	High	Low	Some concerns	High
Farmer (2016)	Low	Some concerns	High	Low	Some concerns	High
Gregoriano (2019)[1]	Low	High	High	Low	Some concerns	High
Haynes (1976)[1]	Some concerns	Low	Low	Low	Some concerns	Some concerns
Lin (2017)[2]	Low	Low	High	High	Some concerns	High
Milam (2005)[1]	High	Some concerns	High	High	Some concerns	High
O'Dwyer (2020)	High	Low	High	Low	Some concerns	High
Pakpour (2015)[2]	Low	Low	Low	High	Some concerns	High
Reddy (2017)	Low	Low	Low	Low	Some concerns	Some concerns
Russell (2020)[1]	High	Low	Some concerns	Low	Some concerns	High
Stacy (2009)	Low	High	Low	Low	Some concerns	High
Tang (2014)[1]	High	High	High	High	Some concerns	High
Tuldrà (2000)	Some concerns	Low	High	High	Some concerns	High
Wildman (2021)[1,2]	Low	Low	Low	Low	Some concerns	Some concerns

[1]Studies coded to have used context-dependent repetition in the intervention content
[2]Studies which measured change in habit strength

3.6 Discussion

3.6.1 Summary of findings

This machine-learning assisted review of RCTs aimed to explore conceptualisations of habit and examine the effectiveness of habit interventions in maintaining adherence to medication in people with long term conditions. The review identified 15 RCTs, published between 1976 and 2021.

The findings indicate that a broad range of definitions, intervention characteristics and BCTs have been used to promote maintained adherence to medication and medication-taking 'habits'. Conceptualisation and use of the term 'habit' often did not align with contemporary habit theory in medication adherence interventions, and with a few exceptions, many of the interventions that were identified do not explicitly include context-dependent repetition as a BCT.

3.6.2 Conceptualisations of habit in adherence interventions

The findings of this review highlight the need to reserve the use of the term 'habit' to cue-dependent automaticity, acquired through repetition of behaviour in a stable context. One study drew on ideas about habit from three older literatures to develop three different intervention arms and examined effects against one another (Barankay et al., 2020). However, recent opinions demonstrate an appetite to aggregate interdisciplinary ideas on habit into a unified concept and theory, for consistency of application (Fleetwood, 2019; Gardner et al., 2021). Only two studies in this review made reference to seminal habit theory and literature (e.g., Gardner & Rebar, 2019; Ouelette & Wood, 1998), and emphasised that both repetition and context stability of the behaviour are necessary components. Other studies described definitions more aligned with lay conceptualisations of habit (see Brown et al., 2019), with focus only on the importance of repetition, but little or no explicit reference to the importance of a stable contextual cue. The remaining studies made reference to habits, or habitual/automaticity of behaviours but gave no elaboration on how this was to be established. This finding is echoed in a systematic review by Rebar and colleagues (2016), examining habit formation in physical activity and we reiterate the importance of Rebar and colleagues' recommendations that, in scientific contexts, the term 'habit' should only be used in line with contemporary habit theory.

Among the six interventions in which context-dependent repetition of behaviour was identified as a BCT, all used it in combination with other BCTs. Most frequently combined were: 'Behavioural practice/rehearsal'; 'Prompts /cues'; 'Action planning' and 'Problem solving'. Some interventions included habit as one of a combination of mechanisms of action, but Gardner & Rebar (2019) presented similar findings in a review of habit formation across a range of health behaviours. However, Gardner and Rebar's (2019) study did not include any medication adherence

interventions. The replication of this observation in medication adherence interventions suggests that, whilst 'Habit formation' (i.e., context-dependent repetition) is labelled as a single intervention technique in the BCT taxonomy (Michie et al., 2013, although see other taxonomies- Kok et al., 2016), a number of techniques might be used in combination with it to facilitate self-regulation and action control in the period whilst habits are forming. This aligns with expert consensus on the function of habit formation (behavioural cueing, behavioural regulation) with which other self-regulatory techniques have been associated, as well as with theoretical models of habit formation (Gardner & Lally, 2018; Gardner et al., 2020).

Only one study utilised incentives (Barankay et al., 2020); this study explicitly drew on economic theories of habit formation for this feature, but the importance of reward also relates to early behaviourist work on habit (e.g. Hull, 1943) in which external reward was thought to facilitate repetition of behaviour. The role of reward is heavily implicated in neurological models of habit formation (see Wood & Rünger, 2016). However use of material (financial) incentives alone is unlikely to have long-term effectiveness for maintaining behaviour beyond termination of their receipt (Mantzari et al., 2015). Barankay and colleagues (2020) combined incentives with other intervention content but did not emphasise the need to repeat adherence with a contextually stable cue. This therefore precluded coding of context-dependent repetition. Nevertheless, the use of external reward, rather than assuming the presence of an intrinsic reward when a goal is achieved during habit formation, is a theoretical concept which is seldom addressed in habit interventions. There is some evidence to suggest that people who find a behaviour more intrinsically rewarding are more likely to repeat behaviours during habit formation and achieve greater habitual automaticity (Gardner & Lally, 2013). This in turn has been evidenced to predict the maintenance of behaviour, via habit strength (Phillips, 2020; Phillips, Chamberland, et al., 2016). However, this evidence base is mostly gleaned from physical activity interventions, not medication adherence interventions. It is both intuitive and scientifically evidenced that physical activity improves mood and mental health and therefore is likely to relate to the degree to which a person experiences reward. It is unclear if the same degree of intrinsic reward can be achieved through adhering to medication. Therefore, the form which 'reward' takes in habit interventions for medication adherence may require some more careful consideration.

3.6.3 Overall effectiveness of the included interventions

The primary motivation to synthesise results among interventions was to understand the effects of habit interventions on medication adherence. However, conceptualisations of what a habit is and the BCTs used in these interventions were diverse. This, along with various different intervention types, various methods to capture and summarise adherence data, and the fact that only six of 18

interventions appeared to use the principal BCT required to facilitate habit formation, prevented meaningful meta-analysis of intervention effects. The focus of the synthesis was therefore on a detailed description of differences in conceptualisations of 'habit interventions' and comparisons of effectiveness among the six interventions which used context-dependent repetition.

The observed positive effects of these six interventions on adherence to medication indicate that context-dependent repetition in combination with other behavioural regulation techniques could be effective in sustaining improvements in medication adherence, for at least six months. This finding must be interpreted with caution as no studies made direct comparisons between interventions with and without habit formation as a BCT, and only one of these studies examined the effects of the intervention on habit strength (Wildman et al., 2021). Small positive effects on habit were found, but findings cannot be generalised beyond this study.

3.6.4 Strengths and limitations

A change was made to the inclusion and exclusion criteria in a deviation from the original review protocol, in order to facilitate a more refined and targeted synthesis of the use and conceptualisation of 'habit' in the existing medication adherence literature. The decision to exclude studies only making reference to 'routines' rather than explicitly using the keywords 'habit(ual)' or 'automatic(ity)' facilitated a detailed analysis of how habit had been conceptualised within this literature, without projecting theory and assumptions of the authors' intentions to build habits for their intervention in the absence of these keywords. Inclusion of studies which intended to form medication-taking routines, without further extension of the language screened for in the full-text screening stage, such as 'action planning' or description of linking medication-taking behaviours to other existing routines would have led to only a partial inclusion of this group of studies. The alternative, to expand the inclusion criteria to include studies describing this, would have been beyond the scope of the primary research objectives of this review. Furthermore, this expansion would involve some pre-specification of the BCTs which would likely lead to habit formation. This approach, whilst valid, is opposite to the objectives of this review, which aimed to identify which BCTs were being used in interventions with the intention to form medication-taking habits.

The heterogeneity in conceptualisations and implementations of habit formation interventions in this review meant   that a meta-analysis could not be conducted. However, valuable insights were gained by turning attention to a narrative characterisation of the differences in conceptualisations of habit. In the opinion of the reviewers, the most important finding from this study is the need to reserve the term 'habit' for instances in which a behaviour is automatically cued by a stimulus. Use of modern theories of habit will support the development of medication adherence interventions.

Quality assessments of the included studies led to the judgement that 12 of the 15 studies presented evidence with a high risk of bias; the remaining two studies presented with at least some concerns. Four of the six studies included in the discussion of intervention effectiveness were coded as high risk of bias. However, given the focus of the review, confidence in the bias of effect sizes was less important than the confidence in determining all of the components of the included interventions. Insufficient detail may have meant that fewer BCTs were coded than were actually present in some interventions. This finding is not unique to this review (Michie & Johnston, 2012), and is echoed across reviews which have attempted to code interventions, across a range of fields (e.g. Bruin et al., 2020; Candy et al., 2018).

An adaptive learning method utilising machine learning classifier was used to assist with title and abstract screening. This is a novel approach which is under rapid development in the interdisciplinary field of computer science applications to text mining for reviewing (O'Mara- Eves et al., 2015). A key success of this approach was the reduction in workload for the reviewers. The iterative approach, combining the speed and power of the text classifier with the conceptual input of the reviewer in steering the learning of the classifier, enabled title and abstract screening completion when 55% of the titles and abstracts retrieved from database searches had been manually labelled. Fewer relevant articles were identified with each 5% increment in screening, with only an additional 8 articles found in the final 299 studies screened. This technique could be powerfully implemented in future reviews, especially when a large number of records are identified in the database search stage.

3.6.5 Implications and recommendations

The key finding of this review is that there are a limited number of theory-informed habit-based interventions, designed to enable sustained adherence to medications in long-term conditions. Interventions that utilise context-dependent repetition as a BCT, in combination with other behavioural regulation techniques, show promising outcomes for eliciting sustained improvements in medication adherence for at least six months. However, few interventions of this nature have been tested in RCTs in the medication adherence literature, and the existing evidence is mostly low in quality. For a number of interventions included in this review, 'context-dependent repetition' (equivalent to Michie et al.'s (2013) 'Habit formation BCT) was coded because the repetition of a behaviour was tied to another daily routine. Authors' explicit use of taxonomies to code this BCT will facilitate future syntheses which aim to aggregate the findings of habit intervention studies in medication adherence. Only three studies in this review included a measure of habit, and few studies evidenced the use of contemporary habit theory to support the identification of relevant BCTs. Recently, literature has been produced specifically to facilitate the design of interventions and

studies of habit formation in health behaviours (e.g. Gardner et al., 2020, 2021). Recommendations from these sources should facilitate this process, even among intervention designers unfamiliar with this rapidly growing literature, or from other disciplines.

3.7 Reflections on Chapter 3 and implications for the following chapters

The literature review presented in Chapter 1 highlighted that there is strong theoretical rationale to promote the formation habits in people prescribed long-term medication regimens to support the maintenance of adherence. Chapter 3 presented a publication of a machine learning-assisted review which was conducted to explore Aim 1 of this thesis. The aim was to examine the conceptualisation of 'habit' and the implementation and utility of 'habit formation' in interventions to support the maintenance of medication adherence. This was in response to recognition of a disparity in findings between the behaviour change and medication adherence literatures. Before descending into a detailed and technical analysis of the best way to capture habit in medication adherence behaviour in the following chapters, it was necessary to address this discrepancy. Future work can benefit from the review's explicit clarification of common misconceptions about what a habit is and how it can be targeted but specifically for the thesis, the review provided two key learning points: first, that many devices are available for research purposes which capture the date and time at which participants with long-term conditions administer their medications; second, that the translation of habit theory into medication adherence contexts is in its infancy and therefore it is timely to start work on metrics of habit which leverage the data made available from these devices, and which are explicitly grounded in contemporary habit theory.

Chapter 4 Properties of the Behaviour Frequency x Context Stability measure and considerations going forward

4.1 Chapter overview

Having addressed Aim 1 in Chapter 3, Thesis Aim 2 was to explore whether the time of day at which people take their medications can be used to derive an objective metric of context stability for use in habit research. As discussed in Chapter 1, combined with an objective measure of frequency of behaviour using electronic devices, an objective measure of context stability could generate an objective implementation of the BFCS measure. Previous research identified this as a promising method to pursue in the objective measurement of 'habit' or likelihood of habitual behaviour.

Chapter 4 begins to explore Thesis Aim 2 with the following objective:

1) To critique previous implementations of the 'behaviour frequency x context stability' measure to understand how this approach could be implemented with objective data (Thesis Aim 2, Objective 1).

The critique highlights a series of issues with the use of a multiplicative score derived from behavioural frequency and context stability metrics and concludes with specific rationale for the trajectory of the following thesis chapters.

4.2 Background

Sometimes referred to as a 'classic' measure of habit (Wood, 2017, p.393), the origin of the frequency in context approach is often attributed to a seminal article published by Ouellette & Wood (1998). Following a meta-analysis of studies examining the effects of past behaviour, subjective norms, attitudes and intentions on future behaviour, Ouellette & Wood (1998) concluded that "past behavior directly contributes to future performance in contexts that support the development of habits. Behaviors that are well practiced and performed in stable contexts are likely to be repeated because they can be performed quickly [and] relatively effortlessly" (p.65).

Ouellette & Wood (1998) supported their findings with a primary research study in which multiple linear regression models predicting future behaviour were constructed using past behaviour and context stability, alongside a range of other key psychological constructs commonly implicated in the

prediction of behaviour. They found that information about both the context *and* frequency of past behaviour were key; that is, only when context was stable and could therefore facilitate habitual performance of behaviour, was frequency of past behaviour useful to predict future behaviour. This was an important turning point in the history of habit research as until this point, habit had typically been represented by past behaviour frequency, alone (e.g. Triandis, 1977).

In 2005, Wood et al., (2005) presented the first published example of what has come to be known as the prototypical 'frequency in context' or 'behaviour frequency x context stability' (BFCS) measure. Wood et al., (2005) measured the 'habit strength' of existing behaviours (i.e., past behaviour) including exercising, reading the newspaper, and watching TV, in the participants' old university, and measured the change in frequency of behaviour (i.e., future behaviour) after moving universities (i.e., changing context). 'Habit strength' in each of these behavioural domains was measured as the multiplicative product of frequency of past behaviour and the stability of the context in which it had been performed:

$$\text{Habit strength (BFCS)} = \text{frequency of behaviour} * \text{context stability of behaviour} \qquad [1]$$

High scores indicated highly frequent and contextually stable behaviours (i.e., likely to be habitual), whereas low scores indicated less frequent or less context-dependent, or less frequent *and* less context-dependent behaviour (i.e., unlikely to be habitual).

Since then, BFCS has been used to measure habit across a range of behavioural domains (e.g., eating healthful or unhealthful snacks, physical activity, flossing, drinking alcohol, handwashing), but notably has been implemented in many ways. For example, the Wood et al., (2005) original implementation of BFCS measured frequency of past behaviour on a scale from zero (low frequency) to three (high frequency), and context stability on a scale from zero to three, which once multiplied resulted in a BFCS scale from zero to nine, where high scores indicated stronger habits. However, two years later, Ji & Wood (2007) took the raw number of times a behaviour was performed over a given period and multiplied it by the context stability score which ranged from zero to three. This resulted, in their sample, in a scale ranging from one to 45 for purchasing fast food and a scale from zero to 69 for watching TV[2]. In the following year, Danner et al., (2008) published another study using BFCS to measure habit in snacking, drinking milk, drinking alcohol and in travel behaviours. In their study, frequency was measured from zero (never) to eight (very frequently), and context measured from one (unstable) to nine (stable), resulting in a scale from zero to 72. More recently, Sheeran & Conner (2019) measured past behaviour frequency by asking participants to indicate on a

---

[2] Note this was sample specific and would change if a participant performed a behaviour more or less frequently than the observed extremes

seven-point scale, 'In the past four weeks, I have eaten five portions of fruits and vegetables per day' and context stability by asking 'Is eating five portions of fruit and vegetables a day something that you would do at the same times and in the same places each time', again on a seven-point scale ('definitely no' to 'definitely yes'). Even within the same study, some researchers have opted to change how BFCS was implemented, despite studying the same behaviour (see Neal et al., 2013). These differences in the implementation of BFCS measures have yet to be addressed in terms of their effects on how the measure, as an index of habit, is to be interpreted.

This discussion is timely; as highlighted in Chapter 1, there is increasing interest in the utility of objective measures of behaviour and psychological constructs in the study of behaviour change. BFCS measures have been identified as an important opportunity for the future of objective measurement in habit research (e.g. Hoo et al., 2019; Naab & Schnauber, 2016). Before pursuing this further, it is important to address the properties of the BFCS measure as it has been historically implemented to ensure future developments in this approach are meaningful and interpretable.

The present chapter aims to highlight a series of properties of BFCS measures as they have been implemented to date, and to present specific rationale for work conducted in the following thesis chapters. This discussion pertains only to the properties of BFCS as a composite term described by the multiplication of its two numeric components, frequency and context. Whilst the specific wording used in measures of habit may also impact on how individuals respond to questions (e.g. Gardner & Tang, 2014), the issue is not discussed here as it is not relevant to the focus of the remaining thesis, which is objective measures.

4.3 Properties of BFCS measures in existing research

The following section discusses some of the properties of BFCS measures, as used in the existing literature. This discussion aims to highlight the effects of *multiplying* frequency by context to create the BFCS score, and the effects of using small vs large and matched vs unmatched scales to capture frequency and context stability.

4.3.1 Example with small, matched scales for frequency and context

The first published use of a multiplicative 'BFCS' measure as an index of habit was in Wood et al., (2005). This was a self-reported frequency in context measure using small, matched scales for its components. Frequency was self-reported on a scale from 0 to 3: 0 (I never perform the behaviour), 1 (monthly or less often), 2 (at least once a week), 3 (just about every day). Context stability was captured by asking participants whether the participant typically performed the behaviour in the same location, again on a scale from 0 to 3: 0 (did not perform the behaviour), 1 (rarely or never in the same location), 2 (sometimes in the same location), and 3 (usually in the same location). The

resulting scale therefore ranged from 0 to 9 and was used to measure 'habit strength' in the prediction of exercising, reading the newspaper, and watching TV.

Table 5 presents calculated BFCS scores and their descriptions for a range of possible responses on the Wood et al., (2005) scale.

Table 5
Examples of some possible values and their interpretation from the Wood et al., (2005) BFCS scale.

Example	Frequency [a]	Context[b]	Scale values	Interpretation	BFCS score[c]
1	0	0	Lowest frequency	I never perform the behaviour; I did not perform the behaviour	0
2	3	0	Highest frequency, lowest context stability	I perform the behaviour just about every day; I never do the behaviour	Not valid

4.3.1.1 Frequency and context are conflated

Valid scores on this BFCS scale were 0, 1, 2, 3, 4, 6, 9. Example 1 (lowest frequency) and example 2 (i.e. highest frequency, lowest context stability) in Table 5 demonstrate that not all values of the scale are valid; it is not meaningful for a participant to respond with frequency equal to one, two or three in combination with the lowest context stability score (i.e. context equal to zero), as these scores would mean that the person is doing the behaviour but also never doing the behaviour. In this implementation of a BFCS measure, the lowest context stability score is a measure of frequency of behaviour and when multiplied by the frequency value, still represents a measure of frequency alone.

4.3.1.2 The BFCS scale is non-continuous

Figure 2 depicts how, with increasing frequency and context scores, the Wood et al., (2005) scale becomes increasingly non-continuous; values of five, seven and eight are unattainable on this scale.
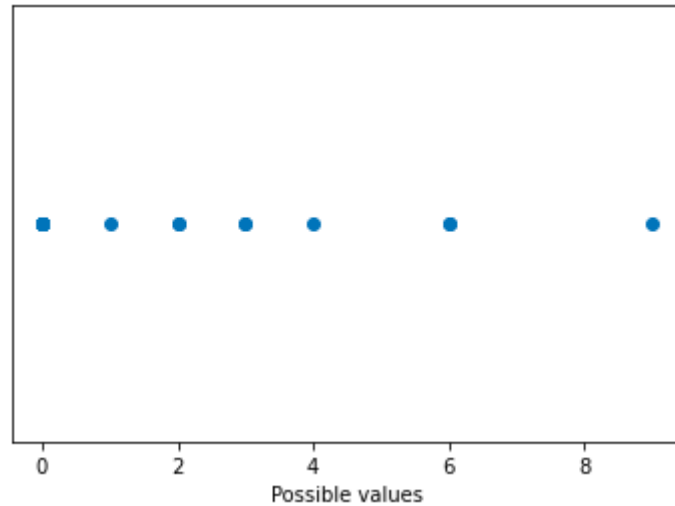
Figure 2
Distribution of possible scores on the Wood et al., (2005) BFCS measure.

Non-continuity is important for the interpretation of differences in BFCS scores, particularly when linear regression analyses are used. Often, (as in Wood et al., 2005) researchers present regression coefficients to demonstrate the size of the effect of the BFCS measure in the prediction of future behaviour. For regression coefficient 'x', we interpret that for a one unit increase in BFCS score, we predict an increase of 'x' in the outcome behaviour. The treatment of ordinal variables as interval in statistical analyses is common in the psychological literature and some have argued should not cause concern (Michell, 2008), unlike treating ordinal scales from 1 to 5 as interval, the multiplication of values results in gaps in valid values toward the higher end of the BFCS scale. Whilst an increase in BFCS score of one at the lower end of the scale is possible and therefore meaningful, it is not at the higher end of this scale, making interpretation of the coefficient difficult; it is not possible to conclude that a step from a score of four to six would predict an increase of x, 2x, or another value altogether in the outcome variable. In fact, the only instance in which the interpretation of increasing a BFCS score by one is valid is when either frequency or context remains equal to one, and the other only increases by one.

4.3.1.3 Changes in component scores result in non-uniform changes in the BFCS score

Multiplication of frequency and context creates non-uniform effects on the BFCS score when the same changes are observed on each component and the starting values are different. Table 6 demonstrates that both participants B and C are one point higher than A on the context or frequency score, respectively, yet the overall BFCS score has not increased by the same amount for both participants; participant B is two points higher than participant A on the BFCS score, but participant C is only one point higher than A. This effect implies that a one unit increase in context score has more impact on habit than a one unit increase in frequency. However, importantly, this

effect would also happen in reverse if all scores were reversed (i.e., a one unit change in frequency could also have more impact than a one unit increase in context). To justify this feature of the BFCS measure, evidence that: i) increasing the score on the component with the lower value has more effect on habit than increasing the component with the higher starting value, and ii) having more closely matched frequency and context is more important than having disparate context and frequency, is required.

Table 6
Calculated BFCS scores using Wood et al., (2005) scale.

	Frequency	Context	BFCS score
Participant A	2	1	2
Participant B	2	2	4
Participant C	3	1	3

4.3.1.4 BFCS scores are commutative

Whilst there is only one way to obtain a BFCS score of one, four and nine on Wood et al.'s, (2005) scale, there are two ways to have a score of two or three, for example:

Frequency=3, Context=1; BFCS=3

Frequency =1, Context=3; BFCS=3

For any instance in which the scores for frequency and context are not matched, this will be the case. As remarked by Gardner (2015), this implies that someone acting in the same context but infrequently, is equally as likely to be acting habitually as someone acting frequently in different contexts. Whilst this is not mutually exclusive with the assumptions described in relation to non-uniformity in change, it requires further demonstration that the value of context and frequency in determining the likelihood of habit are equal.

4.3.2 Example of the use of BFCS with unmatched and/or larger scales

The example from Wood et al., (2005) demonstrates the properties of BFCS when equal scales are used for capturing frequency and context, with a relatively small scale. However, not all studies have implemented BFCS with matching scales for the two components, and many have used larger scales. Danner et al., (2008) measured both frequency and context on nine-point scales, but each component was captured using different bounds: frequency was measured from 0 (never) to 8 (very frequently), whereas context was measured from 1 (unstable) to 9 (stable). As in Wood et al.,'s

(2005) study, the multiplicative product of frequency and context was used to indicate 'habit strength'. The following sections describe some further properties of BFCS when scales for frequency and context are not matched and are captured on larger scales.

4.3.2.1 Frequency and context are not conflated

The context scale used by Danner et al., (2008) improves on that used by Wood et al., (2005) in that context stability is no longer conflated with frequency of the behaviour; the lowest context stability score (1= unstable context) is not an indicator of frequency.

4.3.2.2 The BFCS scale is still non-continuous

Danner et al., (2008) report that multiplication of their frequency and context scales results in a 'relatively continuous' BFCS measure, ranging from 0 to 72. However, this scale is subject to the same interpretation issues due to scale non-continuity as Wood et al., (2005). To demonstrate, Figure 3 shows how the lower end of the scale is relatively continuous but with increasing values, larger gaps between valid values are observed. Because of the larger scales, these gaps are also larger in value as the scale maximum is approached, with the largest step of eight points between the highest possible values (64 and 72).
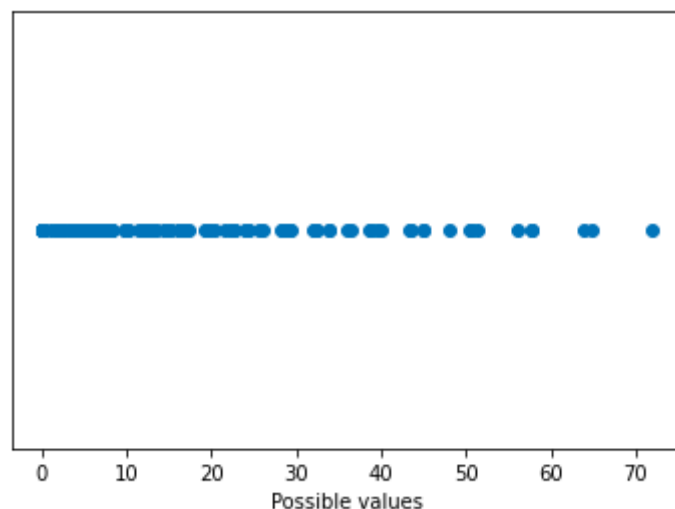
Figure 3
Distribution of BFCS scores on Danner et al., (2008) scale.

4.3.2.3 Changes in component scores and changes at different ends of component scales result in non-uniform changes in the BFCS score

The same effects on BFCS scores are also observed as Wood et al., (2005) for differences in context or frequency when one component remains constant and the other increases or decreases (

Table *7*).

Table 7
Demonstration of the effects of change in BFCS score when frequency and context change but are imbalanced.

	Frequency	Context	BFCS	Change from baseline
Baseline	7	1	7	
Increase context by 1	7	2	14	+7
Increase frequency by 1	8	1	8	+1

As in Wood et al., (2005) the effect demonstrated in

Table *7* implies that an increase of one point in the lower of the two scores has more effect on habit than an increase of one in the higher of the two scores. However, because the scales are larger, this effect is also numerically larger.

Furthermore, larger scales mean that differences are also observed when a change occurs at the low compared to the high end of the scale. As demonstrated in Table 8, an increase of one at the low end of the scale results in a much smaller increase than the same increase at the high end of the scale. When BFCS measures are implemented in this way, they imply that an increase in one point at the higher end of the scale has more effect on habit than an increase in one at the lower end of the scale.

Table 8
Demonstration of the effects of change at the low vs high end of the Danner et al., (2008) BFCS scale (these effects are demonstrated with equal start values for frequency and context, but this would also apply for unequal start values).

	Frequency	Context	BFCS	Change from baseline
Low scorers				
Baseline	1	1	1	
Increase context by 1	1	2	2	+1
Increase frequency by 1	2	1	2	+1
High scorers				
Baseline	7	7	49	
Increase context by 1	7	8	56	+7
Increase frequency by 1	8	7	56	+7

4.3.2.4 BFCS scores are commutative

When scales are unmatched the issue of commutative scores remains, but in a slightly different manner than that observed in Wood et al., (2005). The fact that scales are unmatched means that, numerically, high frequency is not valued as highly as high context.

Table *9* demonstrates the effects of this on the BFCS score.

Table 9

Demonstration of the effects of having unmatched scales on the interpretation of BFCS scores.

Description	Frequency (scale 0 to 8)	Context (scale 1-9)	BFCS
Highest frequency, 3rd highest context	8	7	56
2nd highest frequency, 2nd highest context	7	8	56
3rd highest frequency, highest context stability	6	9	54
Highest frequency, 2nd highest context stability	8	8	64
2nd highest frequency, highest context stability	7	9	63
Lowest frequency, 2nd lowest context stability	0	2	Not valid
2nd lowest frequency, lowest context stability	1	2	2

For different combinations of frequency and context scores, the same BFCS value can be obtained but has different interpretations in terms of the relative strength ratings on the respective component scales. For example, a BFCS score of 56 can be obtained with the second highest frequency and context score, but also with the highest frequency and the third highest context score. However, having the highest frequency and second highest context score results in a higher BFCS score (BFCS= 64) than having the highest context and second highest frequency score (BFCS=63). Whilst subtle, the assumptions implied by these inconsistencies have not yet been empirically evidenced or theoretically justified.

4.4 Considerations for future use of BFCS measures with respect to objective measurement of frequency and context

4.4.1 The issue of conflation of frequency and context

During scale development, habit researchers should be vigilant to the idea that frequency and context must be, at least conceptually, independent of one another. As in the implementation by Danner et al., (2008) context scales should have valid scores which can be measured independently from the frequency scale, meaning the lowest end of this scale indicates 'unstable' contexts.

Whilst in principle this can be relatively easily avoided when BFCS is implemented as a self-report measure, it may be overlooked if described using objective data. When data about the context in which a behaviour occurs is captured automatically, the availability of data to calculate context stability is directly related to the frequency of the behaviour. For example, a researcher may wish to

describe the context stability of an individual's behaviour by calculating the interquartile range (IQR) of the times of day at which the behaviour happened, over the period of a week; low IQR would indicate high context stability, high IQR would indicate low context stability. At the extreme, the behaviour may have only occurred once during that week, therefore only one data point about the context will be available and IQR will equal zero, resulting in the highest possible context stability score. In any instance in which too few data points are available to accurately reflect the actual context stability of the behaviour, there is a risk of over-, or under-estimating context stability. When this occurs, information about frequency will leak into context stability and could cause issues relating to multicollinearity.

There are two potential solutions to delineate frequency from context:

1) Decide a minimum number of data points required to be included for analysis, or;
2) Collect a specified number of data points with which to calculate context stability for all individuals, meaning all participants included in the analysis will contribute the same number of data points to calculate 'context stability'.

The latter option is more conservative and will ensure that context is, at least conceptually, completely delineated from frequency.

4.4.2 The issue of non-continuous BFCS scales

As demonstrated, non-continuous BFCS scales can cause difficulties in the interpretation of their effects in standard linear regression modelling, as is frequent in the habit literature. A potential solution is to treat frequency and context as percentages. For example:

$$\text{BFCS= (percent frequency* percent context stability)/100} \qquad [2]$$

Treating both frequency and context as percentages means both can take decimal point values which, when multiplied, will produce a continuous scale (see Figure 4). Furthermore, the extra step in formula [2] of dividing the multiplicative value by 100 will rescale the measure to take values from zero to 100. This may facilitate more meaningful interpretation of BFCS scores than scales (such as Danner et al., (2008)'s 0-72 scale), as this is a scale that users will be accustomed to working with.
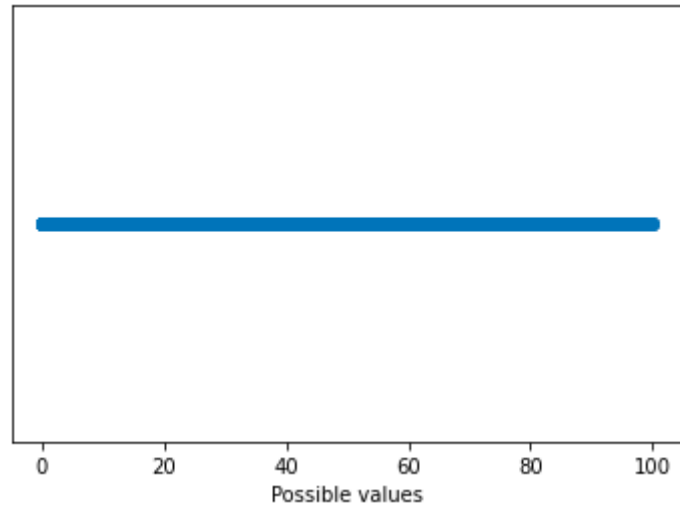
Figure 4
Distribution of BFCS scores when continuous percentage scales for frequency and context used.

This approach requires knowledge of the expected denominator of each component, such as how often the behaviour was expected (if prescribed) or the maximum expected frequency, over the specified period of measurement.

4.4.3 The issue of changes in component scores resulting in non-uniform changes in the BFCS score

Using percentage scales alone does not resolve the issue of non-uniformity of change. The disparity in the effects of change at low vs high ends of the scale persist; although more subtle than in the example from Danner et al., (2008), an increase of one at the low end of the scale results in a much smaller increase than an increase of one at the high end of the scale (Table 10).

Table 10
Demonstration of the non-uniformity effect when differences in percentage scores are observed at different points on the BFCS scale.

	Frequency	Context	BFCS	Change from baseline
Low scorers				
Baseline	10	10	1	
Increase context by 1	10	11	1.1	+.1
Increase frequency by 1	11	10	1.1	+.1
Mid-range scorers				
Baseline	50	50	25	
Increase context by 1	50	51	25.5	+.5
Increase frequency by 1	51	50	25.5	+.5
High scorers				
Baseline	90	90	81	
Increase context by 1	91	90	81.9	+.9
Increase frequency by 1	90	91	81.9	+.9

Furthermore, when frequency and context do not take the same value, changing either score by one still does not result in uniform increase in BFCS score (Table 11).

Table 11

Non-uniformity in BFCS score changes when composite percentage scores change.

	Frequency	Context	BFCS	Change from baseline
Baseline	80	10	80	
Increase context by 1	80	11	88	+8
Increase frequency by 1	81	10	81	+1

When calculated as described in [2], BFCS can only increase by one point when one domain increases by one and the other is equal to 10. As before, greater changes will always be observed if the change occurs in the domain with the smaller value than if change occurs in the larger.

This property is the resultant effect of multiplying two numbers together and could only be resolved by combining the two values with another method, such as by taking the mean. When the mean of two numbers is calculated, the effect of change is equal at low, medium and high ends of the scale (Table 12) and is not affected by changes when the values of context and frequency are not the same (Table 13).

Table 12

Effect on BFCS score when changing frequency or context at low vs high ends of the scale when taking the mean

	Frequency	Context	BFCS	Change from baseline
Low scorers				
Baseline	10	10	10	
Increase context by 1	10	11	10.5	+.5
Increase frequency by 1	11	10	10.5	+.5
Mid-range scorers				
Baseline	50	50	50	
Increase context by 1	50	51	50.5	+.5
Increase frequency by 1	51	50	50.5	+.5
High scorers				
Baseline	90	90	90	
Increase context by 1	91	90	90.5	+.5
Increase frequency by 1	90	91	90.5	+.5

Table 13

Change in BFCS when calculating the mean and observing change from imbalanced frequency and context

	Frequency	Context	BFCS	Change from baseline
Baseline	80	10	45	
Increase context by 1	80	11	45.5	+.5
Increase frequency by 1	81	10	45.5	+.5

When the mean is used to create a composite score, the assumption is that both frequency and context contribute equally to the overall BFCS score. Until the relative importance of each component of the composite BFCS score is empirically demonstrated, use of the mean to produce a composite score could provide a more parsimonious method of implementing a BFCS score due to fewer implied assumptions.

4.4.4 The issue that BFCS scores are commutative

Finally, even on the assumption that frequency and context contribute equally to habit, and if percentage scales are used to describe frequency and context, and if the mean is used to produce a balanced composite BFCS score, the score remains commutative. Without further interrogation of scores post-hoc, it is impossible to state whether any differences between or changes within individuals are driven by context stability or frequency. This property cannot be resolved in any instance in which two measurements are combined to produce a composite score.

4.5 Rationale to focus on the development of an objective context stability metric rather than using the BFCS approach

The narrative among discussions on methods to measure habit (e.g., Gardner, 2015; Hagger, 2019; Rebar et al., 2018; Wood & Neal, 2009) clearly mark the multiplication of behaviour frequency and context as a composite score which alone constitutes a 'habit measure'. This is the approach that has typically been used, either to predict future behaviour directly (e.g. Danner et al., 2008; Hoo et al., 2019), or to represent habit as a mediator (e.g. Galla & Duckworth, 2015), or more frequently as a moderator (e.g. Adriaanse et al., 2011; Danner et al., 2008; Friedrichsmeier et al., 2013; Ji & Wood, 2007; Neal et al., 2013; Sheeran & Conner, 2019; Wood et al., 2005) of other constructs in the prediction of behaviour. In all of these studies, BFCS has been entered into statistical models as a single variable using the multiplicative composite score of frequency and context, and described as a

representation of 'habit strength'[3]. This approach has been recommended by Hoo et al., (2019) as a promising avenue for future methods of measuring habit without the need for self-reported input.

However, the issues discussed in this chapter highlight that more targeted research is needed to understand the mechanisms by which both frequency and context interrelate with one another to result in habitual and/or maintained behaviour. This knowledge is required to meaningfully interpret results from studies using multiplicative composite scores of frequency and context. Whilst not directly acknowledged in relation to BFCS measures, these questions align with priorities that have been identified in the wider field of habit research. In a recent consensus exercise, habit experts called for a deeper understanding of the importance of frequency and context as two independent factors which may influence the degree of habitual automaticity and maintenance of behaviour Gardner et al., (2021). Whilst none of the specified research questions in this publication explicitly pertain to the measurement of habit, a number do require repetition (i.e., frequency within a specified timeframe) and context stability to be measured independently, sometimes alongside other constructs, and explored in terms of their importance for habitual automaticity and maintenance of behaviour (see Gardner et al., 2021).

Existing studies examining these questions in the acquisition of habits are few (e.g. see Kaushal & Rhodes, 2015; Schnauber-Stockmann & Naab, 2018), but some studies have explored the predictive utility of frequency, context and the interaction between them in statistical models predicting future, ongoing behaviours (e.g. Norman & Cooper, 2011; Ouellette & Wood, 1998). For at least two decades, researchers have discussed the utility of frequency of 'past behaviour' as a construct in explanatory models of behaviour, particularly in relation to habit; whilst the predictive utility of past behaviour has been reliably demonstrated (Gardner et al., 2012; Labrecque & Wood, 2015; Verplanken & Orbell, 2003), Ajzen (2011, 2016) argues that the effects of past behaviour may be a combination of: i) residual variance not explained by other constructs specified in predictive models; and ii) a measurement artefact.

Studies that explore the moderation effect by entering past behaviour, context stability and the interaction term in analyses to predict future behaviour are important for the first of these concerns, as they can demonstrate whether some aspect of past behaviour is explained by degree of context stability. There is some support for this moderation effect. In their primary research study, Ouellette

_____

[3] The recent literature critiquing existing measures of habit has been explicit that FiC is not a measure of habit strength, rather a measure of the likelihood that habit has formed because it cannot capture the degree of automaticity with which behaviour is initiated on encountering the associated context (e.g., Gardner, 2015): *"Frequency-in-context measures estimate habit indirectly, based on the likelihood that habit has formed in conducive conditions"* (Rebar, Gardner, Rhodes & Verplanken, 2018; p.38)

& Wood, (1998) conducted a regression analysis and found the interaction between past behaviour and context stability significantly predicted future recycling behaviour and television watching, after adjusting for effects of other key psychological constructs implicated in the prediction of behaviour (although p-values are not reported for any other variables). In a study of breast self-examination (BSE), Norman & Cooper, (2011) also found that only the interaction between past behaviour and frequency was a significant predictor of future BSE, even after adjusting for other constructs including habitual automaticity (measured on the SRHI).

Other studies have explored the relationship between context stability and concurrent, ongoing behaviours (e.g. Dean et al., 2021; Maher et al., 2021; Phillips et al., 2021; Pimm et al., 2015). Whilst not as robust as predicting behaviour in the future in terms of causality, these cross-sectional studies can be informative as they can describe variables which may be important in perpetuating long-term, real-world behaviours; in the maintenance stage of behaviour with no additional intervention, these studies offer some insight into the explanatory value of context stability in maintenance of behaviour. In their survey study, Dean et al., (2021) classified a range of behaviours which could help to save water as 'contextually stable' or not, and found that those classified as contextually stable were predictive of participants stating that they maintained that behaviour. In another self-reported survey study of physical activity behaviours, Pimm et al., (2015) found that cues associated with time and people were associated with the amount of self-reported physical activity behaviour. This type of study has therefore suggested that at least some of the variance in ongoing, existing behaviours can be explained by the context stability of the behaviour.

Together these studies indicate that context stability can explain to some degree the maintenance of ongoing behaviour. However, these studies do little to contend the second of Ajzen's (2011, 2016) concerns. The measurement artefact argument states that the way participants are asked to report on their past and future behaviour tends to be more similar than the way other constructs are measured. Furthermore, the self-report nature of the studies so far discussed are prone to recall bias as they ask generally about behaviour, sometimes over extended timelines of multiple years (e.g. Dean et al., 2021). This concern, whilst valid for critiquing much of the existing literature, can be overcome with the use of objective measures of both behaviour and context.

Objective measures have been introduced in a number of more recent studies to explore the role of context stability in maintained behaviours. Phillips et al., (2021) utilised objective measures of medication adherence behaviour and temporal context stability, captured using electronic pill bottles. They found that variability in context did not explain differences in adherence to pills taken in the morning compared to those taken in the evening, but the authors did not report the extent to

which context stability explained variability in adherence behaviour, generally. Rand et al., (2020) found higher attendance at the gym in the fourth quarter after joining was associated with higher context stability (i.e., attending at the same place and time, captured using electronic swipe cards on entry to the gym) in the third month after joining. As exemplified by these studies, there is some potential to capture 'context stability' using objective measures, but methods are not uniform and there is no current consensus on the best method. Studies which have explicitly set out to identify which measures of context stability associate most highly with habitual automaticity and/or behaviour have only been able to do so with self-report measures of context (Tappe & Glanz, 2013; Wohn & Ahmadi, 2019) or a mixture of objective and self-report measures (Maher et al., 2021).

Therefore, a systematic exploration of methods to measure context stability is timely and is necessary before further work can be conducted to pursue the BFCS and Hoo et al., (2019) approach.

4.6 Conclusion and implications for the following thesis chapters

The objective of this Chapter was to critique previous implementations of the BFCS measure to understand how this approach could be implemented with objective data (Thesis Aim 2, Objective 1). This chapter has highlighted several conceptual limitations with the BFCS measure and the need to develop a stand-alone objective metric of context stability.

In developing this metric, it is useful to first demonstrate its utility in explaining maintained behaviours; only among a group of people in the maintenance phase of their behaviour can we expect some people to have established a habit and therefore be able to detect whether context stability is being captured, and understand the extent of its predictive utility.

The following chapters will conduct a systematic exploration on how to measure context stability with lessons learned from the discussion in the present chapter. An objective measure of context stability must:

1) Not be conflated with frequency of behaviour;
2) Explain variability in concurrent, ongoing behaviour;
3) Explain the predictive ability of past behaviour on future, ongoing behaviour.

Chapter 5. The ACtiF Programme and randomised controlled trial of the CFHealthHub intervention: a case study for the exploration of context stability measurement in the maintenance of medication adherence

5.1 Chapter overview

Chapter 5 presents the dataset used to explore the remaining thesis aims and objectives. The data was derived from the 'Adherence to nebulised medications in adults with Cystic Fibrosis' (ACtiF) randomised controlled trial. This was part of a programme of research to develop the CFHealthHub intervention, which aimed to improve adherence in people living with cystic fibrosis (CF). An overview of CF disease characteristics, management and the ACtiF programme is given in this Chapter.

5.2 Background to the ACtiF programme: Cystic Fibrosis and nebulised medication adherence

5.2.1 Cystic fibrosis

Cystic Fibrosis (CF) is an autosomal recessive disease caused by mutations to the Cystic Fibrosis Transmembrane conductance Regulator (CFTR) gene (Riordan et al., 1989; Rommens et al., 1989; Tsui et al., 1985). In 2021, 10,908 people with CF were registered in the UK, with a median age of 21. Median predicted survival age for people with CF born in the UK between 2017 and 2021 is 53.3 years of age (Cyctic Fibrosis Trust, 2022), compared with 50.6 for those born between 2016 and 2020.

The prevalence of CF is likely to increase in line with increasing predicted survival for people with CF, with some estimates indicating an increase of 75% from 2015 to 2025 (Burgel et al., 2015).

5.2.2 The CF phenotype

Typical clinical presentations of CF include lung disease with chronic bacterial infection, pancreatic insufficiency, diabetes, liver disease and congenital absence of the vas deferens (O'Sullivan & Freedman, 2009; Ratjen et al., 2015). Clinical manifestation of CF characteristics and the associated

complications are cumulative and progressive (Bell et al., 2020; O'Sullivan & Freedman, 2009). Most important for many people with CF is the degeneration of lung function (Esther et al., 2019; Nguyen et al., 2014; Sturgess & Imrie, 1982).

5.2.3 Lung disease in CF

For many people with CF, progressive decline in lung function ultimately results in lung failure, with respiratory disease currently accounting for the highest proportion of morbidity in CF (Chalmers, 2020; Cyctic Fibrosis Trust, 2022; Turcios, 2020).

The known roles of CFTR dysregulation in lung disease pathogenesis include reduced surface liquids and decreased mucus transport in the airway epithelium, and increased obstruction of the airways, which in turn increases likelihood of infection (Ratjen et al., 2015). Over time, repeated infection, inflammatory responses, pulmonary exacerbations and the development of broncheactasis leads to the accumulation of tissue damage and degeneration in lung function.

5.2.4 Pulmonary exacerbations

The term pulmonary exacerbation describes an acute episodic increase in symptoms of chronic lung infection, for which a clinician decides to prescribe antibiotics (Wagener et al., 2013, 2018).

After exacerbation, 25% of individuals fail to recover to baseline lung health (Sanders et al., 2010). Exacerbations are associated with reduced health-related quality of life (Britto et al., 2002; Solem et al., 2016) and poor long term outcomes including more frequent and poorer response to subsequent exacerbations (Hoo, Wildman, et al., 2018; Sanders et al., 2017).

Bacterial, viral and fungal infections relate to increased likelihood of exacerbation (Amin et al., 2010; Nixon et al., 2001; Wark et al., 2012). Therefore, management, and in particular, prevention of infection to avoid 'rescue' treatments in response to an exacerbation, forms a key part of CF routine treatment (Wildman & Hoo, 2014).

5.2.5 CF management

Given the association between lung disease and morbidity in CF, treatment of lung disease is central to long term CF management in adult care (Elborn, 2016; Flume et al., 2007; Mogayzel et al., 2013).

5.2.5.1 'Rescue' therapy

In the event of a pulmonary exacerbation, it is recommended that intravenous (IV) antibiotics are administered (Flume et al., 2009) with some suggestion that better outcomes are observed when treated in hospital (Thornton et al., 2004).

5.2.5.2 Preventative therapy

Preventative therapies include a daily regimen of inhaled or nebulised medications. Inhaled therapies are particularly favoured because of their ability to deliver targeted medications with higher concentration to the airways, with fast action onset and minimal systemic side effects due to minimal absorption into other areas of the body (Heijerman et al., 2009; Hewer, 2012).[4] Furthermore, with the technological development of nebuliser systems, administration of drugs at home is becoming easier, faster and more efficient (Daniels et al., 2013).

A typical regimen would include a mucoactive agent to thin mucus and facilitate people with CF to clear mucus from the lungs, and is commonly prescribed alongside an antibiotic (National Institute for Health and Care Excellence, 2017). A regimen of this kind would result in three prescribed doses of medication per day in at least two sessions (i.e., morning and evening). Decline in lung function or new or worsening lung infection can result in additional nebulised medications being added to the daily regimen (National Institute for Health and Care Excellence, 2017).

With the rising age of the CF population and increasing number of recommended medications, a larger cohort of older adults with more complex CF and comorbid complications, people with CF are being prescribed increasingly complex treatment regimen. The burden of treatment is large, sometimes taking up to 2 hours a day to administer (Sawicki et al., 2009). Successful self-management of medication adherence in CF is therefore critical and researchers and clinicians continue to call for strategies to support people with CF to self-manage daily treatment regimen, especially to inhaled therapies.

5.2.6 Medication adherence in CF

Evidence of adherence to inhaled medications from clinical trials in CF suggests adherence is high; in a study by Pugatsch and colleagues (2016) average adherence was synthesised across eight trials conducted between 2008 and 2013 at one CF centre. Seven of these studies reported average adherence over 78%, including three studies with an average adherence of 100% (Pugatsch et al., 2016). Elsewhere, Lin et al (2017) found self-report inhaled antibiotic treatment adherence to be between 76-87%. These studies indicate much higher average adherence than observed in other long-term conditions but it is likely that they are examples of over-estimates, influenced by methodological and contextual factors. In the synthesis by Pugatsch and colleagues (2016) the

---

[4] There have been some recent advances in the treatment of CF which include gene and protein modulating therapies which are delivered via alternative mechanisms (Cystic Fibrosis Trust, 2023; Lee et al., 2021), however these were mostly introduced after the ACtiF programme and CFHealthHub RCT. Nebulised medications remain an integral part of current CF management (National Institute for Health and Care Excellence, 2017).

included studies were small, from only one centre, participation of individuals was not limited to one study and participants were also engaged in pharmaceutical studies examining the effects of novel drugs. It is possible that adherence rates at that centre were unrepresentative of other centres, and participants in trials were motivated to engage in novel treatments with the expectation of benefits which could relieve them of their complex daily regimen. In addition, this synthesis found that longer follow-ups were associated with significant decline in adherence.

Furthermore, self-reported measures of adherence have been commonly used in this literature despite being poorly calibrated with actual adherence. In their seminal study of 105 people with CF, Daniels and colleagues (2011) compared electronic ('objective') adherence measures with self-reported and clinician assessed ('subjective') adherence for 78 patients. Average objective adherence was 36%, compared with 80% and 50-60% for self-report and clinician assessed adherence, respectively. It is important to note that 25 patients that were using electronic devices to enable automatic adherence data capture, failed to bring their device to clinic to enable download of the data and therefore this figure may overestimate adherence. Other measures of adherence from observational studies, including medicines possession ratios, self-reports, pharmacy refills and electronic data capture have also demonstrated much lower adherence than that reported in CF clinical trials, averaging adherence around 50% for key CF medications, and consistently below optimal adherence of ≥80% (Burrows et al., 2002; Eakin et al., 2011; Latchford et al., 2009; Modi et al., 2006; Quittner et al., 2014).

5.2.7 Interventions to improve adherence in CF

CF is similar in profile to other LTCs in both rates of non-adherence and in lack of existing, effective interventions to improve adherence. To date, medication adherence interventions in CF have largely demonstrated little or no improvement (Bishay & Sawicki, 2016; Quittner et al., 2019). Whilst popular models of health and behaviour change have been discussed in the context of CF (Bernard & Cohen, 2004), few interventions have succeeded in adequately applying theory to intervention development to improve adherence in people with CF, or using objective, sensitive measures of adherence.

In summary, effective preventative treatments for CF exist, many of which are inhaled therapies delivered by a nebuliser. Studies measuring adherence with sensitive, objective tools in routine care have demonstrated that adherence to medication in CF is low, demonstrating adherence levels comparable with other long-term conditions (World Health Organisation, 2003). Finally, interventions targeting adherence in CF have demonstrated little success in improving adherence outcomes, with minimal application of theory to intervention design.

5.3 Overview of the ACtiF Programme

The ACtiF Programme was initiated in 2014, with the aim of developing a theory and evidence-based intervention (Michie & Prestwich, 2010) to improve Adherence to nebulised medications in adults with Cystic Fibrosis (ACtiF). This was a large, UK based programme of research consisting of three, overarching work-packages (WPs).

5.3.1 Early ACtiF WPs

WP 1 primarily focused on the development of a mechanism to objectively measure adherence data using eTrack devices (PARI GmbH, Germany) and display this on a digital platform. The eTrack is a nebuliser device fitted with a Bluetooth chip, enabling time and date-stamped data capture and transfer. The digital platform, 'CFHealthHub', displayed daily adherence using data from the eTrack, alongside an array of educational and self-management resources. The website was developed with two interfaces; one was tailored for people with CF and was accessible via a website and mobile app. The second was for healthcare professionals who were trained to deliver a behaviour change intervention complementary to the CFHealthHub platform.

WP 2 applied the Behaviour Change Wheel framework (Michie et al., 2014; Michie et al., 2011) and COM-B (Michie et al., 2011) model to develop the behaviour change intervention. The intervention incorporated the adherence displays and self-management resources made available via CFHealthHub. To understand how different aspects of COM-B (Michie et al., 2011) relate to adherence in CF, qualitative studies were conducted in WP 2 (Arden et al., 2019; Drabble et al., 2019).

The conceptual framework developed from this work is depicted in Figure 5 (recreated from the trial protocol (Wildman et al., 2018).
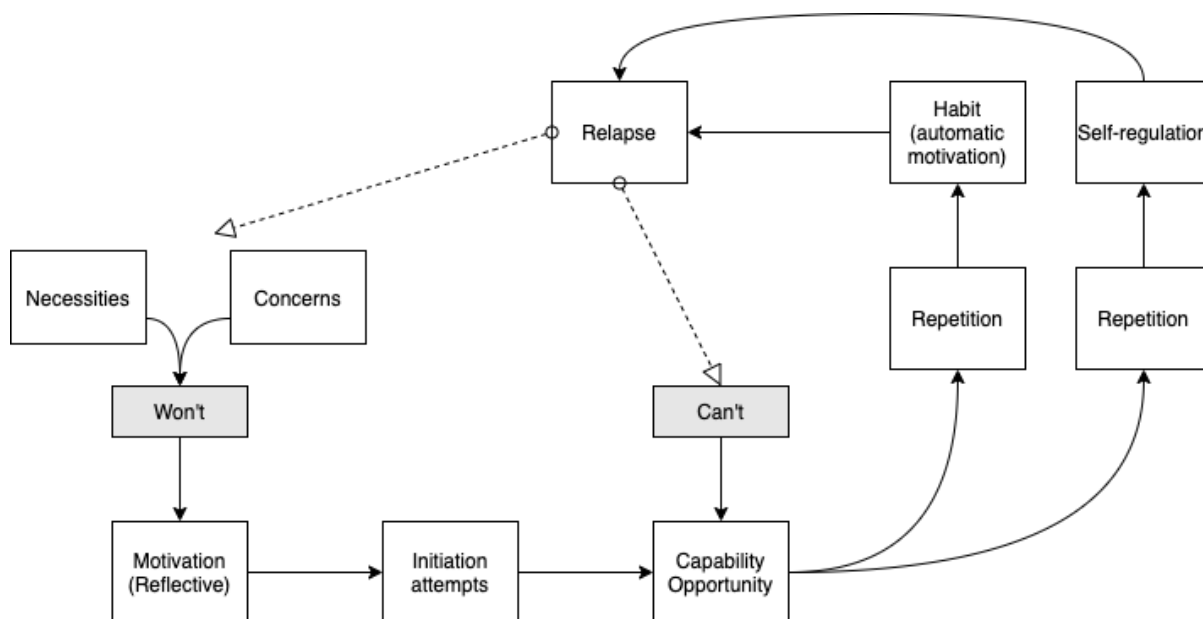
Figure 5
The conceptual framework used to develop the CFHealthHub intervention, recreated from Wildman et al (2018).

Constructs and processes from Social Cognitive Theory (SCT; Bandura, 1991), Control Theory (Carver & Scheier, 1982) and habit theory (Gardner, 2015) combine in this model, to explain the process of maintained behaviour change in medication adherence. First, perceived necessities and concerns (Horne & Weinman, 1999) for the behaviour converge to form outcome expectancies. Even when favourable outcome expectancies (motivation) are held, SCT posits that the individual must also have sufficient perceived self-efficacy to translate motivated intentions into action (capability, opportunity). Specifics of the relevant beliefs surrounding the COM-B constructs are described in detail in Arden and colleagues (Arden et al., 2019) and Drabble and colleagues' (2019) publications. In the case of sufficient motivation, opportunity and capability the individual may begin attempts to initiate the behaviour. Control Theory conceptualises this as an effortful process of goal monitoring, in which the individual observes the discrepancy between their current behaviour and goals, and through feedback loops, regulates their behaviour to meet their goals. For example, if an individual observes their current adherence is 50% because they have not taken their second treatment of the day, they can initiate their second treatment to meet their goal of 100%. Finally, habit theory was incorporated in the model as a mechanism by which medication taking behaviour can be maintained over time. In this context, continued repetition of medication-taking behaviour in a specific context will shift the behaviour from a process of effortful self-regulation to a process which is automatically initiated by contextual cues, following cue-response association formation during repetition.

A full report of the development of the CFHealthHub intervention, along with a TiDIER description, is available elsewhere (Arden et al., 2021). The mechanisms described in this model formed the foundation for the CFHealthHub Intervention, delivered in the full-scale RCT in WP 3.

5.3.2 The ACtiF full-scale RCT (WP3)

WP 3 consisted of a pilot and external, full-scale RCT, evaluating the effectiveness of the CFHealthHub Intervention. The trial was a two-armed parallel-group RCT conducted between October 2017 and June 2019. This was a multicentre trial of 19 UK CF centres, including sites in England, Scotland, Northern Ireland and Wales. A total of 608 participants were randomised to receive the CFHealthHub intervention (n=305), or usual care (n=303).

Intervention arm participants received the CFHealthHub Intervention. Participants were first issued with an eTrack (PARI GmbH, Germany) and access to the CFHealthHub web and app platform. Medication-taking data from the eTrack was combined with prescription data entered on the CFHealthHub interventionist portal, to calculate adherence. Participants could access this alongside a range of resources designed to deliver selected behaviour change techniques, in response to the problem conceptualisation and behavioural analysis described in WP 2. Participants were supported to use CFHealthHub by trained interventionists, who delivered a complementary, manualised behaviour change intervention in face-to-face or telephone intervention sessions. A minimum of six sessions (a phase) was scheduled to be completed over the 12-month follow-up period. Subsequent phases could be initiated if a participant: requested additional support, demonstrated a 20% reduction in adherence within four weeks or was prescribed intravenous antibiotics for a pulmonary exacerbation.

Usual care participants also received an eTrack with time and date-stamped recording of inhalations enabled, but adherence data was not made visible to participants or healthcare professionals treating participants in this arm. No intervention sessions were delivered to this arm.

Participants were followed up for 12 months (+/- 1 month) from the date of consent, and inhalation data collection continued beyond this for all participants still consented into the trial until June 2018.

Full reports of the ACtiF programme (Wildman et al., 2021), pilot RCT (Hind et al., 2019), full-scale RCT protocol (Wildman et al., 2018) are available elsewhere.

5.3.3 Sample characteristics

Baseline characteristics of participants involved in the ACtiF full-scale RCT are summarised in Table 14. A total of 608 participants were randomised and 607 went on to contribute data to the trial dataset.

Table 14

ACtiF trial full sample baseline characteristics.

Baseline Characteristic	Mean (SD)	Median [LQ, UQ]	n(%)
Intervention arm			305 (50.2%)
Age	30.7 (10.7)	28.1 [22.6, 36.1]	
Female			310 (51.1%)
IV days (home)	13.2 (21.2)	0 [0, 19]	
IV days (hospital)	12.7 (21.6)	0 [0,15.5]	
FEV1%	59.5 (23.1)	59.1 [40.0, 77.9]	
IMD quintile	3.04 (1.36)	3 [2,4]	
Number of prescribed doses	3.26 (1.60)	3 [2, 4.29]	
Max days in trial	463.8 (123.2)	486 [430.5, 538]	
CFHH Pseudomonas status: chronic			349 (57.9%)

IV= intravenous antibiotics; FEV1%= forced expiratory volume in first second (percent); IMD quintile= Index of Multiple Deprivation quintiles; CFHH Pseudomonas status: chronic=diagnostic criteria fulfilled for chronic pseudomonas status using the CFHealthHub criteria tool (Zhe Hui Hoo, Coates, et al., 2018).

5.4 Summary

Chapter 4 defined and described BFCS measures and their desirable properties and limitations. The chapter concluded that habit research could leverage time-stamped adherence data to develop an objective measure of context stability. Chapter 5 described the ACtiF programme, including the time-stamped adherence data collected during the RCT of the CFHealthHub intervention.

The following chapters will derive objective measures of context stability from the CFHealthHub RCT time-stamped adherence data to address Thesis Aim 2, which was to explore whether the time of day at which people take their medications can be used to derive an objective metric of context stability for use in habit research.

Subsets of the full trial dataset are used in the exploration of Aim 2 in the subsequent chapters. Rationale for subsample selection is justified in each study. Characteristics of subsamples are summarised in each chapter and direct comparisons of each with the primary dataset are available in Appendix F.

Chapter 6 Derivation and exploration of variables to describe context stability in objective medication adherence data

6.1 Chapter Overview

Chapter 6 begins the exploration of thesis Aim 2, Objective 2, which was to present a systematic approach to identifying and selecting a metric of context stability, derived from objective data on the times of day at which people initiate medication adherence behaviour. The specific objectives of Chapter 6 were:

1) To describe a range of methods that could be used to produce objective variables that describe context stability, with theoretical improvements on previous measures (Thesis Aim 2, Objective 2.1);

2) To verify the interpretability of these variables in terms of their theoretical relationship with concurrent medication adherence behaviour (Thesis Aim 2, Objective 2.2).

To address these objectives, Chapter 6 presents novel methods to derive variables that can describe context stability from objective medication adherence behaviour using time-stamped behavioural data. These methods were selected to improve upon the issues and to address the desirable properties of a metric of context stability that were identified in Chapter 4. The novel methods are presented alongside other metrics that have been used in the literature and a preliminary exploration and comparison of their relationships with behaviour (i.e., the relationship with concurrent medication adherence behaviour) is presented.

6.2 Background

There is no commonly agreed method or measure to describe context stability; as with 'past', or 'future behaviour', is it a theoretical construct that has been defined and measured in different ways in the literature. This is true of both self-report measures of context stability, and those derived using objective data.

Focusing on the latter in line with the topic of this thesis, a range of examples have emerged in the recent literature; these are mainly in the domains of physical activity and medication adherence, and

all utilise technological devices to automatically capture timestamps of the onset of a behaviour (i.e., temporal context stability), and use this data to construct a context stability score. One study calculated a context stability score to describe consistency in the initiation of physical activity (Rand et al., 2020); when participants scanned a pass to enter their gym within a 3-hour window on the same day of the week , this counted toward an aggregated context stability score; doing this twice in a month resulted in a score of 1 and scores could increase up to 3 for repeating this more frequently. Another study recorded whether physical activity occurred within a four-hour period (i.e. 8am-11:59am; 12noon-3:59pm; 4pm-8pm) and indexed context stability by taking the entropy of frequencies of occurrence across the three temporal windows (i.e. the degree to which the probability of behaviour occurring in the morning afternoon and evening, were equal, accounting for the number of options; Maher et al., 2021). In the domain of medication adherence, one study aggregated time-stamped medication adherence data to calculate the variance in times that the behaviour was initiated over the full study period (Phillips et al., 2021), while another used the standard deviation, aggregated into consecutive 7-day periods over the study period and combined this with percentage adherence to construct a behavioural frequency x context stability (BFCS) measure (Hoo et al., 2019).

The diversity described in these studies demonstrates that there is yet to be a consistent approach to the measurement of context stability using objective temporal data, and each of the existing methods has some limitations. As discussed in Chapter 4, objective metrics of context stability have often been conflated with frequency of behaviour by failing to account for the number of observations used to calculate it (e.g. Hoo et al., 2019; Maher et al., 2021; Phillips et al., 2021) and/or by actively including information about frequency in the calculation (e.g. Hoo et al., 2019; Rand et al., 2020). Doing so may result in over-estimation of the relationship between context stability and behaviour. Furthermore, the context stability measures described do little to account for day-to-day changes in routines. This is important when considering researchers' expectations about what a context stability variable derived from temporal data alone can capture. In theory, any feature of a performance environment could constitute a cue for a habit, even abstract 'environmental' cues such as feelings (e.g. Verplanken, 2006; see also discussion by Mazar & Wood, 2019, p.19). To date, no research exists that can definitively and entirely identify what constitutes a cue to behaviour once a habit is established (see discussion in Chapter 1). This difficulty is compounded with evidence that suggests people are often poor at identifying cues for their own habitual behaviour (e.g. Gantman et al., 2017). In the absence of this knowledge, or vastly more expensive and perhaps intrusive objective methods to measure more completely the aspects of individuals' performance contexts, the stance of the present thesis is that current focus should be on

maximising the predictive ability of measures derived from temporal data. This requires acknowledgement that the time of day is not the cue per se, rather some other cue which likely coincides with the same or similar times of day with each repetition.

In the context of medication adherence behaviour, two studies have attempted to understand context stability of behaviour as a predictor of behaviour frequency, using information about the time of day collected on electronic medication devices (Hoo et al., 2019; Phillips et al., 2021). In both studies, 'context stability' was inferred from variability in the times at which medication-taking behaviours were performed. In a sample of 123 participants, Hoo et al., (2019) found a combined measure of frequency and context stability correlated with future behaviour (r=0.4) but a one unit change (on a scale of 1 to 7) in this measure was associated with only a 0.3% increase in adherence. Similarly, Phillips et al., (2021) found that differences in adherence between morning and evening sessions of medication taking were not reflected in the context stability of dosage timings. The extent to which variability in times that medication-taking behaviours occur is a suitable proxy for context stability is unclear from these studies and must be met with careful consideration of the limitations of methods used, as discussed in the previous paragraph.

More advanced methods were applied by Stecher et al., (2021) who created a measure of context stability to describe day-to-day consistency in the use of an app. Dynamic time warping (DTW) was applied to describe the similarity in *length* of use among participants using a meditation app. DTW compares the similarity between two time-series (e.g. engagement with an app on day 1 – i.e. time-series 1, and day 2- i.e. time-series 2); with some constraints, the DTW algorithm compares values in time-series 1 to the *most similar* values it can find in time-series 2. Compared to more standard statistics, this method could be advantageous to habit researchers as it can produce a quantitative summary of stability but with some flexibility around what is considered 'stable'; observations do not need to occur at exactly the same time to be considered the 'same' context; i.e. a difference of a few minutes from one day to the next would not affect the context stability score. By allowing similar behaviour repeated within 60 minutes of another to be captured as repetitions of behaviour in the same context, a degree of flexibility was embedded into the definition of 'contextually stable' behaviour.

Figure 6, taken from Stecher et al.,'s (2021) article, demonstrates an instance in which this might be useful: the plots are a binary visualisation of times during the day in which the app was in use (y=1) or not in use (y=0); the figure demonstrates two ways to compare the time series on day 1 (TS1; top) to the time series on day 2 (TS2; bottom), either using Euclidean distance (left) or DTW (right); Euclidean distance is a simple calculation of the distance between TS1 and TS2 using 1:1 mapping

between observations, which here results in a 'distance' of 2. However, visual inspection of the plots clearly shows the same overall pattern of using the app for 2 units in both TS1 and TS2; DTW seeks the most similar points along the time series and calculates the difference between these, which in this instance results in a distance score of 0. In their implementation, Stecher et al., (2021) applied a constraint such that two points can only be matched within 60 minutes of one another.
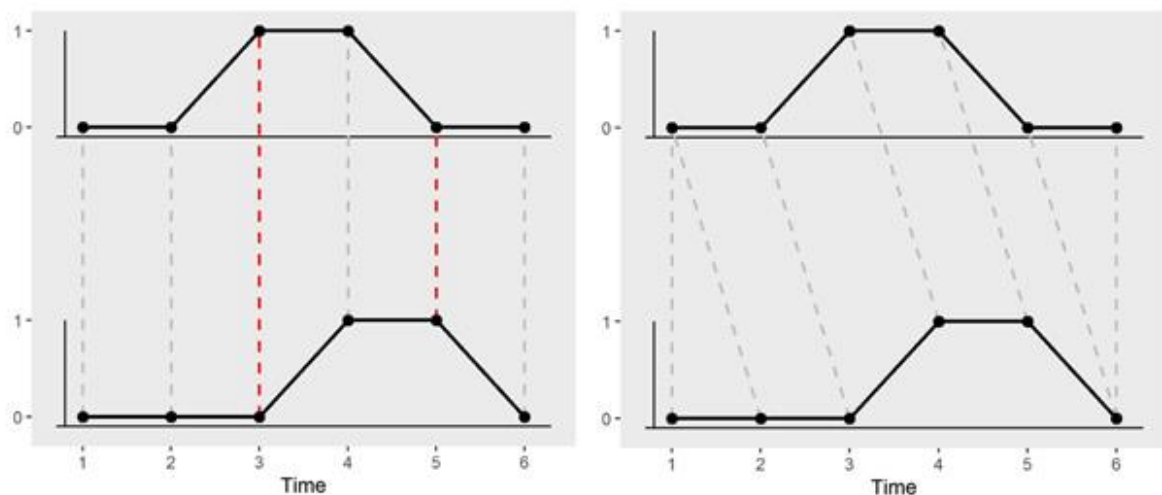
Figure 6
Visualisation of how the DTW algorithm calculates distance (i.e. numerical difference) between 2 time-series. The top row shows timeseries 1 and bottom row shows time series 2. Vertical scale on each plot: 0= app not in use, 1 = app in use. Plots on the left of the panel represent how Euclidean distance is calculated (i.e. lines are drawn perpendicular to the x-axis and with 1:1 mapping between points along the time-series); plots on the right demonstrate how DTW distances are calculated (i.e. by findings the smallest numerical difference between points on each time series, with some set rules). (Plots taken from Stecher et al., 2021, pg. 4, Figure 1)

By applying DTW to their data, Stecher et al., (2021) were able to capture the day-to-day similarity between the *length and frequency* of app use from one day to the next. In considering this as a measure of context stability, this has approach also has several limitations which relate to expectations of human behaviour in the real-world and recent empirical findings and their meaning for habit theory.

First, recent evidence suggests that measures of context stability comparing similarities from day-to-day are weaker in the prediction of behaviour vs week-to-week comparisons (Phillips et al., 2022); this could be due to peoples' routines revolving on weekly cycles. For example, for many people working Monday to Friday, the time at which they eat breakfast on Saturday and Sunday is likely to be different than during the week, but this difference is consistent week-to-week. Comparing the whole of week one to the whole of week two may enable detection of similar behavioural patterns more readily than comparing a weekday to a weekend day (e.g., Friday to Saturday). Second, recent

theory dissociates different types of habit which may exist in real-world behaviours. Theorists have distinguished between habits of instigation vs habits of execution (Gardner et al., 2016; Phillips & Mullan, 2022), whereby habitual instigation refers to habitually selecting action and starting to perform a behaviour, whereas habitual execution refers to habitually carrying out each of the sub-actions required to complete the behaviour. Empirical evidence has demonstrated that instigation is a better predictor of frequency of behaviour, including medication adherence (Gardner et al., 2016; Hoo, Gardner, et al., 2019; Phillips & Gardner, 2016) and has led researchers to focus attention on this aspect of habitual behaviour (Phillips & Mullan, 2022). DTW, as implemented by Stecher et al., (2021) is primarily capturing similarities in execution, as the calculation of distance includes the length of time for which the behaviour was performed. Therefore, Stecher and colleagues' approach to the measurement of context stability is suboptimal when research is interested in habitual instigation of behaviour.

As exemplified by this critique, existing measures of context stability using time-stamped data are diverse and each has their own limitations. Chapter 4 highlighted that lack of attention to the properties of variables caused by the ways in which they are measured can lead to issues with interpretation and predictive validity of context stability. These issues are important to consider in the exploration of novel methods to summarise objective behavioural data. The present chapter begins the process of systematically exploring metrics to summarise context stability using objective behavioural data, with the following aims:

1) To describe a range of methods that could be used to produce variables which describe context stability, with theoretical improvements on previous measures (Thesis Aim 2, Objective 2.1);
2) To verify the interpretability of these variables in terms of their theoretical relationship with concurrent medication adherence behaviour (Thesis Aim 2, Objective 2.2).

6.3 Methods

6.3.1 Raw data preparation

The eTrack™ records the time at which a participant administers a new medication. In order to process this data, timestamps were first converted into the number of minutes into the day, with each day starting just after midnight or 0:00:00, that each medication was taken (e.g., a medication administered at 07:00:30 would be represented as 420.5 minutes).

Second, the time at which each day started was shifted. Previous studies examining the times at which medications were taken in CF have utilised a 05:00:00-04:59:59 clock scale to distinguish one 'day' from the next (Zhe Hui Hoo, Wildman, et al., 2019). This was following the observation that

many people with CF take medications very late at night, spanning across the threshold of midnight, and there is evidence of this behaviour in this sample. An example is given in Figure 7 for a participant that was prescribed more than one dose a day; focusing only on the inhalations labelled in blue, it is clear that whilst all inhalations in blue are labelled as the first observed inhalation on each day, the inhalations recorded between 00:00 and 02:00am are more similar in time to inhalations recorded around 23:00, which are mostly orange and green and labelled as second and third inhalations.
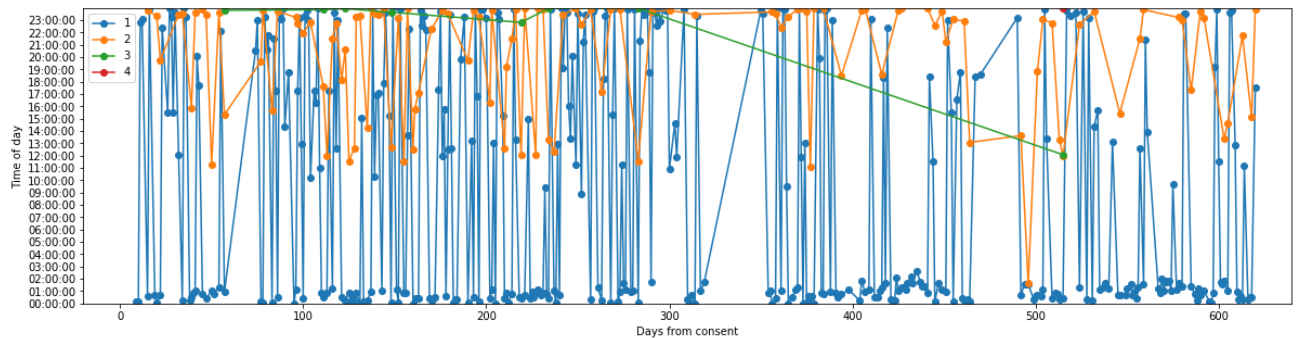
Figure 7
The times at which a participant took each recorded inhalation over the duration of their participation in the study. The first inhalation on any given day using a 00:00-23:59 clock is plotted in blue.

These were more likely late, last inhalations on a given day, rather than very early, first inhalations on the following day. Given this observation, the 'end' of each day was shifted from 23:59:59 to 04:59:59 to avoid artificially inflated measures of variation in observed times at which medications were taken. Once this transformation was applied, a behaviour observed at 07:00:30am was represented as 120.5 minutes from the start of the day (i.e., 5am).

6.3.2 Feature generation

A combination of simple descriptive statistics and a set of more complex variables were derived from the times at which medications were taken to produce a set of potential 'context stability' variables.

6.3.2.1 Simple descriptive statistics

The following variables were calculated: range of the times of day that inhalations taken (RANGE); inter-quartile range of the times of day that medications were taken (IQR); standard deviation of the times of day that medications were taken (SD); entropy (i.e. equality of probabilities among the possible outcomes; to define 'possible outcomes' as a discrete variable, each day was divided into 72 non-overlapping bins of 20 minutes and observations categorised into respective bins; ENT); entropy in the differences in time of day between consecutive inhalations (DIFF_ENT); average of the differences in time of day between consecutive inhalations (DIFF_AVG); standard deviation of the

differences in time of day between consecutive inhalations (DIFF_SD). Entropy variables were calculated by applying the SciPy Statistics 'Entropy' method to frequency distributions of observations in 20-minute bins. All variables were measured in units of minutes.

6.3.2.2 Context stability variables derived from Recurrence Quantification Analysis

Recurrence Quantification Analysis (RQA) is a non-linear method of quantifying recurring patterns in longitudinal data without *a priori* knowledge of any underlying repeating structure. When applied to the present dataset, a recurrence (i.e., a behaviour repeated in the same temporal context) would be identified if the time at which a medication session was initiated on a given day was repeated within some pre-specified window of time on the following days.

RQA has recently emerged as a potentially useful technique in the study of longitudinal behaviour change (Heino et al., 2021). Applied through the lens of habit theory, RQA could enable a rich description of the context stability of medication-taking behaviours with objective, longitudinal data.
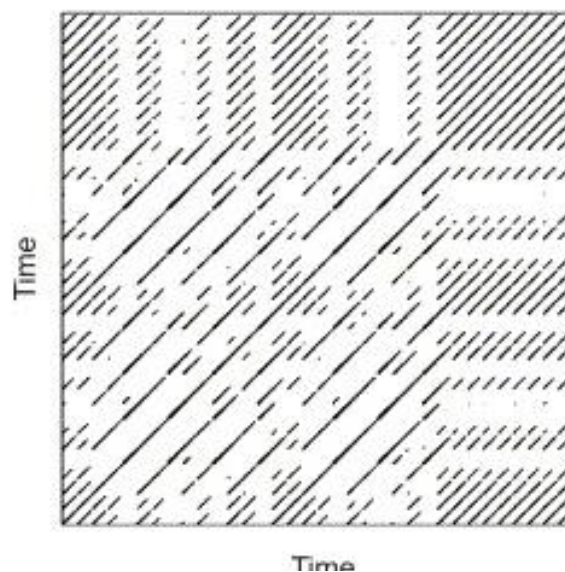
Figure 8
Example of a recurrence plot taken from Marwan et al., (2007), Figure 1(b), pg 242. For each value along the x-axis that recurs within a pre-specified radius elsewhere in the time series, a point is plotted in the matrix.

The number of minutes into the day at which a medication was taken was identified for each participant, over a given number of days (i.e. 14 or 28 days), and plotted along the x and y axis in a two-dimensional matrix known as the Recurrence Plot (RP; see example in Figure 8 taken from Marwan et al., (2007), pg. 242). The solid diagonal line on the RP is known as the line of identity and this represents the intersection of the same observation on each axis (i.e., the same observation in the sequence of longitudinal data sampled).

One RP was produced for each participant. For each observation along the x-axis of the RP, the recorded value (i.e., number of minutes into the day that an inhalation occurred) was compared against all other values along the y-axis (i.e., the number of minutes into the day that inhalations occurred on the other days sampled). A 'recurrence' was recorded if an observation on the y-axis was found to have a value within a pre-specified radius (e.g., within 60 minutes) of the reference observation. For each identified recurrence, a point was plotted at the intersection on the RP; for values not within the pre-specified radius, i.e., a non-recurrence, the intersection was left blank. This resulted in a binary visualisation of the recurrence dynamics for each participant's behaviour. An example of a recurrence plot for a 14-day period with a recurrence radius threshold of 60 minutes is given in Figure 9.
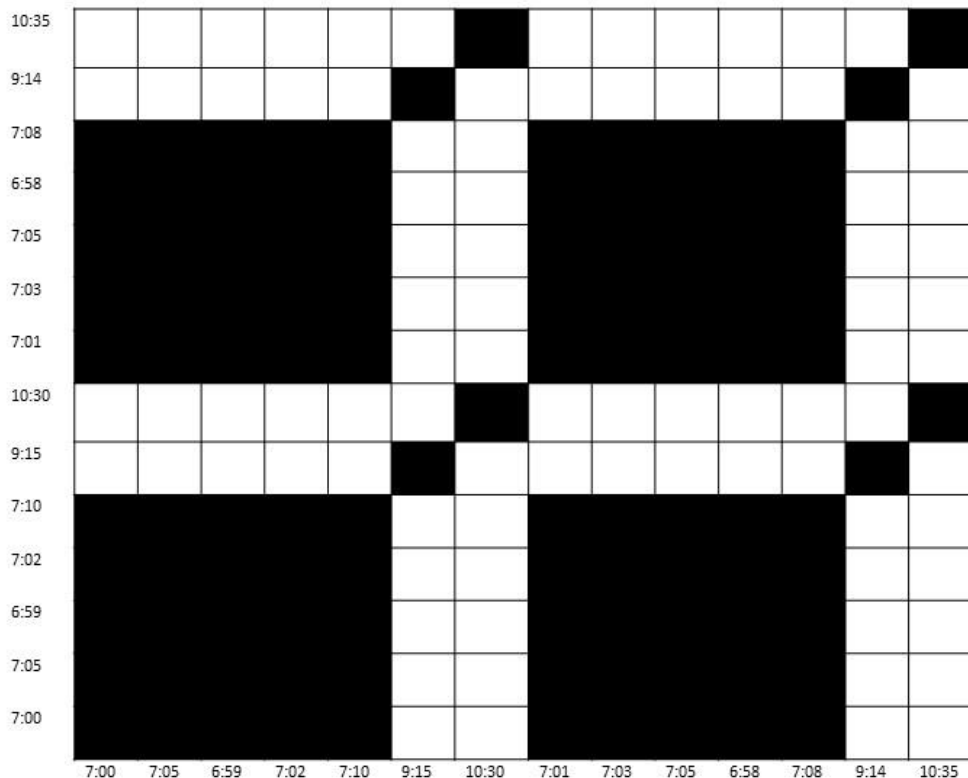
Figure 9
Example of a RP for a 14-day period; recurrence threshold was set to 60minutes for this plot. For each observation, the time at which it occurred is plotted along the x-axis and again along the y-axis. For each x-axis value, all values of the y-axis are examined to identify any observation within a 60-minute radius. If an observation is within 60 minutes of the reference observation, a point (i.e., a solid cell) is plotted at the intersection; otherwise, the intersection is left blank.

RQA techniques can produce a range of quantitative variables that capture the recurrence dynamics of behaviour (Weber & Zbilut, 2005). These describe the vertical and diagonal structures of the RP

and all depend on hyperparameters which are described in the subsequent section. Eight RQA variables were calculated with the potential to represent context stability:

- **Recurrence rate (RR):** the proportion (%) of recurrence points on the RP. Higher RR would indicate higher proportions of medication taking behaviours which occurred in the same context. A value of 100% would imply that all selected observations for the participant were recorded within the recurrence threshold, i.e., if the threshold was set to 60 minutes and RR calculated over 14 observations, this would mean all doses of medication were taken within 60 minutes of one another.[5]

- **Determinism (DET):** the proportion (%) of recurrence points which fall on a diagonal line with one or more recurrence point. Higher DET could indicate recurrent shifts in the context that a medication was taken (e.g., a shift from 9 to 11am from Friday to Saturday, each week).

- **Laminarity (LAM):** the proportion (%) of recurrence points which fall on a vertical line with at least one other recurrence. Higher LAM indicates more instances of consecutive recurrences.

- **Trapping time (TT):** the average length of vertical lines of recurrence points which fall on a vertical line with at least one other recurrence. This is represented as an actual number therefore is related to the number of observations used to create the RP. Higher TT indicates, on average, longer streaks of consecutive recurrences.

- **Longest vertical line length (VMAX):** the length of the longest vertical line in the RP, i.e. the longest streak of recurrences. This is represented as an actual number therefore is related to the number of observations used to create the RP. Higher VMAX indicates at least one instance of a longer streak consecutive recurrences.

- **Entropy in the vertical lines (VENT):** the entropy in the lengths of vertical lines of recurrence points. Higher VENT could indicate more variability in the number of consecutive recurrences at a time.

- **Average white vertical line length (W):** the average number of consecutive non-recurrent points. This is represented as an actual number therefore is related to the number of observations used to create the RP. Higher W indicates, on average, more observations in between recurrences.

---

[5] This is essentially the same as could be achieved by implementing DTW with timeseries representing behavioural instigation.

- **Entropy in the white vertical line length (WENT):** the entropy in the lengths of vertical white lines (i.e. with no recurrent points- see W). Higher WENT indicates more variability in the number of occasions between medications taken in the same context.

RQA hyper-parameter selection

The calculation of all RQA variables depends on the pre-specification of two hyper-parameters: minimum line length and recurrence radius. Recurrence metrics are most useful as comparative tools, as opposed to raw characterizations of the data, due to their dependence upon the hyper-parameters used.

The minimum length of a 'line' of recurrence points was set to two[6]; this is common practice in the use of RQA across fields, including psychology (Heino et al., 2021). Furthermore, this had theoretical grounding in that this could capture two observations that occurred at the same time on a two-day streak such as a weekend.

Recurrence radius (rad) describes the maximum difference between two observations that constitutes a recurrence and must be pre-specified. Applied to the present dataset in which observations are represented by the number of minutes into a day, rad=10 minutes would record a recurrence point for a medication taken at 2:25 and 2:19pm, but not for a medication taken at 2:36pm as this is more than 10 minutes later.

There is currently no evidence base to guide what the radius for a 'recurrence' should be set to. As previously discussed, context stability variables derived from objective temporal data have been summarised in many ways, and most have implicit assumptions as to how close two events can happen in time to constitute a likely recurrence of a behaviour in the same context. Hoo, Wildman, et al., (2019) specified a maximum context stability score at a standard deviation of 180 minutes, whilst in the domain of physical activity, Rand et al., (2020) classed a behaviour as in the same context if it was within a three-hour (180 minute) time window and Maher et al., (2021) applied a fixed, four-hour (240 minute) time window. Phillips et al., (2021) did not apply a cap or radius for recurrence but measured variance in the times at which medications were taken; the average variance in times medications were taken in this study was approximately 6.5 to 7 minutes. This is a very small window compared to three or four hours and should be interpreted with caution as this study was conducted with a small sample (n=51) with likely very high average adherence (average of ~9 pills missed over ~28 days on which 2 pills were expected per day; approximate sample average

---

[6] Applies to DET, LAM, TT variables.

of 84% adherence). To date, there has been no systematic exploration of the temporal window of recurrence expected for behaviours performed in the same 'context'.

Rather than apply one arbitrarily selected radius, each of the selected RQA variables were calculated with 15 different radius settings (5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 70, 80 and 90 minutes). The maximum radius of 90 minutes was selected because many regimens require more than one medication session in a given day due to the requirement for more than one dose of the same medication to be taken after an interval. The minimum expected interval between these sessions is three hours and therefore to minimise the risk of capturing an early inhalation from the second session as a recurrence from the first session, the three-hour window was halved, into 90 minutes.

Imputation for RQA variable missing values

Some RQA variables cannot be calculated if the RP has no recurrence points or if it is completely saturated. This only applies to some variables, and values were imputed accordingly because in the context of this study, missing values are still meaningful. Details are presented in Table 15.[7]

Table 15
Imputed values for missing RQA variable values

Variable	Imputation
DET	Missing values were imputed with a value of 0. Minimum DET is 0 but will report missing if no recurrence points off the line of identity.
TT	Missing values were imputed with a value of 1.9. Minimum TT is 2 but will report missing if no lines (i.e., no recurrences) of length 2 or more.
W	Missing values were imputed with a value of 0.9. Minimum value was 1 but will report missing if there were no observed non-recurrences (i.e., no time between recurrence points and the recurrence plot is saturated).

In total, 127 variables were calculated, including those derived using simple, descriptive statistics, and those derived using RQA at 15 different radius settings (Appendix E). Each of these 127 variables were candidate variables for describing context stability.

---

[7] Unlike other studies in which missing adherence datapoints were imputed, imputation in this form does not conflate frequency with context stability as the same number of observations were used to calculate the RP from which the variables are derived.

6.3.3 Data-handling

6.3.3.1 Constraints applied to enable comparability between participants

The times of, and adherence to, the first observed inhalation on each day was used for all subsequent variable derivation and analyses. Regimen were frequently complex in the trial sample, with as many as 12 prescribed medications to be taken on a single day. In the example of a people with CF prescribed two doses of the same medication which should be taken at separate times on the same day, it is theoretically possible that they could have a habit for taking their first, or second medication, or both, or neither. Only 60 of 608 participants (~11%) were prescribed only one dose of medication for the duration of the study. For this subset of participants, we could be certain that if a medication was taken, it was the only prescribed dose for that day and therefore context stability for a specific behaviour could be confidently calculated from the time at which it was taken. However, whilst the use of this subset of participants could reduce a potential source of noise in the calculation of context stability, it was decided that the inclusion of more data in subsequent analyses was more beneficial than use of only 11% of the dataset. Therefore, to separate out discrete behaviours and their associated habits, and to make comparisons between individuals with different treatment regimen, the first inhalation observed on each day was used to calculate context stability.

6.3.3.2 Time-series segmentation

Time-series of the times at which the first medication was taken on each day for each participant were segmented into 12 consecutive, mutually exclusive 28-day periods, hereafter referred to as intervals.[8]

6.3.3.3 Constraints applied to delineate context stability from adherence

Among the critiqued studies that used objective data to describe context stability, the impact of using few observations in the calculation of context stability was not considered. In the instance of a participant that has only taken one medication over the period of interest, a measure of entropy or inter-quartile range would indicate extremely high context stability (IQR=0; ENT=0). In theory, low IQR and low entropy should indicate high context stability, but this interpretation cannot be usefully applied in these examples as the number of observations used to calculate the variable (i.e., frequency of the behaviour) conflates the meaning of the context stability score. Furthermore, a minimum of two observations is required to calculate a standard deviation; to overcome this issue in participants with one or no inhalations, Hoo, Wildman, et al., (2019) imputed missing values with the

---

[8] For simplicity of description in the remainder of the thesis, segments are henceforth referred to as 'intervals', i.e., interval 1 refers to the 1st 28-days since the participant's consent into the trial.

180-minute 'ceiling' score. This is similarly problematic as a context stability score of 180 minutes becomes synonymous with very low frequency of past behaviour.

RQA variables are even more sensitive to the effect of using different numbers of observations to calculate them. To make comparisons between individuals based on RQA variables, the time series used to calculate the variables must be equal in length. For example, 'W' captures the average number of consecutive non-recurrent points in the RP. Higher W indicates more occasions on which medications were taken out of context between medications taken in context, therefore high W indicates low context stability. If only the inhalations observed during a 28-day period of the study were used to construct W, a participant that only took 14 medications during that interval (i.e. 50% adherent) could have a max W=14, whereas a participant that took 28 medications (i.e. 100% adherence) could have max W=28. Neither of these participants would have taken a medication in the same context as another medication but the participant with 50% adherence would appear to have a better context stability score.

In this study, to ensure context stability was delineated from frequency of medication-taking behaviours, and to ensure comparability of context stability scores between individuals, the nearest 14 and 28 inhalations prior to interval 7 of the trial were identified and extracted to compute two sets of context stability variables. Variables were constructed using points closest to the midpoint of the trial follow-up (up to the start of interval 7) because data sampled in the early weeks of the trial were considered unlikely to reflect participants' longer-term behaviour; the white-coat effect (Podsadecki et al., 2008), associated with initial provision of chipped nebulisers, typically created a temporary increase in adherence at the start of the study which was followed by a fall (Wildman et al., 2022). The set containing 28 observations were considered the baseline set, as this could contain a minimum of four cycles of a week-long set of routines. This set was then compared to the 14-observation set, in order to determine whether similar relationships between context stability variables and behavioural frequency could be observed with a smaller number of observations from each individual. This comparison was driven by the practical implications of requiring less data in order to capture context stability, for both research and clinical purposes.

6.3.4 Calculation of concurrent behaviour frequency

To calculate 'concurrent behaviour' it was important to look at adherence over the same period used to calculate context stability variables. The numerator (i.e., number of inhalations administered) was therefore equal to 14 or 28, for respective variable sets, as these were the same 14 or 28 inhalations from which the time-of-day data was derived for calculating context stability. The denominator (i.e., number of doses prescribed) was calculated as the number of days over

which the 14 or 28 observations collected to calculate context stability variables were taken. For example, the nearest 28 inhalations prior to the start of interval 7 were extracted for participant A and presented in Figure 10.
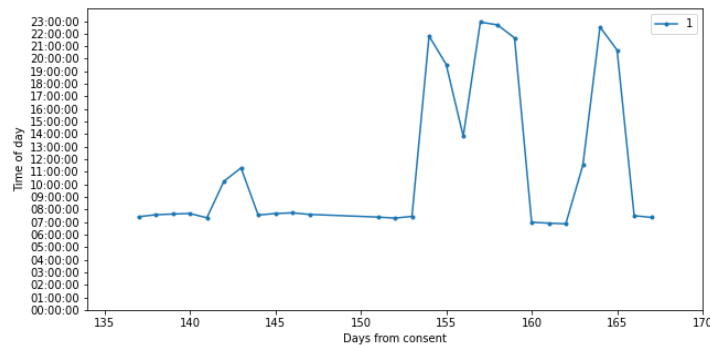
Figure 10
The 28, first inhalations recorded to closest to interval 7 for participant A (adherence to first inhalation over this period = 90.3%)

To collect 28 inhalations in total for this participant, the first observation was on day 137 and the last observation was on day 167, totalling a 30-day period. Substituting this information into the numerator adjusted adherence formula:

$$Concurrent\ adherence = \frac{28\ inhalations}{30\ days} \times 100$$

$$Concurrent\ adherence = 90.3\%$$

6.3.5 Analysis

First, Pearson's correlations between each of the 127 potential context stability variables and concurrent behaviour were compared for the sets of variables derived from 14 and 28 observations. The purpose was to determine whether similar relationships with concurrent behaviour could be observed for variables derived from 14 observations as for those derived using 28. Similarity was determined by comparing correlations, significance and the precision of estimated confidence intervals.

Interpretability of each variable was subsequently assessed by evaluating whether the hypothesised explanation of each variable matched the bivariate correlations with concurrent adherence behaviour and the range of observed values.

Variables were considered to be interpretable as a metric of context stability if:

i)   The actual and hypothesised correlation with concurrent behaviour was consistent in direction;

ii)   The range of values for each variable could be meaningfully interpreted.

Variables were calculated in Python (v 3.8.3). The Pandas library was used for simple context stability variables and PyRQA (v 8.0.0; Rawald et al., 2017) was used to derive RQA variables. Correlation estimates between context stability variables and concurrent behaviour, along with their associated p-values and confidence intervals were calculated in SPSS (v 27.0.1.0).

6.4 Results

6.4.1 Dataset overview

A total of 526 of 608 (86.5%) participants had sufficient data to calculate context stability and concurrent behaviour with at least 14 inhalations over the first 168 days of the data collection period; 473 of 608 (77.8%) participants had sufficient data to calculate context stability and concurrent behaviour with at least 28 inhalations over the first 168 days of the data collection period.

Among the 14-observation dataset, the minimum adherence to medication was 8.6% (median concurrent behaviour=82.4%; [LQ,UQ]= [44.1,100]%; Figure 11). For the 28-observation dataset, the minimum adherence to medication over this period was 16.9% (median concurrent behaviour= 84.9%; [LQ,UQ]=[53.9%, 100%]; Figure 11).
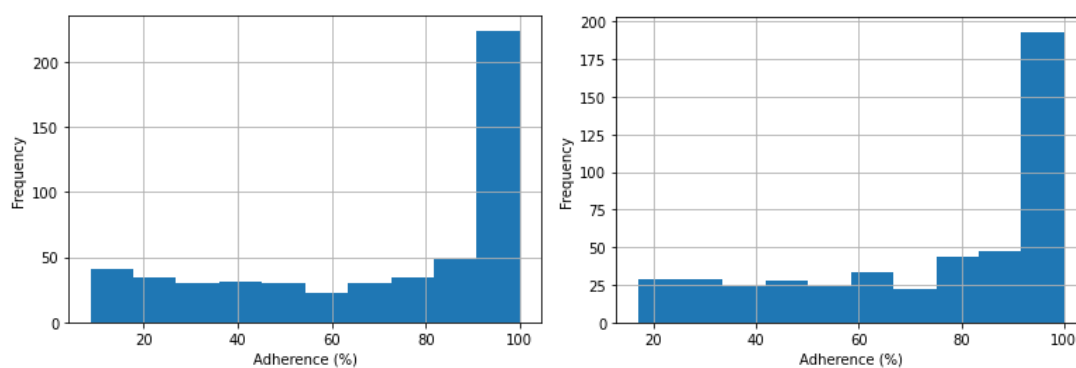
Figure 11
Distribution of adherence to first dose of medication during the period over which the nearest 14 inhalations (left) (n=526) and nearest 28 inhalations (right) (n=473) closest to interval 7 were extracted.

For participants with the lowest adherence over this period, inhalations recorded as early as the first interval of the trial were included in the sample of inhalations used to calculate context stability variables with 14 observations (n=36 participants, 6.8%). However, most participants had 14

inhalations recorded within interval 6 (i.e., within one interval prior to interval 7; n=349, 66.3%). This was similar for the sample of inhalations used to calculate context stability variables with 28 observations; 19 participants (4.0%) sampled data from as early as the first interval of the trial to collect 14 observations, but the majority (n=346, 73.2%) sampled data from only intervals 5 and 6.

6.4.2 Comparison of variables calculated with 14 and 28 observations

For both the 14 and 28 inhalation sets, all potential context stability variables (n=127 in each) were significantly associated (p<0.05) with concurrent adherence. The mean size of correlation coefficient was approximately the same for both sets (mean r=0.26 (SD=0.05) using 14 observations; mean r=0.27 (SD=0.07) using 28 observations). Confidence intervals surrounding correlation coefficient estimates were also approximately equal in size indicating similar precision (mean CI width=0.16 using 14 observations; mean CI width=0.17 using 28 observations). Small differences in precision were expected given the larger sample available for calculation of variables with 14 observations. These findings satisfied that similar relationships could be observed between context stability variables and concurrent behaviour with only 14 observations, as could be obtained using 28 observations. Therefore, the 14-observation set were taken forward to explore variable interpretability.

6.4.3 Bivariate correlations of context stability variables with concurrent behaviour

Bivariate correlations between each context stability variable and concurrent behaviour ranged from weak to moderate (Table 16). The weakest observed relationship was for W at radius 90 (r= 0.14), the strongest was LAM at radius 90 (r=0.34).

6.4.4 Exploration of the interpretability of context stability variables

Interpretability was operationalised as evidence that the actual and hypothesised correlation with concurrent behaviour was consistent in direction and that the range of values for each variable can be meaningfully interpreted. These conditions were upheld for all variables except DET and VENT (Table 16). Therefore, a total of 13 of 15 potential variable categories (n=97 individual potential context stability variables) were identified as suitable for further analysis in Chapter 7. The issues identified in the excluded variables are discussed, below.

Table 16

Bivariate correlations between variables and behaviour and checks for interpretability.

Variable	Theoretical interpretation	Expected direction of effect	Observed correlation with concurrent adherence	Range of observed values	Take forward to Chapter 7?
Range, in minutes, of the time of day at which inhalations taken (RANGE)	Captures the earliest and latest that a patient took their first dose of medication of the day	-ve Low scores indicate high context stability	-0.30*	26.53 to 1422.92 minutes	Yes
Inter-quartile range, in minutes, of the time of day at which inhalations taken (IQR)	Captures the spread of times that the first medication of the day was taken	-ve Low scores indicate high context stability	-0.27*	2.72 to 886.71 minutes	Yes
Standard deviation, in minutes, of the time of day at which inhalations taken (SD)	Captures the spread of times that the first medication of the day was taken	-ve Low scores indicate high context stability	-0.29*	7.27 to 493.98 minutes	Yes
Entropy of the time of day at which inhalations taken (ENT)	Captures the equality in probability of inhalations occurring at different times of day, when the day is separated into 20-minute intervals.	-ve Low scores indicate high context stability	-0.29	0.51 to 2.64	Yes
Entropy in the differences in time of day between consecutive inhalations (DIFF_ENT)	Captures equality in probability of inhalations occurring at different intervals from day to day, when intervals are separated into lengths of 20-minutes.	-ve Low scores indicate high context stability	-0.23*	1.09 to 2.56	Yes
Mean, in minutes, of the differences in time of day between consecutive inhalations (DIFF_AVG)	Captures the average of the differences between times of day at which consecutive inhalations were taken.	-ve Low scores indicate high context stability	-0.32*	6.90 to 685.92 minutes	Yes
Standard deviation, in minutes, of the differences in time of day between consecutive inhalations (DIFF_SD)	Captures the spread of the differences between times of day at which consecutive inhalations were taken.	-ve Low scores indicate high context stability	-0.26*	4.94 to 460.97 minutes	Yes

Table 16
Continued

Variable[a]	Theoretical interpretation	Expected direction of effect	Observed correlation with concurrent adherence[b]	Descriptive statistics[c]	Take forward to Chapter 7?
Recurrence Rate (RR)	Captures the proportion (%) of observations of times at which a medication was taken within a pre-specified radius of another observation (i.e. a recurrence point).	+ve Higher RR indicates higher proportions of medication taking events in the same context	0.24 to 0.33*	7.14 to 100%	Yes
Determinism (DET)	Captures the proportion (%) of recurrence points (see RR) which fall on a diagonal line with one or more recurrence point.	+ve Higher DET could indicate shifts in the context a medication was taken which is later repeated (i.e. shift from 9 to 11am from Friday to Saturday, each week).	0.18 to 0.33*	0 to 100%[d]	No- the upper observed value was 100% which is difficult to interpret.
Laminarity (LAM)	Captures the proportion (%) of recurrence points (see RR) which fall on a vertical line with at least one other recurrence.	+ve Higher LAM indicates more consecutive days with medications taken in the same context.	0.21 to 0.34*	0 to 100%	Yes
Trapping Time (TT)	Captures the average length of vertical lines of recurrence points (see RR) which fall on a vertical line with at least one other recurrence.	+ve Higher TT indicates lengthier, sustained periods for which medications were taken in the same context.	0.14 to 0.24*	1.9 to 14[e]	Yes
Maximum vertical line length (VMAX)	Captures the longest time a patient was able to maintain recurrences (see RR).	+ve Higher VMAX indicates evidence that the longest sustained period in the same context.	0.19 to 0.28*	1 to 14	Yes
Entropy in vertical line length (VENT)	Captures the entropy in the lengths of vertical lines of recurrence points (see RR).	-ve Higher VENT could indicate more variability in the lengths of time the patient is able to sustain behaviour in the same context.	0.15 to 0.27*	0 to 2.07	No- the observed association between VENT and adherence is in the opposite direction to that which was expected.
Average white vertical line length (W)	Captures the average time between recurrence points (see RR).	-ve Higher W indicates more occasions of medications taken out of context, between medications taken in the same context	-0.14 to -0.33*	0.9 to 12.5[f]	Yes

Table 16

Continued

Variable[a]	Theoretical interpretation	Expected direction of effect	Observed correlation with concurrent adherence[b]	Descriptive statistics[c]	Take forward to Chapter 7?
Entropy in white vertical line length (WENT)	Captures the entropy in the lengths of vertical white lines (i.e. with no recurrence points- see RR).	-ve Higher WENT indicates more variability in the number of occasions that a medication is taken out of context, between those taken in context.	-0.28 to -0.32*	0 to 2.48	Yes

[a]All RQA variables were calculated for recurrences in intervals of 5,10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 70, 80, 90 minutes.

[b]Range represents the range of observed correlations between each variable at each of the pre-specified radii with adherence at interval7

[c]Descriptive statistics are given as ranges summarised across each of the observed statistics for each variable at each of the pre-specified radii

[d] Missing values were imputed with a value of 0 (minimum DET is 0 but will report missing if no recurrences off the line of identity).

[e]Missing values were imputed with a value of 1.9 (minimum TT is 2 but will report missing if no lines (i.e., no recurrences) of length 2 or more).

[f]Missing values were imputed with a value of 0.9 (minimum W is 1 but will report missing if there were no observed non-recurrences (i.e., no time between recurrence points and the recurrence plot is saturated)).

*Correlations significant at p<.01

Variable[a]	Theoretical interpretation	Expected direction of effect	Observed correlation with concurrent adherence[b]	Descriptive statistics[c]	Take forward to Chapter 7?

6.4.5 Variables excluded from further exploration

6.4.5.1 Determinism

Determinism (DET) describes the proportion of recurrence points that lie on a diagonal line of minimum length 2. The proposed theoretical interpretation of DET in relation to context stability of medication taking was that higher DET could indicate shifts in the context a medication was taken which is later repeated (i.e., shift from 9 to 11am from Friday to Saturday, each week). Concerns were raised about what this variable was representing in this dataset as, whilst it was expected that a positive association could exist between DET and future behaviour, it was unexpected that the upper observed value of DET would be 100%.

An example of how DET can be meaningfully applied and interpreted in RQA is displayed in the recurrence plot (RP) in Figure 12. DET=88.7% in this RP, reflecting that most recurrence points lie on diagonal lines. Here, DET usefully describes the repetitions in patterns of change in this time series.
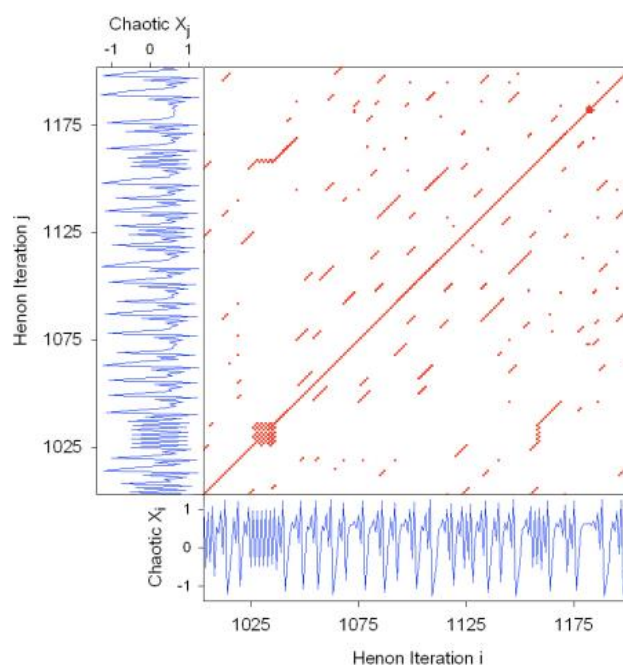
Figure 12
Taken from Weber & Zbilut (2005) p.47. Recurrence plot of a system, DET=88.7%

In RQA analysis, a truly random signal should have determinism score of zero (i.e. if someone was taking their medication randomly) but DET can deviate from zero when the radius parameter is high enough to capture arbitrary recurrences (Weber & Zbilut, 2005). This score is therefore sensitive to changes in the pre-specified radius parameter.

When RQA is used to describe the recurrences in a single system (as in Figure 12), the radius parameter is informed either by existing evidence or by a detailed exploration of the effects of changing the radius parameter. However, the purpose of the application of RQA in this context is to make comparisons between different time series, rather than to produce a detailed description of a single time series. Plots for 18 different radius parameter settings were created to enable an exploration of the optimal radius parameter for the prediction of predict future behaviour. All radius parameters were uniformly applied to all participant's time series and this, as anticipated, leads to a diversity in the saturation of RPs between individuals. The utility of RQA in this context is in extracting the RQA variables that can explain the variability in RP saturations in a meaningful way.

Exploration of a participant's RP with DET=84.7% (Figure 13) demonstrates that DET is likely confounded with plot saturation. DET is high in this plot as many points are lying on diagonals with other points, but the patterning of this RP could be more intuitively described by the distribution of vertical black and white points, rather than by diagonals.
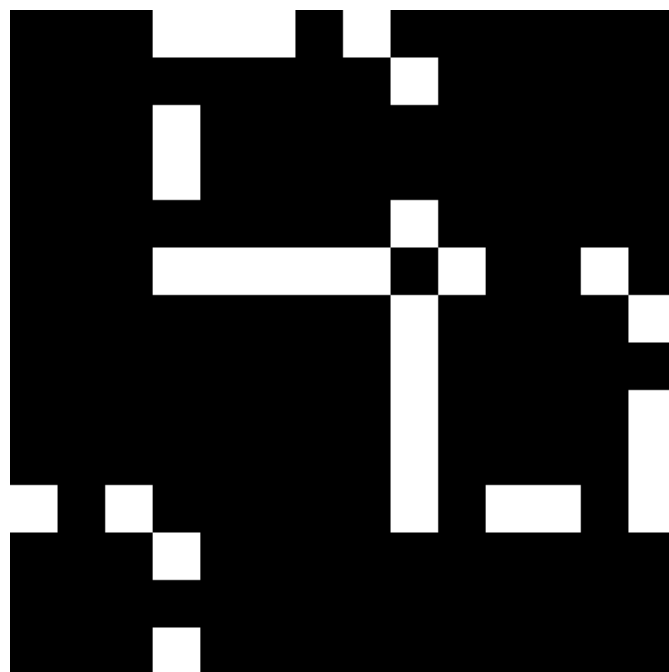
Figure 13
Participant's RP calculated with radius= 60 minutes, DET=84.7%

Despite DET=84.7% at radius=60 minutes, this individual does not appear to be having lots of shifts of greater than 60 minutes in the times at which their medications are taken which are then later repeated, according to the proposed theoretical interpretation of DET as applied to context stability. Rather, this person appears to have many consecutive days in which they take their medication

within 60 minutes of another day, with a few exceptions. The latter interpretation is reflected in the plot of raw inhalation times (Figure 14).
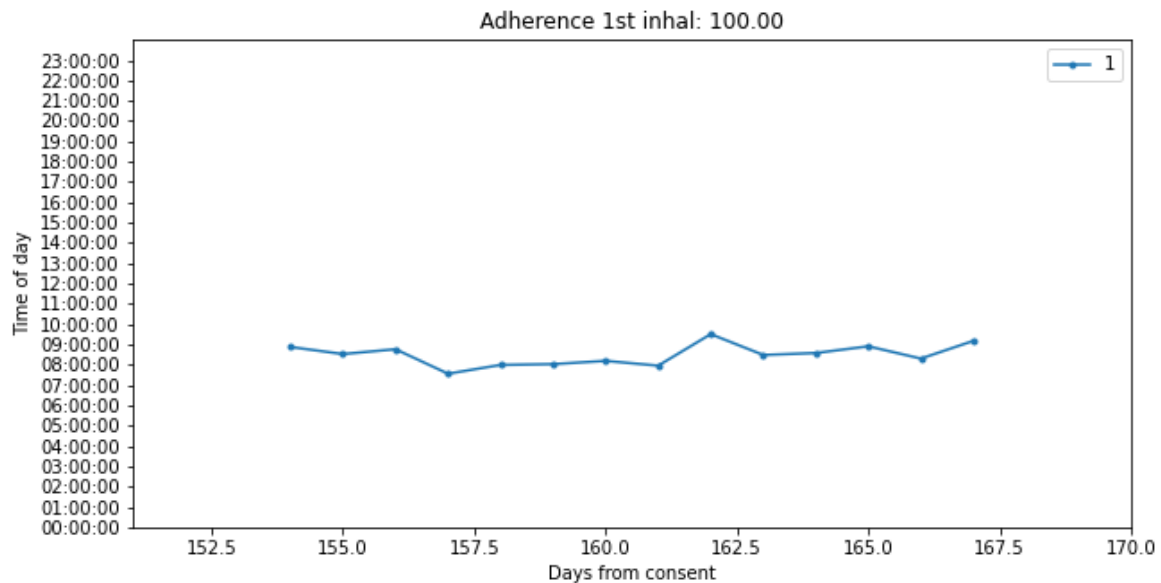
Figure 14
Times at which the 14 sampled datapoints used to calculate context stability were observed for one participant. This participant mostly takes their medication within an hour radius. This could most usefully be described in terms of many recurrences of the same context (e.g., by a LAM or RR variable), rather than by consistently changing routines each day, throughout the week then repeating this again the next week (i.e., by a DET variable).

Figure 14 makes clear that the theoretical interpretation of DET as applied to this data does not reflect the raw data. Further inspection of the full dataset indicated that as the radius parameter was increased, DET appeared to converge toward recurrence rate (RR) and laminarity (LAM) scores, (e.g., r=.88 and 0.85, respectively when RAD=90mins). Therefore, it appears that the patterning of RPs described by DET can be more intuitively explained with RR and/or LAM, and that DET is not a useful variable for the purposes of describing context stability in this dataset.

6.4.5.2 Entropy in the lengths of vertical lines

Vertical line entropy (VENT) describes the complexity of vertical line structures in the RP of minimum length 2; i.e., if the length of vertical lines on the RP were all of equal length, VENT=0. The proposed theoretical interpretation of this variable in relation to context stability was that higher VENT could indicate more variability (i.e., less context stability) in the lengths of time the people with CF is able to sustain behaviour in the same context. Therefore, it was expected that VENT would be negatively associated with future behaviour. However, the observed correlation between VENT and adherence was positive (r=0.15 to 0.27, dependent on the pre-specified radius parameter).

More detailed analysis of the behaviour of this variable indicated that VENT may also be confounded with RP saturation. To exemplify this, two plots are presented in Figure 15; the left of this figure presents a people with CF with a highly saturated RP and high VENT (RR=84.7%, VENT=1.88) whilst the right presents a people with CF with a comparatively sparse RP and minimum VENT (RR=18.4%, VENT=0). The right RP has a mixture of recurrence points of vertical length one and two meaning the participant is only managing to maintain behaviours in the same context for short segments. VENT only calculates entropy in lines of a minimum length of two, therefore entropy was equal to 0 in this plot. Meanwhile, the right RP is almost the opposite as it contains mostly recurrence points broken up by a few *non*-recurrences of length 1 and 2. The positioning of these short breaks in the recurrence patterns meant that recurrence streaks are broken up into segments of many different lengths, resulting in high entropy. This feature of this entropy variable reduces its interpretability for describing context stability and therefore was excluded on this basis.
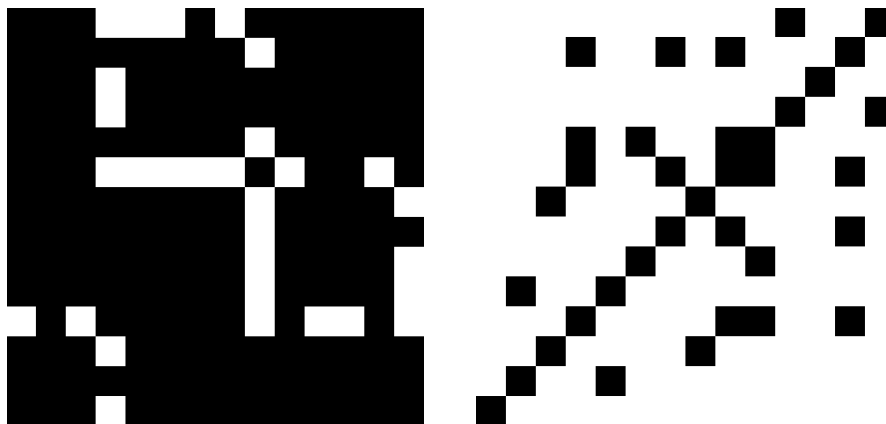
Figure 15
RPs for two participants from the ACtiF trial. Left RP: RR=84.7%, VENT=1.88; Right RP: RR=18.4%, VENT=0.

6.5 Discussion

6.5.1 Summary of findings

Chapter 6 aimed to explore a range of methods to describe context stability using objective medication adherence data, and to examine the interpretability of these variables with respect to their relationship with behaviour. A range of simple statistics and a more complex method called Recurrence Quantification Analysis were applied to the times of day at which people with CF took their medications to produce a total of 15 variable categories, including 127 potential context stability variables, which were available for examination. A total of 13 variable categories (n=97 individual variables) were classified as suitable to progress for further analysis in Chapter 7 which will aim to select a metric of context stability from this pool.

The principal contribution from the analyses presented in Chapter 6 was the novel application of Recurrence Quantification Analysis (RQA) to temporal, behavioural data, with the purpose of describing context stability. RQA is emerging as a useful technique the field of psychological and social research (e.g. see Heino et al., 2021; Jenkins et al., 2019). Applied to the problem of context stability measurement, RQA demonstrated several properties that improve on the limitations of previous attempts to measure context stability with objective temporal data. First, this method enables the capture of habits triggered by routines that happen at similar times on not only on a day-to-day basis, but also those that occur week-to-week. For example, if a person takes their medication after brushing their teeth, but they reliably brush their teeth an hour later on a Sunday than the rest of the week, this could still be captured as a recurrent point. Second, the same number of data points for each participant were used to construct each measure, therefore successfully delineating the variance captured from that described by frequency. This enables non-conflated examination of the relationship between context stability and behaviour which previous implementations of BFCS and context stability measures have failed to achieve. Variables derived using the methods presented in this chapter demonstrated comparable correlational effect sizes to those observed from self-reported measures of context stability (e.g. Maher et al., 2021; Norman & Cooper, 2011; Pimm et al., 2015).

6.5.2 Limitations

The variables derived in this Chapter are intended as proxy measures of context stability; it is not the intention of this thesis to present times of day as explanatory cues to behaviour. These measures are a pragmatic solution to the fact it is very difficult to discern cues to action. The proposed solution to the development of a metric is an attempt to best capture a proximal measure of context stability, which is readily available to researchers with electronic devices that collect information on the initiation of behaviour.

There was no concrete information in the present dataset to determine whether, for a person prescribed more than one medication per day, an observation many hours later than other 'first' observations was truly the participant completing a 'late first session', or if the participant perceived themselves to have skipped the first session and instead was completing their 'second' session of the day. Conceptually, each session could be thought of as a separate behaviour, and each could have different cues and degrees of habitual automaticity. This is a limitation of any approach that passively captures a behaviour that occurs multiple times a day and does not require additional input from research participants to label their behaviour. Attempts were made to reduce noise in the data caused by objective measurement vs subjective experience of behaviour, by adjusting the 'daily' clock from 12-12am, to 5-5am. This was to align the measurement of days with the

participants' likely experience and was also in line with previous adherence and context stability research in the CF population (see Hoo, Wildman, et al., 2019).

6.6 Conclusion and implications for the following thesis chapters

Chapter 6 identified 97 potential context stability variables which appear to be associated with concurrent behaviour (adherence to medication), including some which have important conceptual improvements on measures previously used in the literature. These variables will be taken forward to Chapter 7 in which feature selection methods will be applied to select which variable(s) to use.

Chapter 7 Selection of a context stability metric and evaluation of its utility

7.1 Chapter Overview

Chapter 6 presented novel methods to measure context stability and initiated the exploration of the relationship between variables derived from the time of day at which people with CF took their medications and their adherence behaviour. The process resulted in 97 potential variables that could objectively describe context stability in medication adherence data. Therefore, the first objective of Chapter 7 was to fulfil the remainder of Thesis Aim 2 (i.e., to present a systematic approach to identifying and selecting a metric of context stability, derived from objective data on the times of day at which people initiate medication adherence behaviour) with the following objective:

1) To identify which of these variables explain the most variance in concurrent medication adherence behaviour and use this information to define an objective metric of context stability (Thesis Aim 2, Objective 2.3).

Following this, Chapter 7 began the exploration of Thesis Aim 2, Objective 3, which was to evidence the utility of the defined objective metric of context stability in explaining variance in maintained and future behaviour, and in related constructs in evidence of predictive, convergent and divergent validity. The specific objectives addressed were:

1) To examine the amount of variance explained in concurrent behaviour by the objective metric of context stability (Thesis Aim 2, Objective 3.1);

2) To examine the amount of variance explained in participants' self-reported habitual automaticity for taking their medication by the objective metric of context stability (Thesis Aim 2, Objective 3.2);

3) To examine the amount of variance explained in perceived effort participants required to take their medication by the objective metric of context stability (Thesis Aim 2, Objective 3.3);

4) To examine whether distinct variance in concurrent behaviour is explained by reflective motivational constructs and the objective metric of context stability (Thesis Aim 2, Objective 3.4).

These aims and objectives are addressed in two parts within the following chapter.

7.2 Part 1: Selection of a context stability metric

7.2.1 Background

The selection of a subset of most useful variables or 'features' from a large number of variables is a common task in the field of statistics and machine learning, and is increasingly utilised during the application of machine learning techniques to medical and healthcare problems (Remeseiro & Bolon-Canedo, 2019). The purpose is to condense an array of available features by removing redundant or irrelevant information, thereby optimising the data given to a prediction model (Guyon et al., 2006). Gunyon et al., (2006) identify four situations in which reducing the number of features is useful: i) particularly relevant to the field of machine learning, some datasets are so large they need to be condensed in order reduce storage requirements and increase computation speed; ii) reduction of the feature set can increase predictive accuracy of models in some instances; iii) reducing the number of features given to an algorithm can increase understanding of the processes that led to specific predictions; and finally, iv) if the algorithm is to be applied in real-world problems, reducing the number of features means less data needs to be collected again in the future.

A large variety of feature selection methods have been developed over several decades, and range from relatively simple to complex and computationally expensive (Guyon et al., 2006; Remeseiro & Bolon-Canedo, 2019). The performance of each technique varies according to the data complexities and the specific task; no method ubiquitously performs 'best' across feature selection problems (Bolón-Canedo et al., 2012).

Applied to the present task, the purpose of reducing the number of potential context stability variables from those articulated in Chapter 6 both to improve interpretability and to reduce the burden of variable derivation from time-stamped behavioural data. In this task, simple methods such as individual relevance ranking are likely to be inappropriate. This is a univariate approach in which individual features are ranked according to their univariate relationship with the outcome variable. For example, metrics such as Pearson's correlation could be evaluated for each feature and the outcome variable, and the features ranked from highest to lowest by magnitude of the observed r. The present task has a high risk of multicollinearity due to the fact all variables were derived from the same raw data, meaning that whilst many features may have individual utility, they may become redundant when in competition with other variables. This may not be an issue if in search of a single variable, however habit researchers have commonly devised multifaceted measures of context stability which incorporate multiple 'domains' into a single score (e.g. McCloskey & Johnson, 2019; Pimm et al., 2015); this may be relevant to the variables available for the present analysis if features

show low(er) relevance individually but can explain additional, incremental variance when used in combination with other variables. Therefore, an approach which enables all variables to 'compete' alongside one another, accounting for feature dependencies (Guyon et al., 2006) would be more appropriate than the univariate approach.

Broadly, the process of feature selection involves the generation of a bank of features, definition of an evaluation criterion and estimation of this criterion for each of the features, or feature subsets. The criterion applied in this instance will be the variable(s) which explain the most variance in behaviour.

The following analyses presented in Chapter 7, Part 1 implement a multivariate feature selection technique known as the Least Absolute Shrinkage and Selection Operator algorithm (LASSO; Tibshirani, 1997) in order to identify which of the 97 variables from Chapter 7 explained the most variance in concurrent medication adherence behaviour. The specific objective was:

1) To identify which variables explain the most variance in concurrent medication adherence behaviour and use this information to define an objective metric of context stability (Thesis Aim 2, Objective 2.3).

7.2.2 Methods

7.2.2.1 Context stability

The 97 context stability variables carried forward from Chapter 6 were retained for analysis in the present chapter. These were RANGE, IQR, SD, ENT, DIFF_ENT, DIFF_AVG, DIFF_SD, RR, LAM, TT, VMAX, W and WENT. Each RQA derived variable (RR, LAM, TT, VMAX, W, WENT) calculated at each radius setting (5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 70, 80 and 90 minutes) was included as a separate variable. For example, RR_rad_5, RR_rad_10 were two separate variables describing RR calculated at 5 and 10-minute radii, respectively. Variables for which high scores indicated lower context stability (RANGE, IQR, SD, ENT, DIFF_ENT, DIFF_AVG, DIFF_SD, W, WENT) were reverse coded by taking scores away from the maximum score, so that high scores indicated high context stability for all context stability variables. This resulted in a total of 97 variables available for analysis. Following the findings from Chapter 6, all context stability variables were calculated using 14 time- and date-stamped adherence observations from each participant.

7.2.2.2 Concurrent behaviour

As in Chapter 6, concurrent behaviour was calculated as the percent adherence to the first medication of the day over the period from which the 14 inhalations used to calculate context

stability were collected. For example, a participant with 14 inhalations collected over 20 days would have concurrent behaviour equal to 70% (i.e., 14/20*100).

7.2.2.3 Analysis

LASSO regression

A Least Absolute Shrinkage and Selection Operator (LASSO) algorithm with repeated 10-fold cross-validation was implemented to identify the variable(s) of context stability with greatest predictive utility for concurrent behaviour. Multicollinearity was expected between the various context stability measures. LASSO regression handles this by shrinking coefficients of variables to zero if they are highly correlated with another variable which has more explanatory power. This characteristic identifies LASSO as a feature selection method as this process essentially removes the effects of less predictive variables from the model. In practice, calculation and interpretation of many similar variables is both impractical and not parsimonious and application of the LASSO regression method in this analysis enabled the identification of the context stability variables with greatest utility in the prediction of concurrent behaviour.

All variables were standardised prior to analysis to ensure the penalisation term was applied equally to all variables, thus enabling comparability between variables in the interpretation of their relative importance.

One-hundred repetitions of a 10-fold cross validation procedure were completed (depicted in Figure 16). During this process, the dataset is first shuffled, then randomly split into 10 folds (i.e., subsets of the data) with equal number of observations in each fold. For each of the 10 folds, the LASSO algorithm returns a coefficient estimate for each variable by fitting the model to the data in nine of the 10 folds and then calculates the predictive utility of the fitted model on the 10th fold. Each fold is held out once per repetition such that all participants act, at one point, as a hold-out test set. This process was set to repeat 100 times, resulting in a total of (10 x 100) or 1000 models fit to the data, and 1000 coefficient estimates. Use of the cross-validation procedure enabled the estimation of the out-of-sample generalisation score for the predictive utility of the LASSO model but more importantly for the present analysis, a description of the magnitude and precision of coefficient estimates across the 1000 models fit.

Out-of-sample generalisability scores were calculated as the coefficient of determination (R^2), representing the proportion of variance in the outcome variable, in this case concurrent behaviour, that is explained by the independent variables in the model.

The specific objective of Part 1 was to understand which of the variables retained from Chapter 6 explained the most variance in concurrent adherence behaviour. This was determined in terms of 'variable importance'. Variable importance was interpreted from coefficient size and precision; for each context stability variable, these were respectively calculated by taking the average and 95% confidence intervals (95%CI) of the 1000 coefficient estimates for each variable.

LASSO models can be inconsistent in feature selection tasks when variables are highly correlated, therefore the robustness of the variable importance analysis was tested by first conducting the analysis with the 14 inhalations closest to interval 7 (hereafter referred to the interval 7 analysis). The procedure was then repeated with data closest to intervals 10 and 12 in which context stability variables and concurrent behaviour were calculated using the nearest 14 inhalations to interval 10 and 12 (hereafter referred to the interval 10 and 12 analysis), respectively.

LASSO regressions were implemented in Python (v3.8.3) using Scikit-Learn's LASSO regression software (v1.0.2).
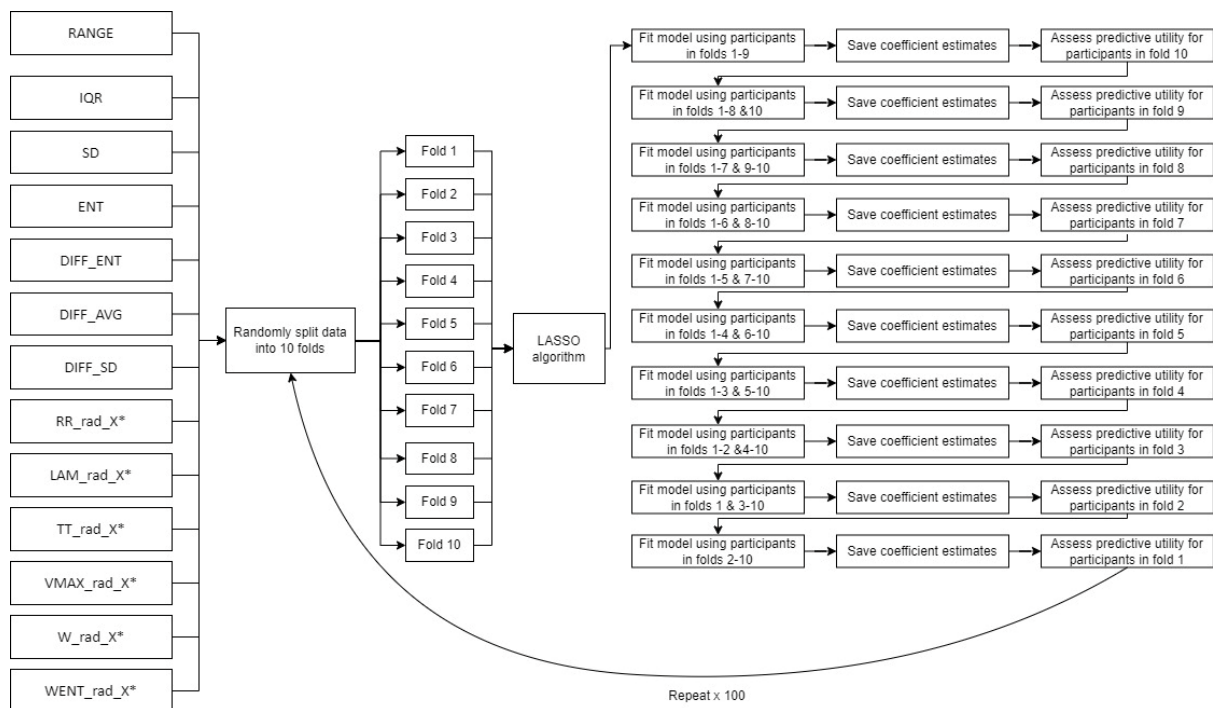
Figure 16
Flow diagram of LASSO algorithm with 100 repetitions of a 10-fold cross validation procedure.
Variables calculated at each of 5-, 10-, 15-, 20-, 25-, 30-, 35-, 40-, 45-, 50-, 55-, 60-, 70-, 80- and 90-minute radii in RQA analysis.

7.1.3 Results

7.1.3.1 Overview of the dataset

As in Chapter 6, the analysis at interval 7 was conducted using data from 526 participants; inclusion in analysis required a minimum of 14 observations between the start of the trial and the end of interval 6 (see Appendix F, Table F1 for comparison with original trial sample).
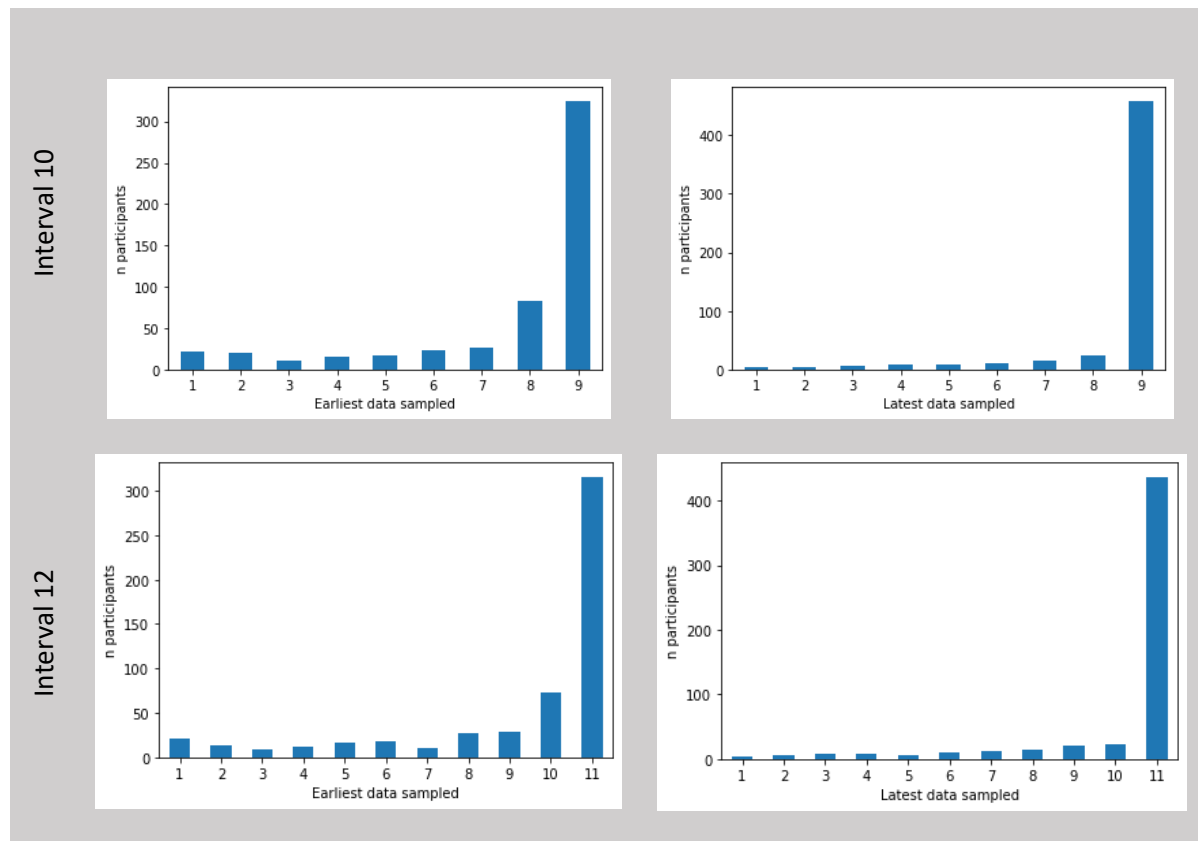
Figure 17 shows that most participants had enough data from interval 5 or 6 to calculate context stability and concurrent behaviour for the interval 7 analysis (n=419; 79.7%). However, data were sampled from as early as the first interval of the trial for 36 (6.8%) participants. Very few participants (n=5) had data sampled exclusively from interval 1 (n=5) or interval 2 (n=5).

The pattern of data sampling was similar at intervals 10 and 12 to interval 7 (Figure 17; see Appendix F, Table F2 and F3 for comparison with original trial sample). A total of 542 and 546 participants had sufficient data for inclusion in analysis at intervals 10 and 12, respectively. Most participants had enough data in the two intervals closest to interval 10 (n=408; 75.3%) and 12 (n=389; 71.2%) to sample 14 observations for analysis[9].

---

[9] The data sampled in each analysis was not mutually exclusive; for some participants, some of the data points used in the interval 7 analysis would have also been sampled in the interval 10 and/or interval 12 analysis if they were very low adherers.
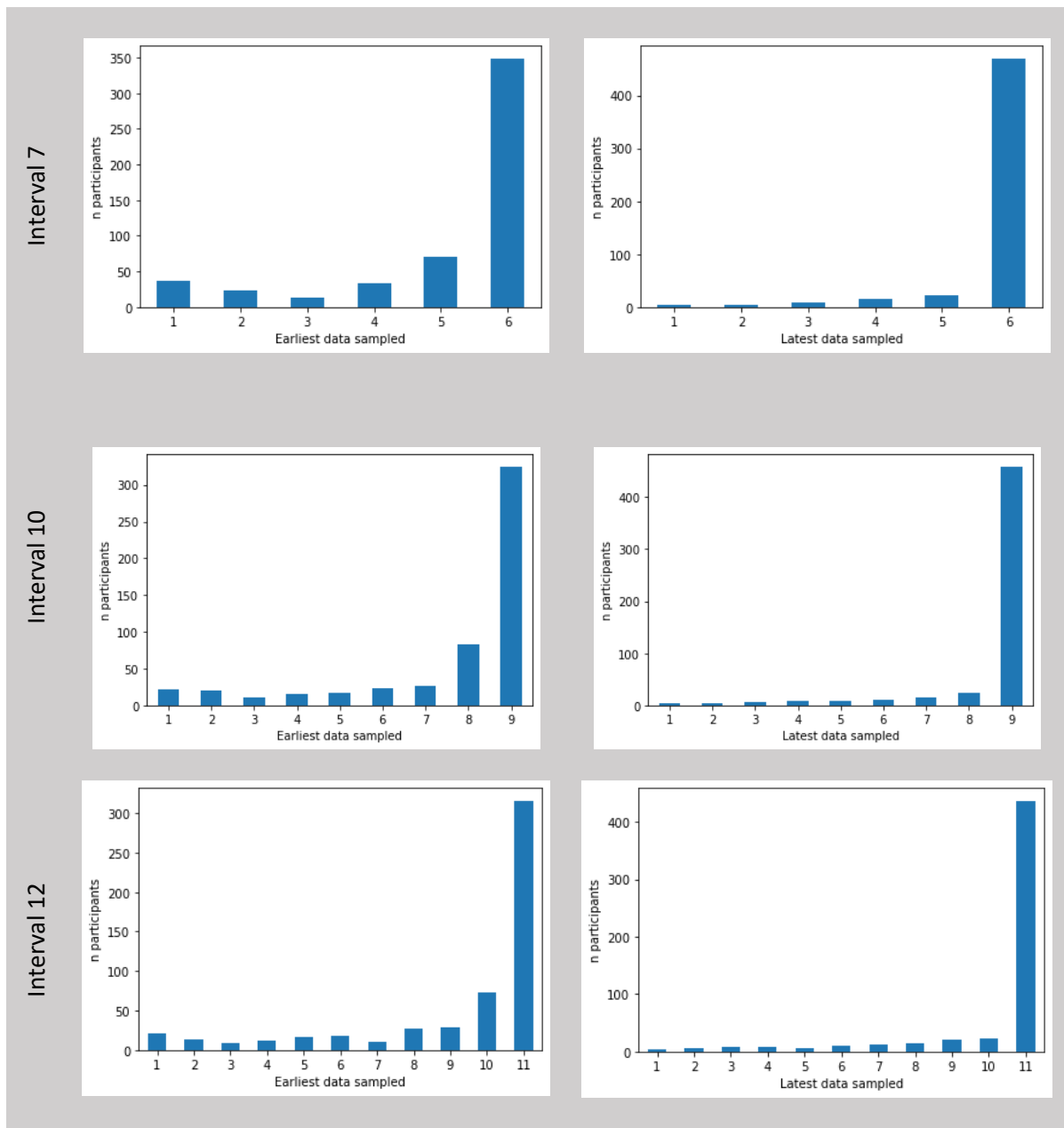
Figure 17

Earliest and latest intervals data were sampled from participants included in the interval 7 LASSO analysis (n=526), interval 10 (n=543) and interval 12 (n=546).

Bivariate correlations with concurrent behaviour were largely similar across datasets closest to intervals 7, 10 and 12 for each of the context stability variables (Table 17). Small-to-medium correlations with concurrent behaviour were observed across all context stability metrics at intervals 7, 10 and 12.

Table 17

Bivariate correlations between each of the context stability variables calculated with data closest to intervals 7, 10 and 12, with concurrent behaviour from these periods.

Variable	Observed correlation with concurrent behaviour at interval 7	Observed correlation with concurrent behaviour at interval 10	Observed correlation with concurrent behaviour at interval 12
RANGE	-0.30**	-0.30**	-0.30**
IQR	-0.27**	-0.25**	-0.22**
SD	-0.29**	-0.29**	-0.28**
ENT	-0.29**	-0.29**	-0.32**
DIFF_ENT	-0.23**	-0.21**	-0.27**
DIFF_AVG	-0.32**	-0.27**	-0.26**
SD_DIFF	-0.26**	-0.23**	-0.22**

Variable[a]	Observed correlation with concurrent behaviour at interval 7[b]	Observed correlation with concurrent behaviour at interval 10[b]	Observed correlation with concurrent behaviour at interval 12[b]
RR	0.24 to 0.33**	0.24 to 0.29**	0.26 to 0.33**
LAM	0.21 to 0.34**	0.20 to 0.25**	0.24 to 0.30**
TT	0.14 to 0.24**	0.11 to 0.16*	0.15 to 0.22**
VMAX	0.19 to 0.28**	0.16 to 0.24**	0.19 to 0.27**
W	-0.14 to -0.33**	-0.22 to -0.32**	-0.21 to -0.29**
WENT	-0.28 to -0.32**	-0.26 to -0.33**	-0.26 to -0.30**

**Correlations significant at p<.001
*All correlations were significant at p<.001, except TT_rad_20 and TT_rad_30
[a]All variables calculated with intervals set to 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 70 80 & 90 minutes
[b]Correlations summarised for r's observed across all 15 radius settings

7.3.1.2 Overall predictive utility

When all context stability variables were entered into the LASSO regression for the interval 7 analysis, the mean amount of variance explained (R^2) in the 1000 training sets was 16.2% (SD=0.9%). The generalisability score (R^2 for the 1000 hold-out test sets) indicated that 10.6% (SD=8%) of the variance in concurrent behaviour could be explained by the context stability measures, collectively. Mean predictive utility in the training datasets was similar at interval 10 (16.3%, SD=0.9%) and interval 12 (16.6%; 0.9%). This was the same for the hold-out sample when all context stability

measures were entered into a LASSO regression; findings were similar to interval 7 (R^2=10.6%, SD=8%) at interval 10 (R^2=9.3%, SD=9%) and interval 12 (R^2= 11.3%; 8.2%).

7.3.1.3 Exploration of CS variable importance

Figure 18 shows the mean and standard deviation of the coefficient estimates for each of the 10 context stability variables with largest mean coefficient estimates, across the 1000 models for the interval 7, 10 and 12 analyses. All context stability variables were standardised, therefore Figure 18 demonstrates the most important context stability measures for predicting concurrent behaviour in the LASSO regression for each interval analysis. For interval 7, the largest average coefficient estimate was for W_rad_5, DIFF_AVG and LAM_rad_90. The largest average coefficient estimates at interval 10 were W_rad_70, LAM_rad_15, RANGE and ENT; at interval 12, the highest-ranking variables were ENT, LAM_rad_90, RANGE, W_rad_90 and LAM_rad_5.
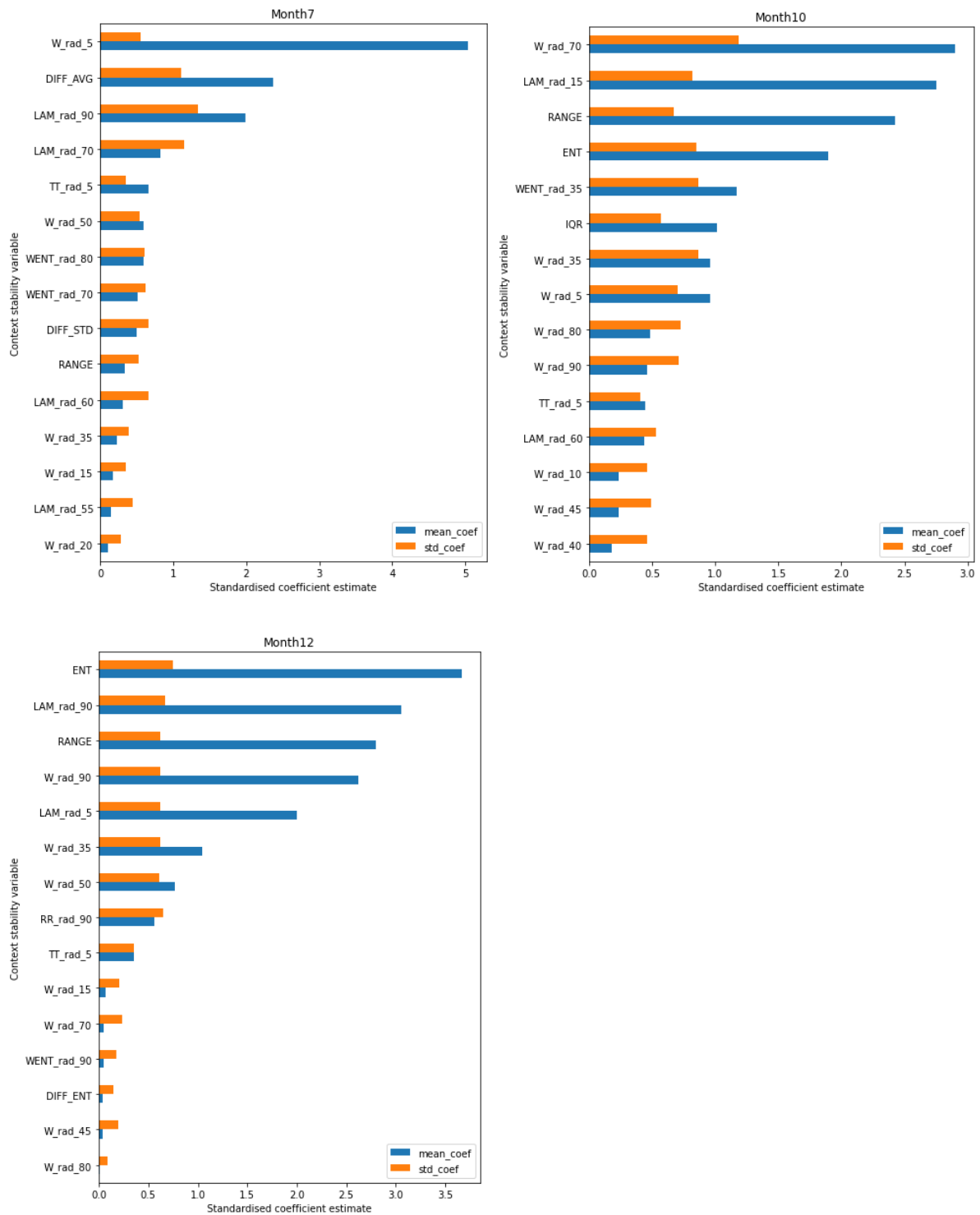
Figure 18

Mean and standard deviation of coefficient estimates for the 10 context stability variables with greatest mean coefficient estimates across 1000 folds at interval 7 (top left), 10 (top right) and 12 (bottom left).

Table 18 presents mean coefficient estimates, 95% confidence intervals and the proportion of the 1000 models in which the coefficient was shrunk to 0 (P0; i.e. the proportion of occasions on which a variable's effects were essentially removed from the model) for all three analysis intervals.

Table 18

Summary of variable coefficients across 1000 LASSO algorithm estimates from 100 repetitions of the repeated 10-fold procedure at intervals 7, 10 and 12.

Variable importance analysis with data closest to interval 7				
Importance rank	Variable name	Mean coefficient estimate	95% CI of coefficient estimate	Proportion of models in which coefficient estimate was equal to 0 (P0)
1	W_rad_5	5.04	[3.96, 6.12]	0
2	DIFF_AVG	2.37	[0.20, 4.54]	0.04
3	LAM_rad_90	1.99	[-0.64, 4.61]	0.13
4	LAM_rad_70	0.82	[-1.45, 3.08]	0.50
5	TT_rad_5	0.66	[-0.01, 1.33]	0.03
6	W_rad_50	0.60	[-0.46, 1.65]	0.26
7	WENT_rad_80	0.59	[-0.61, 1.79]	0.32
8	WENT_rad_70	0.52	[-0.70, 1.73]	0.41
9	STD_DIFF	0.50	[-0.81, 1.80]	0.47
10	RANGE	0.33	[-0.70, 1.36]	0.56

Variable importance analysis with data closest to interval 10				
Rank	Variable name	Mean coefficient estimate*	95% CI of coefficient estimate	Proportion of folds in which coefficient estimate was equal to 0**
1	W_rad_70	2.90	[0.58, 5.23]	0.00
2	LAM_rad_15	2.75	[1.15, 4.35]	0
3	RANGE	2.43	[1.12, 3.74]	0.01
4	ENT	1.90	[0.23, 3.56]	0.02
5	WENT_rad_35	1.17	[-0.52, 2.86]	0.15
6	IQR	1.01	[-0.09, 2.12]	0.05
7	W_rad_35	0.96	[-0.73, 2.65]	0.23
8	W_rad_5	0.96	[-0.41, 2.32]	0.13
9	W_rad_80	0.49	[-0.93, 1.90]	0.48
10	W_rad_90	0.46	[-0.92, 1.84]	0.47

Variable importance analysis with data closest to interval 12				
Rank	Variable name	Mean coefficient estimate*	95% CI of coefficient estimate	Proportion of folds in which coefficient estimate was equal to 0**
1	ENT	3.67	[2.20, 5.14]	0
2	LAM_rad_90	3.05	[1.75, 4.36]	0
3	RANGE	2.80	[1.58, 4.01]	0
4	W_rad_90	2.62	[1.40,3.84]	0
5	LAM_rad_5	2.00	[0.79, 3.21]	0.00
6	W_rad_35	1.05	[-0.17, 2.26]	0.13
7	W_rad_50	0.77	[-0.43, 1.97]	0.19
8	RR_rad_90	0.56	[-0.70, 1.83]	0.35
9	TT_rad_5	0.35	[-0.35, 1.05]	0.24
10	W_rad_15	0.07	[-0.34, 0.48]	0.83

Table 18 indicates that, of the 97 variables entered into the LASSO regression, there is likely no *single* variable that reliably, best predicts concurrent behaviour. However, inspection of the categories and combinations of variables that were chosen most frequently across the three timepoints and with largest effects revealed that a subset were repeatedly found to be more useful than others. These were W at very low and high radius settings, LAM at low and high radius settings, RANGE and ENT. Interpretations of these variables are presented in Table 19.

Table 19

Interpretations of the most frequently selected variables in the LASSO analysis.

Variable	Interpretation*
W	The average number of observations in which medications were taken out of context, consecutive with at least one other observation
LAM	The percentage of observations that were taken in context, consecutive with at least one other observation
RANGE	The range (in minutes) between the earliest and latest medication taken
ENT	The entropy** in the times of day that medications were taken

*Use of the term 'context' here refers to the number of minutes at which the radius is set; e.g. for W_rad_90, the variable would indicate the average number of observations in which medications were taken greater than 90 minutes earlier or later than any other observations, consecutive with at least one other observation

**Entropy was calculated as the entropy of observations across 20-minute bins.

7.3.1.4 Post-hoc exploration of the relationships between objective context stability variables

When W with high radius was selected, a low radius LAM variable was also selected at intervals 10 and 12. At interval 7, low radius W and high radius LAM were selected. This indicated that information provided by these two variables in combination may work synergistically to explain variance in behaviour. Furthermore, a third, non-RQA derived variable may also be useful, alongside these two variables, as indicated by the frequency of the selection of RANGE and ENT at intervals 10 and 12, and DIFF_AVG at interval 7, alongside W and LAM variables.

Bivariate correlations were explored to further understand the relationships between variables ranked highest among each of the interval analyses (Table 20). Whilst most highest-ranking variables demonstrated small-to-moderate and significant correlations with one another, near-zero and non-significant relationships were observed between W at high radii, and LAM at low radii, whilst independently demonstrating small to moderate correlations with concurrent behaviour. This suggests than in combination, high radius W and low radius LAM may explain different aspects of the variance in concurrent behaviour. In addition, RANGE demonstrated a consistent association of $r=0.30$ with concurrent behaviour at intervals 10 and 12.

Table 20

Correlations between most frequently selected variables from LASSO at intervals 7, 10 and 12.

Interval 7 Correlation Matrix (n=526)

	W_rad_5	AVG_DIFF	LAM_rad_90	Concurrent behaviour
W_rad_5	1	.42*	.50*	.33*
AVG_DIFF		1	.79*	.32*
LAM_rad_90			1	.34*
Concurrent behaviour				1

Interval 10 Correlation matrix (n=542)

	W_rad_70	LAM_rad_15	RANGE	ENT	Concurrent behaviour
W_rad_70	1	.06	.46*	.24*	.28*
LAM_rad_15		1	.37*	.67*	.24*
RANGE			1	.46*	.30*
ENT				1	.29*
Concurrent behaviour					1

Interval 12 correlation matrix (n=546)

	ENT	LAM_rad_90	RANGE	W_rad_90	LAM_rad_5	Concurrent behaviour
ENT	1	.51*	.43*	.17*	.68*	.32*
LAM_rad_90		1	.52*	.11*	.41*	.30*
RANGE			1	.39*	.23*	.30*
W_rad_90				1	.05	.22*
LAM_rad_5					1	.25*
Concurrent behaviour						1

*Correlation significant at p<.001

Based on the LASSO analysis findings and post-hoc explorations of relationships between the highest-ranking categories of variables, three were selected to be further tested for composite utility in predicting medication adherence behaviour in Part 2; these were W calculated at a 90-minute radius, LAM at a 15-minute radius, and RANGE. This combination was chosen in response to the observation that a high radius W variable, a low radius LAM variable and a simple summary statistic could be more usefully combined to predict concurrent behaviour than in isolation. RQA-derived variables were highly correlated with neighbouring radius settings (e.g., r=0.91 between W_rad_90 and W_rad_80 at interval 7). This may explain slight differences in radius settings of variables selected in each LASSO analysis, and why the approach of identifying variable categories of interest and patterns of radius settings was suitable. The selection of 90-minutes for W was because a high radius setting was required and conceptually, this was the upper limit of what was considered feasibly 'recurrent' due to the scheduling of medications, as described in Chapter 6- *RQA hyper-parameter selection*; W describes the non-recurrent points on the RP and therefore an observation more than 90 minutes from any other observation was chosen to conservatively capture a non-

recurrence. Similarly, the 15-minute radius for LAM was selected to satisfy the need for a low-radius LAM variable; LAM_rad_15 was highly correlated with lower radius settings, was the second most important variable in the interval 10 analysis behind a high radius W variable. Furthermore, because LAM captures the recurrent points on the RP, 15 minutes conceptually satisfied a reasonable amount of time within which a strictly adhered-to daily routine might be initiated. Finally, RANGE was selected as this variable had a consistent relationship with adherence over all three timepoints (r=.30) and conceptually, is most different from the LAM variable (compared to ENT). Used together, these variables would explain context stability in terms of: the most extreme difference in times that medication sessions was initiated over the period of observation (RANGE); the proportion of consecutive observations that were initiated within 15 minutes of one another (LAM_rad_15); the average number of consecutive observations which were initiated greater than 90 minutes apart (W_rad_90).

7.3.1.5 Selected context stability variables: demonstration in participant case studies

Together, these variables indicate high context stability in terms of non-extreme differences in time, frequent and/or lengthier stretches of behaviour repeated within 15 minutes of another observation and shorter stretches of behaviour more than 90 minutes different from another observation. To demonstrate how the three variables combine, Table 21 presents four case studies of participants with increasing context stability (as described by these three variables in combination), alongside their respective scores on each of the domains.

Table 21

Visualisation of recurrence plots and range scores which contribute to overall CS score

CS (%)*	Plot of 14 observations used to calculate context stability	RANGE (mins)	RP radius 15 (from which LAM_rad_15 is calculated)	RP radius 90 (from which W_rad_90 is calculated)
11.5%		1409.90		
50.15%		439.15		
60.15%		165.28		

Table 21

Continued

CS (%)*	Plot of 14 observations	RANGE	RP radius 15	RP radius 90
86.07%		26.53		

*CS score relates to the score calculated following the formula presented in Part 2, section 7.2.2.1 'Composite context stability'.

7.1.4 Part 1 Discussion

7.1.4.1 Overview of findings

A LASSO regression was implemented to identify which of 97 potential context stability variables explained the most variance in concurrent medication adherence behaviour. Objective medication adherence data was used to calculate variables using a combination of simple summary statistics and RQA methods, which were introduced in Chapter 6. The findings indicated that there was no single potential context stability variable that reliably predicted concurrent behaviour better than any other single variable. However, the information gleaned from the LASSO analysis enabled the identification a subset of variables which could work synergistically to explain more variance in medication adherence behaviour, than independently. Based on the findings of an analysis conducted across three timepoints, three variables were selected to describe context stability: RANGE, W_rad_90 and LAM_rad_15. RANGE described the range (in minutes) between the earliest and latest medication taken. Derived from RQA analysis, W_rad_90 and LAM _rad_15 described the recurrence patterns over 14 observations of an individual's medication taking behaviour: W_rad_90 described the average number of observations in which medications were taken greater than 90 minutes later than any other observation, consecutive with at least one other observation; LAM_rad_15 described the percentage of observations that were taken within 15 minutes of another observation, consecutive with at least one other observation. In other words, a score which combined these variables into one score would reflect context stability in terms of non-extreme differences in time, frequent and/or lengthier stretches of behaviour within 15 minutes of another observation and shorter stretches of behaviour more than 90 minutes different from another observation. A case study was presented to demonstrate how these variables relate to the data from which they were derived.

Use of multiple self-reported items assessing different types of context stability is common in the literature (e.g. Judah et al., 2018; Maher et al., 2021; Pimm et al., 2015; Tappe & Glanz, 2013; Wohn & Ahmadi, 2019); this analysis supported that context stability may be best described using multiple dimensions, even when using objective measures. Many self-reported measures have used multiple response items to measure context stability to represent multiple aspects of the environment which may serve as cues to behaviour (e.g., time of day, location, other people). The findings of Part 1 suggest it may be useful to translate this practice into the measurement of context stability with objective data, by selecting multiple variables which describe different characteristics of temporal data.

Of note for future studies wishing to implement context stability measures from objective data, this analysis found that the standard deviation (SD) was consistently outperformed by other variables

which as indicated by their selection both more frequently and with larger effects. This suggests that SD was not the most powerful variable to derive from time-stamped behavioural data to describe context stability in concurrent medication adherence behaviour.

7.4.1.2 Limitations

The lack of consistency in findings across the three time-points meant that the selection of variables was guided first by the data (i.e., the outcome of the LASSO analyses) and then by exploring hypothesis (i.e. post-hoc exploration of commonalities in the findings and relationships between variables) and applying theory (i.e. selection of radius settings in terms of theoretical understanding of what the variables represented and what that would mean for behaviour in the real world). The inconsistency in findings across the different analyses could be driven by multiple factors. First, interval 10 and 12 analyses were more consistent with one another than the interval 7. This may be due to more data being available at these time points or because of more proximity in real-time. Second, it was possible to observe similarities in the categories of variables which were selected but there were inconsistencies in the radius settings that were selected for RQA-derived variables; this could be due to correlation between neighbouring radius settings meaning it was necessary to apply theoretical understanding of what the variables were describing to the LASSO findings to decide on which radius settings to select. Feature selection methods are limited by their instability (Chandrashekar & Sahin, 2014) but steps were taken to overcome this by applying repeated k-fold cross-validation; the results were taken from 1000 subsets of the data therefore the results, as presented, are likely a fair representation of the trends in the available dataset and not a result of sampling issues. The dataset available for this analysis was large (>526 participants) and methods to assess robustness of the findings were taken, but replication of these findings in external datasets is necessary to be confident that the findings are generalisable to other samples, population groups and behaviours.

7.1.4.3 Implications for Part 2

Three variables were selected to describe context stability in a composite measure. Their selection was guided by data-driven methods and theory. Further examination of their ability to uphold construct validity is required.

7.2 Part 2: Examination of the predictive, discriminant and convergent validity of the selected objective context stability variables

7.2.1 Background

Simple statistics such as the variance, standard deviation or entropy of the times of day at which people perform behaviours have been used to describe context stability in objective data (Hoo, Wildman, et al., 2019; Maher et al., 2021; Phillips et al., 2022). As discussed in Chapters 4 and 6, there are some limitations with the approaches implemented in previous research. Furthermore, the additive effects of different objective summary scores, or the utility of RQA-derived context stability variables had not previously been explored. Chapter 6 and Part 1 of Chapter 7 aimed to improve on the existing research through a series of analyses which systematically identified candidate variables for describing context stability from objective data describing the time of day at which medications were self-administered. Three variables (RANGE, W_rad_90, LAM_rad_15) were selected.

Demonstration of construct validity is a multifaceted task in the development of new measures, and is particularly important when constructs (such as habit) cannot be measured directly (Rebar et al., 2019). Chapter 6 presented a detailed examination of independent variable derivation and interpretation, but the predictive, convergent and discriminant validity of these variables used in combination has yet to be examined. Therefore, Part 2 of Chapter 7 began to address Thesis Aim 2, Objective 3, to evidence the utility of the defined objective metric of context stability in explaining variance in maintained and future behaviour, and in related constructs in evidence of predictive, convergent and divergent validity. The specific objectives addressed in Part 2 were:

1) To examine the amount of variance explained in concurrent behaviour by the objective metric of context stability (Thesis Aim 2, Objective 3.1);

2) To examine the amount of variance explained in participants' self-reported habitual automaticity for taking their medication by the objective metric of context stability (Thesis Aim 2, Objective 3.2);

3) To examine the amount of variance explained in perceived effort participants required to take their medication by the objective metric of context stability (Thesis Aim 2, Objective 3.3);

4) To examine whether distinct variance in concurrent behaviour is explained by reflective motivational constructs and the objective metric of context stability (Thesis Aim 2, Objective 3.4).

In the exploration of the first objective, the following hypotheses were devised:

H_1) The addition of RQA-derived variables (LAM_rad_15; W_rad_90) alongside a simple measure of context stability (RANGE), will be statistically and meaningfully increase the amount of variance explained in concurrent behaviour;

And, if H_1 was supported:

H_2) Combining the three measures into one, composite score can explain comparable variance explained in concurrent behaviour, to using three, independent variables.

To address the first objective, concurrent behaviour was selected as the outcome variable for H_1 and H_2; much of the existing literature which uses standalone measures of context stability has examined its relationship with behaviour in cross-sectional designs (e.g. Dean et al., 2021; Furman et al., 2021; Maher et al., 2021; Phillips et al., 2021; Pimm et al., 2015). To compare the effects of the objective metric with existing literature, concurrent behaviour was selected as the outcome for H_1 and H_2[10].

Other aspects of Thesis Aim 2, Objective 3 require an exploration of the convergent and discriminant validity of the objective metric of context stability.

Convergent validity can be supported by demonstrating statistical relationships between theoretically related constructs. In the case of context stability as a marker of habit, theory states that the mechanism by which habit increases the likelihood of a behaviour being performed is by increasing habitual automaticity (Gardner et al., 2020; Mazar & Wood, 2019) and the importance of habitual automaticity has been empirically demonstrated in relation to medication adherence behaviour (Arden et al., 2019; Hoo, Gardner, et al., 2019; Phillips et al., 2016). As each are fundamental features of established habitual behaviour, it follows that people with more contextually stable behaviour should also report greater habitual automaticity in ongoing behaviour. Second, a qualitative analysis of barriers and facilitators to medication adherence found that pwCF reported less *effort* was required to complete treatments when behaviours felt habitual or automatic (Arden et al., 2019). It follows that, in ongoing behaviour, people demonstrating more context stability should also report lower perceived effort.

Discriminant validity can be supported by demonstrating no statistical relationship between constructs that are not theoretically related to one another. Dual process models of behaviour (e.g.

---

[10] Furthermore, subject to replication of these effects, tests for the predictive validity of these measures was planned in prospective designs in Chapter 8.

Strack & Deutsch, 2004) posit that two systems are involved in the generation of behaviour: reflective processes (e.g. intending to do a behaviour on the basis of beliefs and attitudes), and automatic processes (e.g. habitually performing behaviour in response to a contextual cue). These are distinct systems. Therefore, discriminant validity could be demonstrated by showing little or no relationship between context stability and reflective motivational constructs, or if the relationship between context stability and behaviour is not changed when variance explained by reflective factors is also considered; i.e. the strength of the relationship between context stability and behaviour is not confounded by reflective motivation.

On the basis of this theory, and to address Thesis Aim 2, Objectives 3.2 to 3.4, the following three hypotheses were tested to understand the convergent and discriminant validity of the composite context stability measure:

H_3) Context stability would explain variance in participants' self-reported habitual automaticity for taking their medication;

H_4) Context stability would explain variance in perceived effort participants required to take their medication;

H_5) Context stability would show discriminant validity by explaining distinct variance in concurrent behaviour and being unrelated to reflective motivational constructs.

7.2.2. Methods

7.2.2.1 Variables

RANGE, W_rad_90, LAM_rad_15

Variables selected to represent context stability (RANGE, W_rad_90, LAM_rad_15) were calculated as described in Chapter 6. As in Part 1, variables for which high scores indicated lower context stability (RANGE, W_rad_90) were reverse coded by taking scores away from the maximum observed score, so that high scores indicated high context stability for all context stability variables.

Concurrent behaviour

Concurrent behaviour was the outcome variable for testing H_1 and H_2 and was calculated in the same way as in Part 1; this was percentage of days on which the participant took at least one medication per day, over the period from which the 14 observations used to calculate context stability were collected.

Composite context stability

In line with the critique in Chapter 4, the composite context stability score was constructed with a

percent scale (i.e., 0-100%), and in the absence of rationale for the importance of one component over another, components contributed equally to the overall score.

LAM_rad_90 was already represented as a percentage. RANGE and W_rad_90 were converted to percentage scores by dividing the reverse coded score by the maximum score on their respective scales and multiplying by 100. Percent scores on each domain was combined by calculating the mean average.

Effort

A single, self-report item with 5-point scale, asking participants to rate 'Taking my nebuliser treatments requires a lot of effort' (1=strongly agree, 5= strongly disagree). High scores indicated low effort.

Habitual automaticity

The 4-item Self-Report Behavioural Automaticity Index (SRBAI; Gardner et al., 2012) was used to rate how automatically participants felt they initiated taking their nebulised medications ('Taking my nebulised treatments is something I do…': 'without thinking'; 'without having to consciously remember'; 'automatically'; 'I start doing before I realise I'm doing it'; rated '1=Strongly agree; 5=Strongly disagree; scale range 4 to 20). High scores indicated high habitual automaticity.

Reflective motivation

The conceptual model for maintenance of adherence to medication developed by Arden and colleagues (Arden et al., 2021; Wildman et al., 2022; see also Michie et al., 2014) was applied in the current analysis (Figure 19). To account for the other sources of variation in adherence behaviour as posited by this behavioural model, measures of perceived necessity and concerns for nebuliser use were included in predictive models (H_1 to H_4). In H_5, these variables were utilised as measures of reflective motivational constructs.

Items from the Beliefs about Medicines Questionnaire (BMQ Nebuliser Specific, Horne et al., 1999) were adapted to ask 21 questions specifically about perceived necessity (n=7) and concerns (n=14) about taking nebuliser treatments for cystic fibrosis. Items were scored on a 5-point Likert scale from 1 (Strongly agree) to 5 (Strongly disagree).
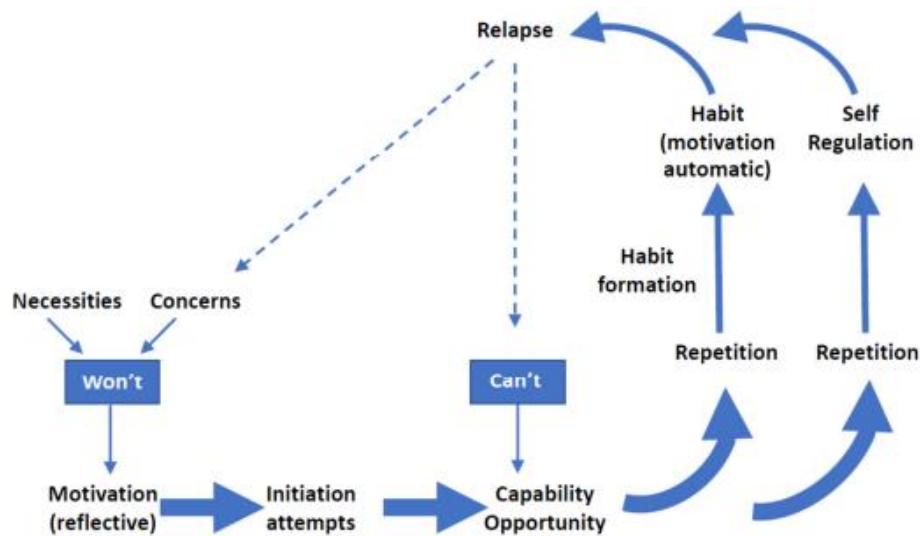
Figure 19
Conceptual model for the maintenance of adherence to medication (figure taken from Arden et al., (2021), Fig.1, pg. 4).

Other covariates

Since data from both the control and intervention arm participants were included in this analysis, the randomisation arm (intervention/control) was included as a covariate in linear regression models.

All measures were calculated with data closest to 12-months for all analyses because covariates (necessity, concerns) were measured at this timepoint (+/- 1 month).

7.2.2.2 Statistical analysis

Procedure for H_1 analysis

A multiple linear regression was conducted in three blocks. First, the baseline model was specified with randomisation arm, necessity and concerns predicting concurrent behaviour (block 1). Second, RANGE was added (block 2), followed by LAM_rad_15 and W_rad_90 (block 3; Figure 20). To examine statistical significance and meaningful change in predictive utility when LAM_rad_15 and W_rad_90 were added in block 3, magnitude and significance of adjusted R^2-change between block 2 and block 3 was examined.
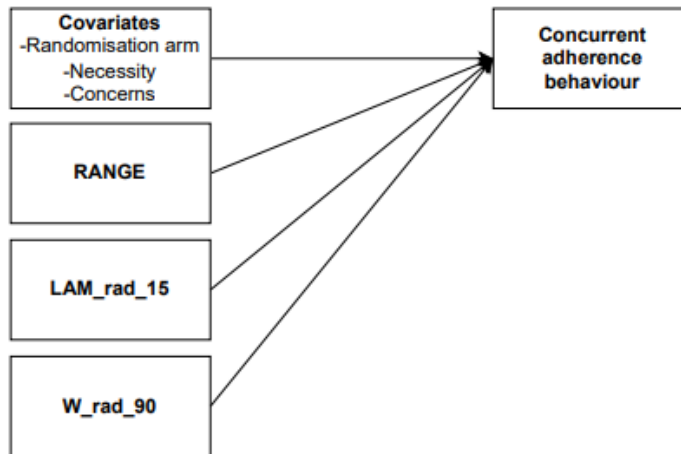
Figure 20
Model 1 testing H_1

Procedure for $H_{2,3,4}$ analyses

Two-stage multiple linear regression models were conducted to test the remaining hypotheses. First, the baseline model was specified with randomisation arm, necessity and concerns predicting concurrent behaviour (block 1). Second, composite context stability (CCS) was added (block 2). Adjusted R^2-change between block 1 and 2 was examined for statistical significance and effect size. Coefficients were also estimated and examined. To test H_2, this regression procedure was implemented with concurrent behaviour as the outcome variable (Figure 21).
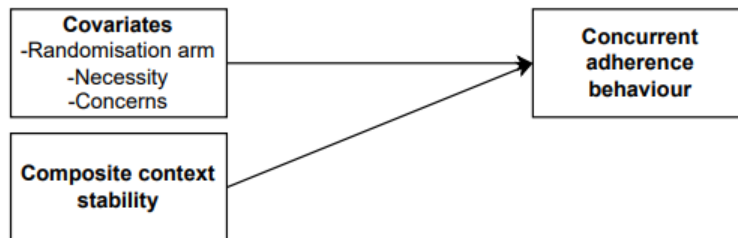
Figure 21
Model 2, predicting concurrent behaviour (H_2)

This procedure was repeated to test H_3 and H_4, substituting the outcome variable with SRBAI (Figure 22), and Effort (Figure 23) as the outcome variables, respectively.

131
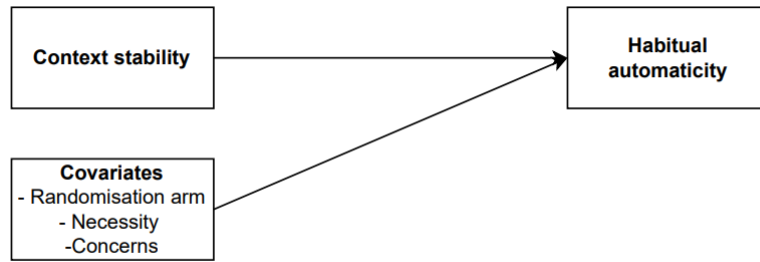
Figure 22
Model 3, predicting SRBAI (H₃)
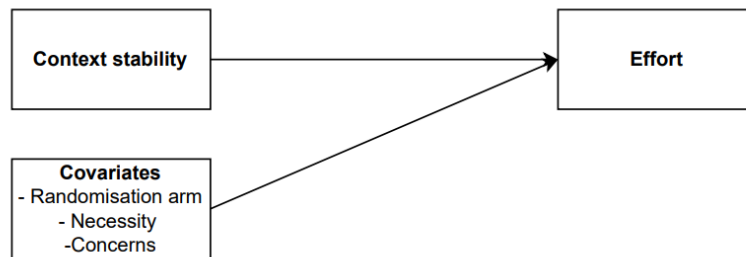
Figure 23
Model 4, predicting Effort (H₄)

Procedure for H₅ analysis

Bivariate correlations between context stability (i.e., automatic motivation) and necessity and concerns (i.e., reflective motivation) were explored, alongside the partial correlation between context stability and concurrent behaviour after accounting for variance explained by necessity and concerns.

All analyses were conducted in SPSS (v27.0.1.0).

7.2.3 Results

7.2.3.1 H₁ and H₂ descriptive statistics

Data was available for 493 participants. Descriptive statistics and bivariate correlations between all variables of interest for H₁ and H₂ are presented in Table 22 and Table 23 (see Appendix F, Table F4 for comparison with original trial sample). Mean concurrent behaviour was 67.52% (SD=32.01) and approximately half of the sample were in the intervention arm (n=254, 51.5%).

Table 22

Descriptive statistics for variables used in H₁ and H₂ analyses.

Variable	Mean (SD)	Median[LQ, UQ]	Min	Max
Concurrent behaviour (CB;%)	67.5 (32.0)	77.8 [41.2, 100.0]	4.69	100.0
Necessity (NEC)	3.67 (0.73)	3.70 [3.10, 4.10]	1.10	5.00
Concerns (CRN)	2.02 (0.55)	2.00 [1.60, 2.40]	1.00	3.60
RANGE (RNG; mins)	810.1 (311.8)	741.2 [565.5, 1106.0]	0.00	1384.2
LAM_rad_15 (LAM)	33.0 (26.1)	29.4 [12.9, 50.0]	0.00	100.0
W_rad_90 (W)	9.12 (1.29)	9.21 [8.54, 9.80]	0.00	12.1
Composite context stability (CCS; %)	54.1 (14.3)	52.9 [43.0, 63.6]	21.9	97.4

LQ=25% quartile; UQ= 75% quartile

Note: RANGE, W_rad_90 have been reverse-coded such that high scores indicate high context stability

Table 23

Bivariate correlations for H₁ and H₂

Variable	CB	IA	NEC	CRN	RNG	LAM	W	CCS
Concurrent behaviour (CB)	1	.18**	.36**	-.17**	.30**	.26**	.23**	.37**
Intervention arm (N (%); IA)		1	.10*	-.10*	.05.	-.03	.08	.03
Necessity (NEC)			1	.23**	.11*	.04	.11*	.11*
Concerns (CRN)				1	-.09	-.09*	-.02	-.10*
RANGE (RNG)					1	.33**	.42**	.81**
LAM_rad_15 (LAM)						1	.04	.78**
W_rad_90 (W)							1	.48**
Composite context stability (CCS)								1

7.2.3.2 H$_1$ results

H$_1$ was supported: randomisation arm, necessity and concerns explained 14.8% of the variance in concurrent behaviour, after adjustment for the number of variables in the model (p<.001; Table 24); the addition of RANGE in block 2 increased adjusted R^2 by 6.1% (overall adjusted R^2=20.9%; ΔF (1, 488)=38.64, p<.001) and the addition of LAM_rad_15 and W_rad_90 in block 3 further increased the overall adjusted R^2 to 24.9%. This indicates both a meaningful and significant increase in explained variance for concurrent behaviour when LAM_rad_15 and W_rad_90 were added to the model (adjusted ΔR^2= 4.0%; ΔF(2, 486)=14.19, p<.001). Approximately 10% of the variance in concurrent behaviour was explained by the selected context stability variables.

Table 24
R-square and significance of multivariable linear regression testing H$_1$

Model	Variables entered	Adjusted R^2	Δ adjusted R^2	Sig Δ
1	Intervention arm, Necessity, Concerns	14.8%	14.8%	<.001
2	Intervention arm, Necessity, Concerns, RANGE	20.9%	6.1%	<.001
3	Intervention arm, Necessity, Concerns, RANGE, W_rad_90, LAM_rad_15	24.9%	4.0%	<.001

7.2.3.3 H$_2$ results

H$_2$ was supported: a significant increase in adjusted R^2 from 14.8% to 25.1% was observed with the addition of the composite context stability score in block 2 (Δ adjusted R^2 =10.3%, ΔF(1,488)=67.9, p<.001; Table 25), compared to a 10.1% increase when entered as three independent variables in analysis 1 (note R^2 was adjusted for number of variables in both models, therefore this likely reflects very similar variance explained).

Table 25
R-square and significance of multivariable linear regression testing H$_2$

Model	Variables entered	Adjusted R^2	Δ adjusted R^2	Sig Δ
1	Intervention arm, Necessity, Concerns	14.8%	14.8%	<.001
2	Intervention arm, Necessity, Concerns, Context stability	25.1%	10.3%	<.001

All variables except concerns contributed significantly to the model explaining concurrent behaviour (Table 26). The largest standardised observed effect was context stability (β=.324, p<.001). An increase of 10% in context stability predicted an increase of 7.3% (95%CI=[5.5, 9.0]) in concurrent behaviour, after adjusting for the randomisation arm and the perceived necessity and concerns regarding taking nebulised medicines.[11]

Table 26
Summary of coefficient estimates from multivariable linear regression testing H_2

Variable	B	SE	95%CI	β	p
Intervention arm	8.70	2.52	[3.75, 13.7]	.136	<.001
Necessity	12.96	1.77	[9.49, 16.4]	.296	<.001
Concerns	-3.02	2.35	[-7.63, 1.60]	-.052	0.20
Context stability	0.73	0.09	[0.55, 0.90]	.324	<.001

B= unstandardised regression coefficient
β= standardised regression coefficient

7.2.3.4 H_3 descriptive statistics

Data was available for 492 participants. Descriptive statistics and bivariate correlations for variables included in the analysis of H_3 are presented in Table 27 and Table 28 (see Appendix F Table F5 for comparison with original trial sample). Mean SRBAI was 12.63 (SD=4.86). The sample demonstrated similar distributions in other variables as to those presented for H_1 and H_2.

Table 27
Descriptive statistics for variables used in H_3 analyses.

Variable	Mean (SD)	Median [LQ, UQ]	Min	Max
SRBAI	12.6 (4.86)	13.0 [8.00, 16.0]	4	20
Necessity	3.67 (0.73)	3.70 [3.10, 4.10]	1.10	5.00
Concerns	2.02 (0.55)	2.00 [1.60, 2.40]	1.00	3.60
Composite context stability (%)	54.1 (14.3)	52.9 [43.0, 63.7]	21.9	97.4

SRBAI= self-report habitual automaticity index; LQ=25% quartile; UQ= 75% quartile

---

[11] See Table 21 for an example of 2 participants with 10% difference in context stability score

Table 28

Bivariate correlations and descriptive statistics for H_3

Variable	SRBAI	IA	NEC	CRN	CS
Habitual automaticity (SRBAI)	1	0.09*	0.31**	-0.29**	0.25**
Intervention arm (N (%); IA)		1	0.10*	-0.10*	0.03
Necessity (NEC)			1	-0.23**	0.11*
Concerns (CRN)				1	-0.10*
Context stability (CS)					1

7.2.3.5 H_3 results

H_3 was partially supported. Randomisation arm, necessity, concerns and context stability explained 18.3% of the variance in habitual automaticity, after adjustment for the number of variables in the model ($F_{(4,487)}=28.51$, $p<.001$). Addition of context stability in block 2 increased adjusted R^2 by 3.8% from 14.5% in block 1 ($F_{(1,487)}=23.6$, $p<.001$). Context stability was a significant predictor after adjusting for the effects of intervention arm, necessity and concerns, but demonstrated small effects (Table 29); an increase of 10% in context stability observed closest to interval 12 predicted an increase of 0.68 (Ba 95%CI=[.41, .96], $p<.001$) on the SRBAI (scale 4-20).

Table 29

Summary of coefficient estimates from multivariable linear regression testing H_3

Variable	B	SE	95%CI	β	p
Intervention arm	.391	.400	[-.395, 1.177]	.040	.329
Necessity	1.587	.280	[1.036, 2.138]	.239	<.001
Concerns	-1.875	.373	[-2.607, 1.143	-.212	<.001
Context stability	.068	.014	[.041, .096]	.200	<.001

B= unstandardised regression coefficient
β= standardised regression coefficient

7.2.3.6 H_4 descriptive statistics

Data was available for 491 participants. Descriptive statistics and bivariate correlations for variables included in the analysis of H_4 are presented in Table 30 and

Table *31* (see Appendix F Table F6 for comparison with original trial sample). Mean effort was 3.19 (SD=1.27). The sample demonstrated similar distributions in other variables as to those presented for previous analyses.

Table 30

Descriptive statistics for variables used in H_4 analyses.

Variable	Mean (SD)	Median[LQ, UQ]	Min	Max
Effort	3.19(1.27)	3.00 [2.00, 4.00]	1	5
Necessity	3.67 (0.73)	3.70 [3.10, 4.10]	1.10	5.00
Concerns	2.02 (0.55)	2.00 [1.60, 2.40]	1.00	3.60
Composite context stability (%)	54.1 (14.3)	52.9 [43.1, 63.7]	21.9	97.4

LQ=25% quartile; UQ= 75% quartile

Table 31

Bivariate correlations and descriptive statistics for H_4

Variable	E	IA	NEC	CRN	CS
Effort (E)	1	0.07	0.17**	-0.416**	0.19**
Intervention arm (N (%); IA)		1	0.11*	-0.10*	0.03
Necessity (NEC)			1	-0.23**	0.10*
Concerns (CRN)				1	-0.10*
Context stability (CS)					1

7.2.3.7 H_4 analysis findings

H_4 was partially supported. Intervention arm, necessity, concerns and context stability resulted in an adjusted R^2 of 19.3% for variance explained in perceived effort for taking nebulised medications at interval 12 (N=491, F(4,486)=30.22, p<.001). Addition of context stability in block 2 increased adjusted R^2 from 17.5% to 19.3%, an increase of 1.8% (ΔF(1,486)=11.83, p<.001). Context stability was a significant predictor after adjusting for the effects of intervention arm, necessity and concerns, but demonstrated small effects (Table 32); an increase of 10% in context stability observed closest to interval 12 predicted an increase of .13 (Ba 95%CI=[.05, .20], p<.001) in effort (scale 1-5).

Table 32

Summary of coefficient estimates from multivariable linear regression testing H_4

Variable	B	SE	95%CI	β	p
Intervention arm	.058	.104	[-.146, .262]	.023	.329
Necessity	.116	.073	[-.028, .260]	.067	<.001
Concerns	-.883	.097	[-1.073, -0.693]	-.384	<.001
Context stability	.013	.004	[.005, .020]	.141	<.001

7.2.3.8 H_5 results

H_5 was partially supported. Data were available for n=493 participants. Bivariate correlations between context stability and reflective motivational factors showed statistically significant but

small effects (r=.11, p=.018; r=-0.10, p=.025 for necessity and concerns, respectively; Table 23). The correlation between context stability and concurrent adherence behaviour was medium strength (r=.365, p<.001, Ba 95%CI=[.289, .437]); the partial correlation between context stability and concurrent adherence behaviour after controlling for reflective motivation found a similar relationship (r=.347 (p<.001, Ba 95%CI=[.271, .424]), indicating minimal change and distinct variance explained.

7.2.4 Part 2 discussion

7.2.4.1 Overview of Part 2 findings

Thesis Aim 2, Objective 3 was to evidence the utility of the defined objective metric of context stability in explaining variance in maintained and future behaviour, and in related constructs in evidence of predictive, convergent and divergent validity. Chapter 7, Part 2 began to investigate this aim. A composite measure of context stability, derived from objective data demonstrated it could explain approximately 10% of the variance in objective medication adherence behaviour, after adjusting for reflective motivational factors. This is comparable with previous studies in medication adherence which used self-reported measures of habit with objective measures of medication adherence (Phillips et al., 2013; Phillips, Cohen, et al., 2016). The addition of RQA-derived variables (LAM_rad_15; W_rad_90) alongside a simple measure of context stability (RANGE), was found to statistically and meaningfully increase the amount of variance explained in concurrent behaviour (H_1 supported) and combining all three variables into one composite score resulted in comparable variance explained to using three, independent variables (H_2 supported in the exploration of Thesis Aim 2, Objective 3.1) meaning a more parsimonious. Further examination of the effects of context stability on medication adherence indicated that a 10% increase in context stability resulted in a 7.3% increase in medication adherence. Proportionally to the covariates in these models, context stability demonstrated the largest effects. This finding supports the Thesis Aim 2, that ongoing medication adherence behaviour could be explained by the degree of context stability that an individual is able to maintain. This is an important addition to the literature as this is the first demonstration of this effect using an objective measure of context stability, which is also conceptually and demonstrably delineated from past behaviour frequency. As this was demonstrated with data closest to the 12-month follow-up point, this analysis presents a cross-sectional understanding of the mechanisms of maintained (i.e., greater than 6 months) medication adherence behaviour.

Less clear support was observed for hypotheses three to five. The composite metric of context stability indicated divergent validity by explaining distinct variance in behaviour from reflective motivational factors, although a small correlation between context stability and reflective

motivational factors was observed (partial support for H_5 in the exploration of Thesis Aim 2, Objective 3.4).

Significant, but comparatively smaller effects than on concurrent medication adherence behaviour were observed for context stability in explaining participants' self-reported habitual automaticity for taking their medication (partial support for H_3 in the exploration of Thesis Aim 2, Objective 3.2). The association between SRBAI and objective measure of context stability (r=.25) was comparable to correlations in previous studies of medication adherence behaviour (e.g. Phillips et al., 2022; Phillips et al., 2021). However, the addition of context stability to the regression model increased the (adjusted) variance explained by 3.8%, and a 10% increase in context stability was estimated to result in a .68-point on the SRBAI scale. Compared to variance explained in behaviour, these were small effects. There is some precedent for this finding; whilst measures of context stability and SRBAI both capture aspects of habit, neither capture it completely and each captures a different aspect which contributes to habitual behaviour (Gardner & Tang, 2014; Rebar et al., 2019; Sniehotta & Presseau, 2012). For example, whilst the SRBAI was presented to the literature as a more parsimonious measure of habitual automaticity (Gardner et al., 2012) it does not capture the cue-dependency of the behaviour (Sniehotta & Presseau, 2012). This is supported by findings from Norman and Cooper (2011) which suggested that SRBAI and context stability could have additive explanatory effects in behaviour.

Similarly, significant but small effects were observed in the amount of effort participants perceived it took to take their medications (H_4 partially supported in the exploration of Thesis Aim 2, Objective 3.3); the addition of context stability increased the variance explained in effort by 1.8%, and a 10% increase in context stability predicted a .13-point increase on the effort scale (range one to five). Whist statistically significant, these were small effects. This finding may relate to the phrasing of the questions on the effort measure; participants were asked to give answers in relation to all of their nebulised medications, and for people with more than one prescribed medication, these questions may force them to conflate their experiences across multiple distinct behaviours. In contrast, the measure of behaviour (concurrent medication adherence) and context stability were both specific to the first medication of the day. This limitation also extends to the SRBAI findings as this scale also asks for responses in relation to all nebulisations taken throughout the day.

Across the models testing hypotheses one to four, the overall amount of variance explained was small (between 18.3% and 25.1%). However, this is a much larger effect than some have presented in the literature (e.g. Pimm et al., 2015 were only able to explain 3% of the variance in activity behaviour using self-reported context stability alone) and is comparable with the findings from

analyses of other explanatory models of health behaviour, particularly those which included measures of context stability (e.g. Furman et al., 2021; Norman & Cooper, 2011), including studies in medication adherence (Phillips et al., 2013; Phillips et al., 2016).

7.2.4.2 Limitations

Due to the present analysis being conducted in secondary data, the design was limited to examining the effects at a single timepoint within the study timeline. Doing so enabled the inclusion of potential confounding effects which may also explain ongoing medication adherence behaviour, according to the conceptual model developed by Arden and colleagues (2021). Unlike analyses in Part 1, this limited conclusions about the robustness of the findings to a single analysis and therefore generalisability of the findings and effects sizes should be tested in an external dataset.

7.3 General discussion

7.3.1 Review of Chapter 7 findings

Chapter 7 presented a systematic approach to identifying a metric of context stability, derived from objective medication adherence data, followed by a series of tests to examine construct validity. In Part 1, feature selection techniques were applied to 97 candidate variables, initially derived in Chapter 6; three variables were selected for their potential for additive predictive utility in the explanation of medication adherence behaviour. This was tested in Part 2; RANGE, W_rad_90 and LAM_rad_15 demonstrated incremental increases in the variance explained in ongoing medication adherence behaviour. This effect was maintained after creating a composite measure of context stability by combining all three scores into a single metric. The effects of the composite score on ongoing behaviour were statistically significant and non-trivial; compared to reflective factors, context stability demonstrated larger effects. Results were less clear in terms of demonstrating convergent and discriminant validity; methodological and conceptual limitations were discussed in the light of these findings, but they generally indicated the expected relationships between related constructs.

7.3.2 General limitations

A minimum of 14 observations was required to calculate this metric of context stability. This limits its use to situations in which sufficient data are available. However, as demonstrated in the critique presented in Chapter 4, this is a limitation of any measure of context stability that wishes to remain independent from frequency of past behaviour.

Compared to existing measures of context stability, relatively complex methods have been applied to derive the metric presented in this chapter; this has resulted in a more complex description of

context stability than previous studies that have used more simple statistics and univariable metrics. However, the advantage of RQA derived variables is the ability to capture recurrences in behaviour with conceptual rigour, and as demonstrated by the feature selection process, and increased explanatory power. Table 21 demonstrates clearly how the three metrics relate to the raw data from which they are derived. The process to select these key variables was systematic and clearly demonstrated that these variables could explain context stability in a theoretically meaningful way, without conflation with behavioural frequency, and with additive predictive utility.

Finally, the analyses presented in Chapter 7 are limited to cross-sectional designs. The objective metric of context stability should be tested in a prospective design to further support predictive utility.

7.3.3 Implications for the following Chapters

The following questions remain to further support the predictive validity of the composite context stability score, and to understand the role of context stability in behaviour:

1) Does context stability predict future behaviour?

2) Does context stability elucidate the relationship between past and future behaviour?

Chapter 8. The role of context stability in future behaviour

8.1 Background

The findings of Chapter 7 supported the construct validity of an objective measure of context stability in terms of convergent and divergent validity, and its ability to explain variance in concurrent medication adherence behaviour. The evidence demonstrated that understanding the context stability of medication adherence behaviour was important for understanding maintenance of adherence. However, the concluding remarks of the chapter highlighted that it was important to understand whether information about context stability could explain why people do behaviour in the *future*.

This distinction is fundamental in the behavioural scientists' goal to understand, and subsequently change, human behaviour; since the 1950's, psychological theories have attempted to explain the determinants of behaviour in order to present candidate mechanisms to change behaviour in the future (Hagger et al., 2021; Prestwich et al., 2014). Therefore, it is important not only to explain variance in behaviour as it happened, but also to explain why it happens in the future.

This is particularly relevant in the study of habits, as the function of habits is to make a behaviour more likely to occur in situations in which it has been performed previously (Rebar et al., 2018); experimental evidence has shown that this effect can override goals (Neal et al., 2011) and occurs in the absence of specific motivations (Neal et al., 2012). This feature has led habit researchers to claim the prediction of (aggregated frequency of) future behaviour is a key indicator of validity in measures of habit (Rebar et al., 2018), as well as indicating support for its importance in explaining the maintenance of behaviour (Kwasnicka et al., 2016).

One of the best known predictors of future behaviour is past behaviour (Ajzen, 2002, 2011); models predicting future behaviour have consistently demonstrated increased variance explained with the addition of past behaviour (e.g. see Ajzen, 2002, 2011; Conner & Armitage, 1998; Kor & Mullan, 2011). But as discussed in Chapter 4, past behaviour is of little use to behavioural scientists because it does not explain *why* future behaviour happens; authors have commented on the difficulty in interpreting the role of past on future behaviour as it is the product of multiple psychological constructs, and could also be resultant of measurement artefact (Ajzen, 2011). Habit may be able to explain at least some of this relationship. Chapters 6 and 7 developed an objective measure of

context stability to address the issue of measurement artefact, but the effects of context stability as a candidate construct in explaining the relationship between past and future behaviour in the maintenance of medication adherence remains to be tested.

Therefore, Chapter 8 addressed the remaining aspects of Thesis Aim 2, Objective 3 with an exploration of the role of context stability in the explanation of future behaviour. Two questions were posed at the end of Chapter 7:

1) Does context stability predict future behaviour?

2) Does context stability elucidate the relationship between past and future behaviour?

Whilst the first question is relatively straightforward to answer, for example, by examining the effects of context stability as an independent variable in a regression model predicting future behaviour (e.g. Dean et al., 2021; Furman et al., 2021; Judah et al., 2018; Maher et al., 2021; Phillips et al., 2021; Pimm et al., 2015), the second may require more consideration. In 1998, the seminal study by Ouellette & Wood (1998) demonstrated how context stability may be able to explain why past behaviour predicts future behaviour using moderation analysis; they concluded that only when contexts were stable did past behaviour significantly predict future behaviour. That is, Ouellette & Wood (1998) suggested that only when behaviour was (likely to be) habitual, would past behaviour predict future behaviour. Whilst widely cited and replicated and therefore useful to explore with the newly developed metric of context stability, this interaction effect may be oversimplifying the relationship; models of maintained behaviour (see Kwasnicka et al., 2016) including medication adherence (Phillips, Cohen, et al., 2016; Wildman et al., 2022) suggest that self-regulation may also play a key role and its role alongside habit is becoming increasingly recognised in that people can still act under the volition of self-regulation when they have a habit (Gardner & Lally, 2018; Gardner et al., 2020). Some habit researchers have even suggested that the ability to flexibly and synergistically switch between automatic (e.g. habit) and reflective systems (e.g. self-regulation) may actually be the key to successful behavioural maintenance (see Dunton et al., 2021). It may be that, rather than habit (or context stability as one aspect of habit) wholly explaining the conditions under which past behaviour predicts future, context stability may explain *some* of the reason for this relationship, not just when the relationship occurs. The latter hypothesis can be tested in a mediation model to understand whether some of the variance in the relationship between past and future behaviour is explained by context stability.

Therefore, the specific objectives of Chapter 8 were:

1) To examine the amount of variance in future adherence behaviour explained by the objective metric of context stability (Thesis Aim 2, Objective 3.5);

2) To examine the moderating effect of the objective metric of context stability in the relationship between past and future behaviour (Thesis Aim 2, Objective 3.6);

3) To examine the mediating effect of the objective metric of context stability in the relationship between past and future behaviour (Thesis Aim 2, Objective 3.7).

8.2 Methods

8.2.1 Variables

Context stability

Context stability was calculated as described in Chapter 7:

CONTEXT STABILITY= (RANGE_PERCENTAGE + LAM_rad_90 + W_rad_90_PERCENTAGE)/3

For analyses 1 and 2, the 14 observations prior to interval 12 were utilised to construct RANGE, LAM_rad_90 and W_rad_90 scores. For analysis 3, the 14 observations closest to interval 12 but later than interval 6 were used.

Past behaviour

Past behaviour was calculated as adherence to at least one medication per day over the period from which the 14 inhalations used to calculate context stability were collected (i.e., closest to interval 12 for analyses 1 & 2, and between interval 6 and 12 for analysis 3). For example, a participant with 14 inhalations collected over 20 days prior to interval 12 would have past adherence equal to 70% (14/20*100).

Future behaviour

Future behaviour was calculated as the adherence to at least one medication per day during the first 14 days of interval 12. For example, a participant with 7 inhalations collected over the first 14 days in interval 12 would have concurrent adherence equal to 50% (7/14*100).

8.2.1.4 Covariates

Necessity and concerns

Items from the Beliefs about Medicines Questionnaire (BMQ, Horne et al., 1999) were adapted to ask 21 questions specifically about perceived necessity (n=7) and concerns (n=14) for taking nebuliser treatments for cystic fibrosis. Items were scored on a 5-point Likert scale from 1 (Strongly agree) to 5 (Strongly disagree). For analyses 1 and 2, necessity and concerns measured at the end of

the trial (i.e., 12 +/-1month) was used; for analysis 3, necessity and concerns measured at baseline was used.

Randomisation arm

As in Chapter 7, the arm which participants were randomised to (either usual care (reference category), or intervention) was included.

8.2.2 Statistical procedure

8.2.2.1 Analysis 1

Analysis 1 aimed to examine the effect of (objectively measured) context stability in the prediction of future adherence behaviour. A two-stage multiple linear regression model was conducted (Figure 24). First, the baseline model was specified with randomisation arm, necessity and concerns predicting future adherence behaviour (block 1). Second, context stability was added (block 2). R^2-change between block 1 and 2 was examined for statistical significance and adjusted R^2-change examined for effect size. Regression coefficients were also estimated and examined for effect size.
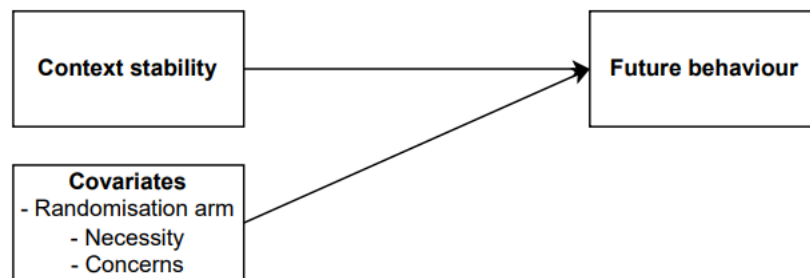
Figure 24
Conceptual model tested in analysis 1

8.2.2.2 Analysis 2

Analysis 2 aimed to examine the moderating effect of (objectively measured) context stability on the relationship between past and future behaviour. A moderation analysis was implemented to test the conceptual model presented in Figure 25. The moderation effect was estimated after adjusting for the effects of covariates (randomisation arm, reflective motivational factors), and implemented using the PROCESS© macro (v4.0, Hayes, 2022). Continuous variables were mean-centred prior to analysis.
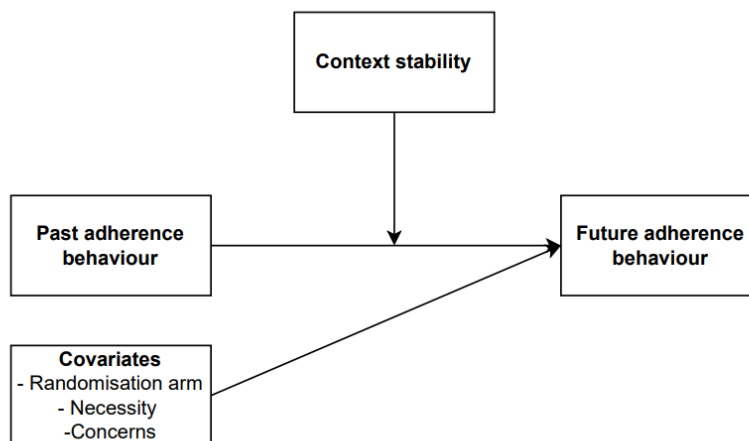
Figure 25
Conceptual model of the moderation effect tested in analysis 2

8.2.2.3 Analysis 3

Analysis 3 aimed to examine the mediating effect of (objectively measured) context stability on the relationship between past and future behaviour. This was explored using regression analysis to test the conceptual model presented in Figure 26. The effects were estimated after adjusting for covariates (randomisation arm, reflective motivational factors). The indirect effect was tested using a percentile bootstrap estimation approach with 5000 samples. The analysis was implemented using the PROCESS macro (v4.0, Hayes, 2022).
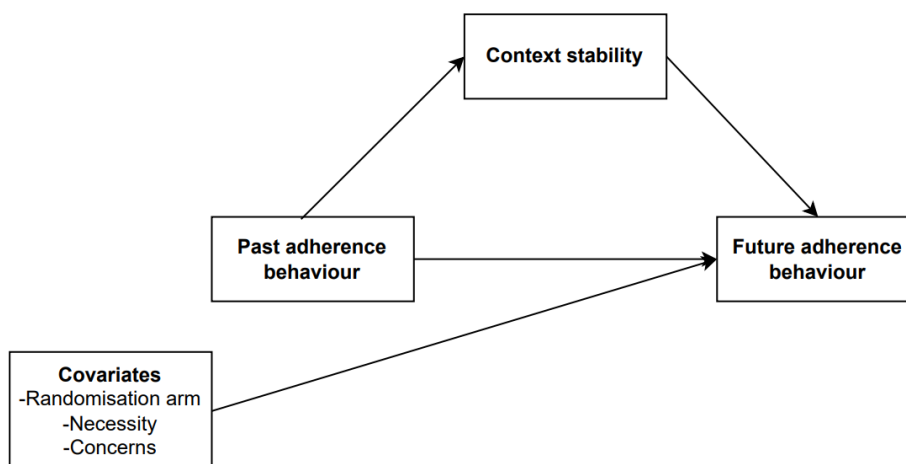
Figure 26
Conceptual model of the mediation effect tested in analysis 3.

All analyses were implemented in SPSS (v27.0.1.0).

8.3 Results

8.3.1 Analysis 1 & 2 descriptive statistics

Data was available for n=472 participants for analyses 1 and 2. Descriptive statistics and bivariate correlations between variables of interest are presented in Table 33 and Table 34 (see Appendix F Table F7 for comparison with original trial sample).

Table 33
Descriptive statistics for variables in analysis 1 and 2.

Variable	Mean (SD)	Median [LQ, UQ]	Min	Max
Future adherence (%)	62.2 (38.3)	78.6 [21.4, 100.0]	0.00	100.0
Necessity	3.67 (0.73)	3.70 [3.10, 4.10]	1.10	5.00
Concerns	2.02 (0.55)	2.00 [1.60, 2.40]	1.00	3.60
Past adherence (%)	68.0 (32.1)	77.8 [42.4, 100.0]	4.70	100.0
Composite context stability (%)	54.5 (14.3)	53.4 [43.4, 63.9]	21.9	97.4

LQ=25% quartile; UQ= 75% quartile

Table 34
Bivariate correlations for variables in analyses 1 and 2.

Variable	FA	IA	NEC	CRN	PA	CS
Future adherence (FA)	1	0.17**	0.36**	-0.19**	0.80**	0.33**
Intervention arm (N (%); IA)		1	0.09*	-0.08	0.17**	0.02
Necessity (NEC)			1	-0.22**	0.36**	0.11*
Concerns (CRN)				1	-0.17*	-0.10*
Past adherence (PA)					1	0.36**
Context stability (CS)						1

N=472
*p<.05
**p<.01

8.3.2 Analysis 1 findings

Addition of context stability in block 2 significantly increased variance explained in future adherence behaviour; adjusted R^2 increased by 8.2%, from 15.0% in block 1 (future adherence predicted by randomisation arm, necessity, concerns) to 23.2% in block 2, when context stability was added (n=472; $\Delta F_{(1, 467)}$=51.0, p=<.001). Context stability was a significant predictor of future behaviour (b=0.78, 95%CI= [0.57, 1.00]; Table 35); a 10% increase in context stability predicted 7.8% increase in future adherence behaviour.

Table 35

Model coefficients predicting future behaviour

Variable	B	SE	95%CI	p
Intervention arm	10.1	3.11	[3.97, 16.2]	.001
Necessity	15.4	2.18	[11.1, 19.7]	<.001
Concerns	-5.73	2.91	[-11.5, -0.02]	.049
Context stability	0.78	0.11	[0.57, 1.00]	<.001

8.3.3 Analysis 2 findings

A total of 64.5% of the variance in future behaviour was predicted by the moderation model (n= 472; $F_{(6,465)}$=140.9, p<.001), but moderation of the effects of past on future behaviour by context stability was very small and non-significant; addition of the interaction term resulted in ΔR^2=0. 07% ($\Delta F_{(1,465)}$=0.98, p=.323) and the interaction term did not significantly predict future behaviour (b=.002, 95%CI=[-.003, .008]; Table 36).

Table 36

Moderation model coefficients predicting future behaviour.

Variable	B	SE	95%CI	p
Intervention arm	3.06	2.16	[-1.18, 7.29]	.157
Necessity	3.95	1.57	[0.86, 7.05]	.012
Concerns	-2.67	2.00	[-6.59, 1.25]	.181
Past adherence	0.88	0.04	[0.81, 0.96]	<.001
Context stability	0.13	0.08	[-0.03, 0.29]	.104
Past adherence x Context stability	.002	.003	[-.003, .008]	.323

8.3.4 Analysis 3 descriptive statistics

Data was available for n=439 participants for analysis 3. Descriptive statistics and bivariate correlations are presented in

Table *37* and Table 38 (comparisons with full trial dataset are available in Appendix F, Table F8).

Table 37

Descriptive statistics for variables used in analysis 3.

Variable	Mean (SD)	Median[LQ, UQ]	Min	Max
Future adherence (%)	68.6 (34.6)	78.6 [42.9, 100.0]	0.00	100.0
Necessity	3.67 (0.73)	3.70 [3.30, 4.10]	1.40	5.00
Concerns	2.08 (0.53)	2.10 [1.70, 2.40]	1.00	4.40
Past adherence (%)	71.1 (33.2)	85.7 [50.0, 100.0]	0	100.0
Composite context stability (%)	55.2 (14.2)	54.3 [44.7, 64.5]	28.4	97.4

Table 38

Bivariate correlations and descriptive statistics for variables in analysis 3.

Variable	FA	IA	NEC	CRN	PA	CS
Future adherence (FA)	1	.13**	.26**	-.09	.69**	.31**
Intervention arm (N (%); IA)		1	-.02	.11*	.14**	-.01
Necessity (NEC)			1	-.14**	.29**	.05
Concerns (CRN)				1	-.15**	-.09
Past adherence (PA)					1	.25**
Context stability (CS)						1

N=439
*p<.05
**p<.01

8.3.5 Analysis 3 results

The total effect model explained 48.8% of the variance in future adherence behaviour (F(4, 434)=103.23, p<.001). The direct effect of past on future behaviour was significant (b=0.66, 95%CI=[58.06, 73.20], p<.001). There was a significant indirect effect of past on future behaviour, through context stability however this was a very small effect (b=.04, Ba 95%CI=[0.02, 0.07]; Figure 27).
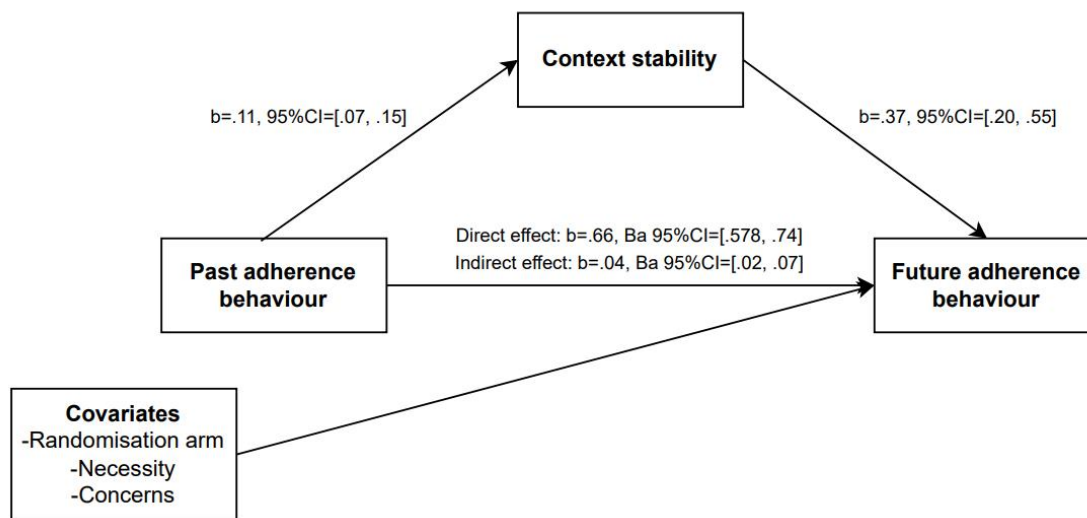
Figure 27

Mediation model with direct and indirect effects predicting future adherence behaviour.

8.4 Discussion

The aim of Chapter 8 was to understand the role of objectively measured context stability in future behaviour, in terms of its ability to predict future behaviour and explain the relationship between past and future behaviour. Three analyses were presented, with the findings that objectively measured context stability explains a significant amount of variance in future adherence behaviour

(results of analysis 1 in the exploration of Thesis Aim 2, Objective 3.5) but demonstrated little utility in explaining why past behaviour predicts future behaviour (results of analyses 2 and 3, in the exploration of Thesis Aim 2, Objective 3.6 and 3.7, respectively).

An additional 8.2% of the variance in future behaviour was explained by objectively measured context stability in analysis 1; this is similar in magnitude to that observed in Chapter 7 when context stability was added to a model predicting concurrent medication adherence behaviour. Furthermore, the predicted effect of a 10% increase in context stability on adherence was similar for both concurrent and future behaviour (b=0.73, 0.78, respectively). The cross-sectional study presented in Chapter 7 improved on previous studies that have examined the role of context stability in concurrent behaviour by ensuring complete dissociation between frequency of behaviour and the metric representing context stability, as well as controlling for the effects of other reflective behavioural constructs. The replication of this effect in a prospective design, further supports that context stability has some explanatory utility in understanding medication adherence behaviour. The similarity of the effect size across cross-sectional and prospective designs is coherent, as in both cases the data were sampled from latter end of the 12-month data collection phase, therefore both concurrent and future behaviour were conceptually sampled from the maintenance phase of the observation period; it is unlikely that behaviour drastically changed for many participants during this period of the study. The loss in predictive utility in the prospective design could potentially be accounted for by individuals whose behaviour changed between the two observation periods. According to theory, truly habitual behaviour can only continue in the presence of the contextual cue with which it is associated (Mazar & Wood, 2019); disruptions to usual routines could reduce or even eliminate cue encounters and henceforth result in the termination of habitual behaviour (e.g. Wood et al., 2005). Whilst these cues may have been consistent and acted upon by a participant in the past, they may be removed or change in the future, and objective measurement of context stability through temporal stability of behavioural initiation in the past alone may not be enough to anticipate the disruption to the cue, and therefore the decline in future behaviour frequency. In summary, the findings indicate that context stability, as operationalised in analysis 1, may be useful in explaining behaviour that is maintained, but may not be as useful for predicting change in behaviour. The sensitivity of the objective metric of context stability could be further explored to examine whether it can detect disruptions to cues and subsequently, whether these signals predict change in future behaviour.

Whilst not trivial, the observed effect sizes for context stability were small compared to those observed for past adherence in the prediction of future adherence behaviour (e.g., the model tested

in analysis 1 did not include past behaviour and the total R^2 indicated that 23.2% of the variance in future behaviour was explained, compared to the model tested in analysis 2 which did include past behaviour and had a total R^2=64.5%). The additional variance explained could suggest that past behaviour may account for other behavioural constructs that were not captured in the conceptual models presented. Recent models of behaviour maintenance suggest that the ability to flexibly and synergistically switch between automatic (e.g. habit) and reflective systems may actually be the key to successful behavioural maintenance (see Dunton et al., 2021). It is possible that the skill and ease with which people can make switches between these two systems is captured by past behaviour but is not by the measurement of the reflective and automatic constructs included in these analyses, alone. Second, the objective measure of context stability used in these analyses may not capture all aspects of habit, or even context stability itself. For example, the inclusion of past behaviour increases variance explained in future behaviour, even in models capturing behavioural automaticity (Gardner et al., 2012) and additive effects of context stability and measures of habitual automaticity have been observed in previous studies (Norman & Cooper, 2011), suggesting that these may capture different aspects of the habit construct.

These considerations may also account for the absence of a moderation effect, as well as explaining some of the significant but very small mediation effect. If it is true that one or multiple unmeasured constructs are important for the maintenance of adherence behaviour in the future, then it follows that the relationship between past and future behaviour would not only be observed when context stability is high. Furthermore, failure to fully capture all aspects of habit (i.e., automaticity was not captured by context stability) and perhaps the way that context stability functions alongside other reflective behaviours, may account for the very small mediation effect. It is also possible that small effects were observed because of the difference in the length of time between measures of past and future behaviour across the different analyses. For analysis 1 and 2, past behaviour for many participants immediately preceded future behaviour and was concurrent with context stability, however the mediation analysis specified past behaviour as observations over the last 14 days of interval 6, and future behaviour as the first 14 days of interval 12, leaving a break of 168 days between measurements. Correlations were weaker between past and future behaviour in this analysis (r=0.69, compared to r=0.80 in analyses 1 and 2), and total variance explained was noticeably lower in the mediation model, compared to the moderation model (R^2=48.8% in the mediation model, compared to R^2=64.5% in the moderation). Past behaviour is a better predictor of future behaviour when it is measured more proximally in time (Dunton et al., 2021). However, the measurement of past and future behaviour as implemented in analysis 3 was necessary to ensure

accuracy in the temporal arrangement of the mediator between the independent and dependent variables, as well as ensuring participants with a range of adherence levels (i.e., low, medium and high) were included in the analysis, as a minimum of 14 observations over the 168-day period was necessary to calculate context stability.

8.5 Conclusions

The findings of Chapter 8 explored the support that objectively measured context stability can explain significant and non-trivial amounts of variance in future behaviour. The research presented in Chapters 6 to 8 have cumulatively explored Thesis Aim 2 which asked whether the time of day at which people take their medications can be used to derive an objective metric of context stability for use in habit research. The utility of the objective metric of context stability has been demonstrated in analyses commonly used in habit research, but the final empirical chapter of this thesis should seek to demonstrate its use in habit research with novel empirical designs and in relation to clinically applied scenarios.

Chapter 9. Demonstrating the utility of an objective metric of context stability in applied clinical research

9.1 Chapter overview

Chapter 9 addresses Thesis Aim 2, Objective 4, which was to demonstrate the application of the metric of context stability developed in the previous chapters to applied clinical research questions.

9.2 Background

The review in Chapter 3 highlighted that there is a need for more research that examines the effectiveness of theory-informed habit interventions in supporting maintenance of medication adherence behaviour. Habit is hypothesised to facilitate maintenance of behaviour by bypassing reflective decision-making processes when self-control is diminished, enabling people to engage in behaviours as they have in the past with minimal cognitive expense (Gardner et al., 2020). However, research remains uncertain on how long it takes to form a habit, indicating there is likely large differences between both behaviours and individuals (Lally et al., 2010). This poses a challenge for habit intervention developers as it makes intervention design and evaluation of effectiveness difficult; questions about how long someone might need to develop a habit, who might need more support than others, when to withdraw additional support and when to evaluate effectiveness must all be considered carefully. These issues may benefit from behaviour, population and intervention-specific research. Chapter 9 presents an example of how the context stability metric developed in the previous chapters may be applied to address some of these issues in the context of two clinical scenarios.

9.2.1 Overview of clinical context

The behaviour and population of interest to the present thesis is medication adherence behaviour in people with cystic fibrosis (CF). The ACtiF Programme (see Chapter 5) developed a complex intervention, called the CFHealthHub Intervention, which aimed to improve adherence to nebulised medication in people with CF (Wildman et al., 2021, 2022). The CFHealthHub Intervention was delivered via a digital interface and in one-to-one intervention sessions with a trained healthcare professional. The roll-out of CFHealthHub has already begun across the UK, with the CFHealthHub Data Observatory in which aspects of the intervention are being implemented in clinical care of patients in the UK (CFHealthHub, 2023).

Two health economic analyses were conducted to evaluate the cost-effectiveness of the CFHealthHub Intervention; the within-trial analysis evaluated intervention cost-effectiveness within the one-year follow-up period of the trial with the intervention delivered as per the trial protocol. This evaluation indicated an incremental cost of £865.91 per patient in the first year, compared to usual care (Wildman et al., 2021). The intervention was expected to become cost-saving when the analysis was extended over the lifetime period but in an NHS setting where resources are limited, clinical teams would benefit from knowledge of some indicators that enable them to identify patients who are more likely than others to require intervention, to know when this intervention is most needed and if there any specific strategies that could be built into the intervention which would enable participants to achieve higher adherence to their prescribed medications.

9.2.2 Clinical scenario 1

Some participants in the ACtiF trial had complex health and social needs and were disengaged from their CF team during the trial; the programme discussion and conclusions suggested that these participants may require an intervention with focused attention on motivations for behaviour and complementary health and social care support. However, many others in the trial did engage with the intervention and CF team but had varying degrees of success in achieving maintained adherence behaviour. It may be useful to clinical teams to know, of those who initially engage in medication adherence behaviour, after how long should expectation of failure to maintain become low, and what characteristics of the behaviour might associate with likelihood of failure to maintain.

Habit theory would expect that those who demonstrate higher context stability would be more likely to form or already have habits (Gardner et al., 2020) and therefore more likely to maintain adherence behaviour for longer. As such, the first two research questions explored in Chapter 9 were:

Research Question 1 (RQ1): After how many days should those most at risk of failing to maintain medication adherence be identified?

Research Question 2 (RQ2): Is the likelihood of failing to maintain adherence behaviour in the future lower for those with greater context stability?

9.2.3 Clinical scenario 2

All analyses conducted in the development of the context stability metric (Chapters 6 to 8) were performed on a single daily repetition of behaviour, by exploring the effects of context stability and the ability to adhere and maintain adherence to the first medication of the day. However, many participants in the ACtiF trial were prescribed more than one medication to take on each day. This is

typical for people with CF as a mucolytic drug is usually prescribed alongside antibiotics, perhaps also among other prescribed medicines and therapies (Ratjen et al., 2015). Therefore, a specific challenge for people designing interventions in this specific population and behaviour in that the inhalation of nebulised drugs may need to be repeated multiple times a day. To my knowledge, currently there exists no research or recommendations for people with CF on how best to distribute their inhalations throughout their day to maximise adherence.

The habit literature may offer some potential, theory-informed strategies to support people with CF in the management of multiple doses of medication to take each day. Recent insights distinguish between the different roles that habit can adopt in the facilitation of behaviour, namely that of 'habitual instigation' and 'habitual execution' of behaviour (Gardner, 2015; Gardner et al., 2016). When the role of a habit is to facilitate the 'instigation' of behaviour, it facilitates the initial generation of action by automatically by-passing decision-making processes and creating an impulse to perform the behaviour when an associated cue is encountered. By contrast, when the role of a habit is to facilitate the 'execution' of behaviour, it facilitates the performance of behaviour through all of the sub-actions required to complete the behaviour in full (Gardner et al., 2016; Phillips & Gardner, 2016). For example, both instigation and execution habits may have a role in making the first cup of tea in a day; the impulse to make a cup of tea may be instigated by coming downstairs in the morning and walking into the kitchen. This could cue the person to walk straight to the kettle and switch it on. Having instigated the first step, the person is facilitated to completion of the behaviour by an execution habit which enables the completion of the rest of the tea-making routine with minimal cognitive effort (e.g., getting the cup out of the cupboard, getting the teabag, and so on). Importantly for habit researchers, evidence has shown that it is the instigation of behaviour which is most predictive of future behaviour and therefore likely the most important focus for intervention designers (Gardner, 2022; Gardner et al., 2016; Phillips & Gardner, 2016). When multiple repetitions of behaviour on a single day are required, as for many people with CF prescribed multiple doses of medication, patients may be more successful at taking more than one dose of medication if they take it with their first dose of the day, and if that first dose is taken in conditions which facilitate habit (i.e., in a contextually stable environment).

The analyses for Clinical Scenario 2 were designed to demonstrate, in the sample of patients from the ACtiF trial, whether patients took their 2nd dose of medication more frequently if taken with a contextually stable (i.e., more likely to be or become habitual) 1st dose of medication, than if they took their 2nd dose in a separate, second session which occurred later in the day.

Research Question 3 (RQ3): Are people more likely to take a 2nd dose of medication if they take immediately after the 1st dose of the day?

Research Question 4 (RQ4): Were participants more likely to take a 2nd dose of medication if the 1st dose was more contextually stable?

9.3 Methods

9.3.1 Variable derivation for the analysis of RQ1 and RQ2

Context stability

For each day of each participant's involvement of the trial, the nth inhalation of each day was labelled (i.e. if three inhalations were recorded on a given day for a participant, each consecutive record was labelled the 1st, 2nd and 3rd respectively). Labelled inhalations were then grouped by 1st, 2nd, 3rd inhalation for each participant. For all 1st inhalations of the day for each participant, the first 14 of observations were selected, starting from the first day of their involvement in the trial. Context stability of those 14 observations was calculated using the method described in Chapter 7. Then, a rolling window was applied in steps of one day (i.e., window one included observations one to 14; window two included observations two to 15, moving chronologically through their days involved in the trial) and the context stability calculation was repeated for each window. This resulted in a longitudinal sequence of data points describing each participant's context stability for their first inhalations of each day. This process was then repeated for the 2nd and 3rd inhalation records for participants who were prescribed three doses of medication for the duration of their involvement in the trial.

Initiation of adherence maintenance

There is no clinically agreed definition of maintenance of adherence (Vrijens et al., 2012b), nor maintenance of behaviour. Derived from the Transtheoretical model of Health Behaviour Change (Prochaska & DiClemente, 1982), the metric of 6-months to indicate maintenance has been widely cited and utilised (e.g. as in the review in Chapter 3). However, this is somewhat arbitrary and, as discussed in the introduction, researchers remain unsure how long it takes to form a habit. Furthermore, a single cut-off defined in days rather than in terms of behaviour is not suitable for answering the research questions posed in the present chapter. Therefore, for the purposes of this chapter, participants were considered to have initiated adherence on the first observation that the participant took their first medication of the day for three consecutive days. For example, if they took their medication on days one, two and three, the first day of initiation of adherence maintenance was classified as day three.

Failure to maintain adherence

Participants were considered to have failed to maintain adherence on the first observation that the participant missed their first medication of the day for three consecutive days, after having previously initiated adherence maintenance (see above). For example, Figure 28 shows an example participant who initiated maintenance of adherence on day 16, indicated by the left vertical red line (i.e. took their first medication on days 14, 15 and 16) and failed to maintain adherence on day 107, indicated by the right vertical red line (i.e. missed their medication on days 105, 106 and 107).
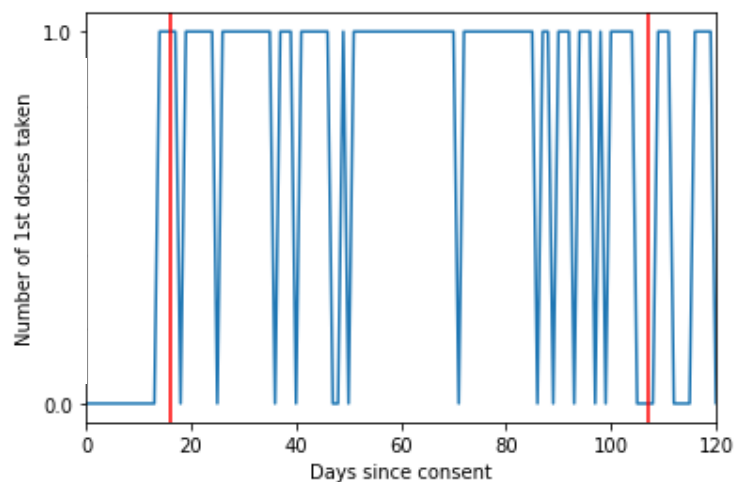
Figure 28
Example of a participant who initiated maintenance of adherence on day 16, indicated by the left vertical red line (i.e. took their first medication on days 14, 15 and 16) and failed to maintain adherence on day 107, indicated by the right red line (i.e. missed their medication on days 105, 106 and 107).

If a participant never failed to maintain adherence (i.e., they maintained adherence to the end of the trial), the analysis censored their data at the last day of their involvement in the trial. For example, Figure 29 shows an example participant who initiated maintenance of adherence on day four of the trial (indicated by the vertical red line) and subsequently did not fail to maintain adherence to their first medication of the day; whilst they missed one or two consecutive days at various points, they did not fail to take medication for three consecutive days. Their data was censored at the last day of their involvement of trial (day 159).
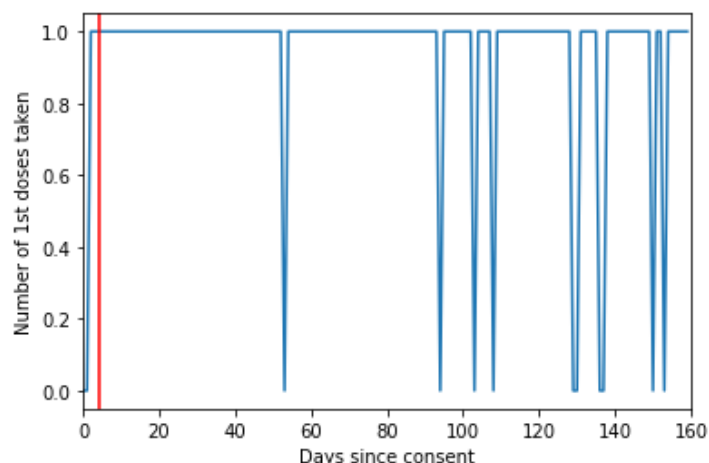
Figure 29
Example of a participant who initiated maintenance of adherence on day four of the trial (indicated by the vertical red line) and subsequently did not fail to maintain adherence to their first medication of the day. Their data was censored at the last day of their involvement of trial (day 159).

Context stability during the period of maintenance

Context stability during maintenance of the first dose of the day was calculated by taking the mean *context stability* of the first dose of the day (see above), on the days between which the participant *initiated adherence maintenance* and *failed to maintain adherence* to the first dose of medication.

Covariates

As in previous chapters, the effects of context stability were estimated after adjustment for baseline perceived necessity and concerns for taking their nebulised medications, and the randomisation arm to which each participant was allocated.

9.3.2 Variable derivation for analysis of RQ3 and RQ4

All variables derived to answer Research Question 3 were aggregated variables which describe behaviour at the level of the participant. Analyses were conducted only on participants prescribed three doses of medication for the duration of their involvement in the trial. This was to ensure that adherence and context stability for the 1st, 2nd and 3rd dose of the day was comparable between participants.[12]

---

[12] Three doses of medication was the most frequent prescription among participants prescribed more than one dose per day thus making this the largest, comparable subsample of data to use.

Likelihood of taking a 2nd dose, given a 1st dose was taken

For each participant, the days on which a 1st dose of medication was taken were selected. Then the percentage of 2nd doses taken on the selected days was calculated to represent the likelihood of taking a 2nd dose of medication, given that a 1st dose was taken.

Mean number of minutes between 1st and 2nd doses

For each participant and for each day on which a 1st and 2nd dose of medication was taken, the number of minutes between doses was calculated. The number of minutes between doses for each participant was then aggregated into the mean number of minutes between 1st and 2nd doses.

Standard deviation of minutes between 1st and 2nd doses

For each participant and for each day on which a 1st and 2nd dose of medication was taken, the number of minutes between doses was calculated. The number of minutes was then aggregated for each participant into the standard deviation of minutes between 1st and 2nd doses.

Mean context stability of the 1st dose

The mean was calculated for the sequence of context stability data points derived from 1st doses taken, for each participant (see Context Stability, above).

Mean context stability of the 2nd dose

The mean was calculated for the sequence of context stability data points derived from 2nd doses taken, for each participant (see Context Stability, above).

9.3.3 Statistical analyses

9.3.3.1 RQ1 analysis

Kaplan-Meier curves were used to identify the rate of failure to maintain adherence to the first medication of the day over time. Participants that did not fail to maintain adherence for the duration of their participation in the study (beyond initiation of adherence), were included in the analysis up to the point of their discontinuation from the study, at which point their data was censored. The Kaplan-Meier curves were plotted up to the number of days at which data were available for greater than 100 participants, to avoid misinterpretation of estimates yielded from small samples.

9.3.3.2 RQ2 analysis

Cox regression analysis was used to examine the effects of *context stability during the period of maintenance*, on the number of days before *failure to maintain adherence*. Estimates of the effects of mean context stability were adjusted for randomisation arm, baseline necessity and concerns about taking nebulised medications.

9.3.3.3 RQ3 and RQ4 analysis

Associations with likelihood of taking the 2nd dose of medication on each day were explored using Pearson's correlations. *Likelihood of taking a 2nd dose given a 1st dose was taken* was correlated with *mean number of minutes between 1st and 2nd doses*, the *standard deviation of minutes between 1st and 2nd doses* taken, the *mean context stability of the 1st dose* and the *mean context stability of the 2nd dose*.

9.4 Results

9.4.1 Descriptive statistics for sample included in analyses for RQ1 and RQ2

A total of 408 participants were included (67.1% of the original RCT sample). Descriptive statistics for the variables included in the analyses for RQ1 and RQ2 are presented in Table 39. Analysis sample characteristics are presented beside full sample characteristics in Appendix F, Table F9.

Table 39
Descriptive statistics of variables included in the Kaplan-Meier and Cox regression analyses.

	Mean (SD)	n (%)
Intervention arm		223 (54.7%)
Baseline necessity	3.70 (0.72)	
Baseline concerns	2.04 (0.54)	
Mean days of maintenance	205.3 (193.8)	
Mean context stability to 1st medication of the day (%)	55.3 (11.0)	

9.4.2 RQ1 analysis: after how many days should those most at risk of failing to maintain medication adherence be identified?

The Kaplan-Meier survival plot is presented in Figure 30. This indicates that 25% of participants that failed to maintain adherence to their first medication of the day, defined as failure to take medication on three consecutive days, had done so by day 35 (SE=3.32 days); 50% of those who failed to maintain medication adherence had done so by day 126 (SE=10.20 days). In this sample, 22.8% of participants (n=93) that initiated maintenance did not fail to maintain for the duration of their involvement in the trial.
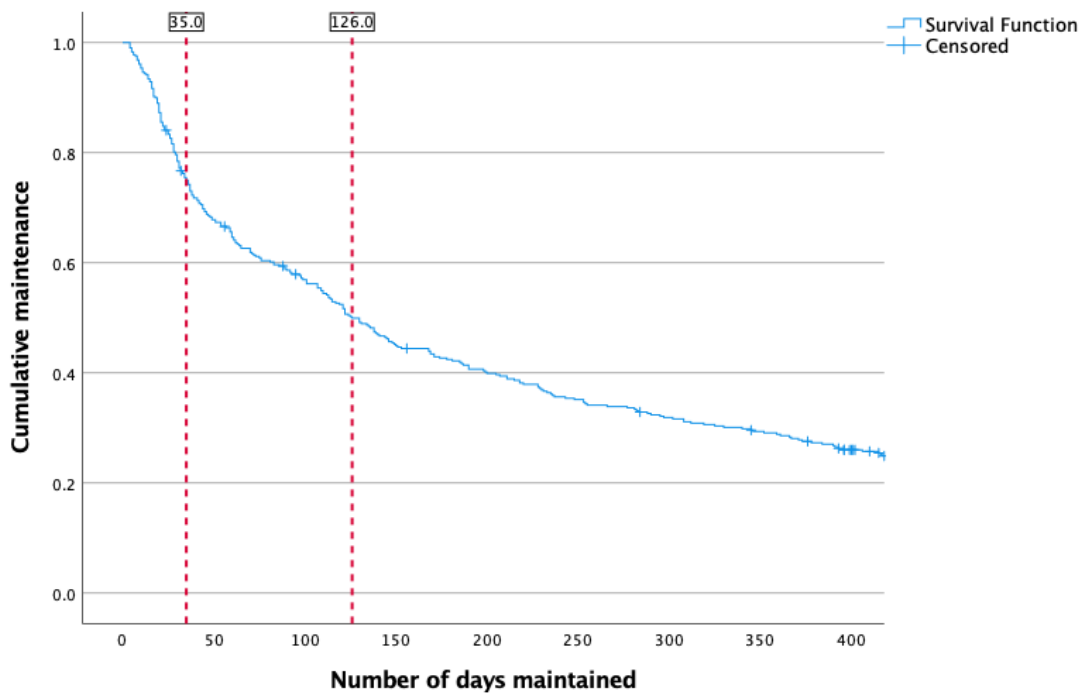
Figure 30
Kaplan-Meier curve of the expected number of days of maintenance

9.4.3 RQ2 analysis: Cox regression to explore whether the likelihood of failing to maintain adherence behaviour in the future is lower for those with greater context stability?

Table 40 presents the results of the adjusted Cox regression analysis to understand the effects of context stability on the number of days until failure to maintain medication adherence. The results indicate, after adjusting for the effects of randomisation arm, necessity and concerns, people with higher context stability had 5% [95%CI= 0.93, 0.96] less probability of failing to maintain medication adherence (p<.001), thereby indicating a highly significant but small effect.

Table 40
Results of the Cox regression analysis to understand the effects of context stability on time to failure of maintenance.

	B (SE)	Hazard Ratio [95% CI]	p-value
Intervention arm	-0.36 (0.12)	0.70 [0.56, 0.88]	.002
Baseline necessity	-0.47 (0.08)	0.63 [0.53, 0.74]	<.001
Baseline concerns	0.21 (0.11)	1.24 [1.00, 1.53]	0.046
Mean context stability	-0.06 (0.01)	0.95 [0.93, 0.96]	<.001

Given the effects of intervention arm on time to survival, estimates for time to fail to maintain medication adherence were reproduced to understand the number of days within which participants with no additional support or contact with a healthcare professional (i.e., the control condition) would be expected to fail to maintain. Results indicated that 25% of participants that

initiate adherence maintenance but receive no additional support will fail to maintain within 31 days (SE=3.21), and 50% will fail to maintain within 109 days (SE=18.16).

9.4.4 Descriptive statistics for the sample included in analyses for RQ3 and RQ4

A total of 67 participants were included in analyses of RQ3 and RQ4 (11.0% of the original RCT sample; 95.7% of the 70 participants prescribed three doses of medication for the duration of their involvement in the RCT). Descriptive statistics for the variables included in the analyses for RQ3 and RQ4 are presented in Table 41. Analysis sample characteristics are presented beside full sample characteristics in Appendix F, Table F10.

Table 41
Descriptive statistics for variables included in analyses for RQ3.

	n	Mean (SD)	Min	Max
Adherence to 1st dose	67	55.4% (35.7)	0.0%	100%
Adherence to 2nd dose given 1st taken	67	55.3% (32.2)	0.0%	100%
Mean mins between 1st and 2nd dose	65	359.9 (214.0)	5.2	849.6
SD mins between 1st and 2nd dose	63	216.8 (87.7)	2.1	419.5
Mean context stability to 1st dose*	50	54.9% (9.8)	36.75%	75.5%
Mean context stability to 2nd dose	50	54.0% (10.2)	29.6%	91.1%

*Given a 2nd dose was recorded on that day

9.4.5 RQ3 analysis: are people more likely to take a 2nd dose of medication if they take it immediately after the 1st dose of the day?

Non-significant effects were observed for the association between the percentage of 2nd doses of medication taken and the mean and standard deviation of the number of minutes between the first and second dose (Table 42).

Table 42
Associations with likelihood of taking 2nd dose of medication (% 2nd doses taken).

	n	Correlation with % 2nd doses taken	95% CI
Mean mins between 1st and 2nd dose	65	0.12	[-0.13, 0.35]
SD mins between 1st and 2nd dose	63	0.05	[-0.20, 0.30]
Mean context stability to 1st dose	**50**	**0.59**	**[0.38, 0.75]**
Mean context stability to 2nd dose	50	0.17	[-0.11, 0.43]

9.4.6 RQ4 analysis: were participants more likely to take a 2nd dose of medication if the 1st dose was more contextually stable?

A large effect size was observed between the percentage of 2nd doses taken and the context stability of the first dose of medication taken (Table 42; r=0.59, 95%CI=[0.38, 0.75]). Figure 31 shows that, in this sample of participants, it was unlikely for participants to take a high percentage of their second doses of medication if they did not have high context stability for their 1st dose of the day (i.e. few data points in the top left quadrant of Figure 31). However, given that neither the mean context stability for the 2nd dose of medication nor the mean or standard deviation of the number of minutes between doses was associated with the percentage of 2nd doses taken. This suggests that having high context stability for the first medication of the day makes participants more likely to take the 2nd medication, but not necessarily because the participant is taking it at the same time as the 1st.
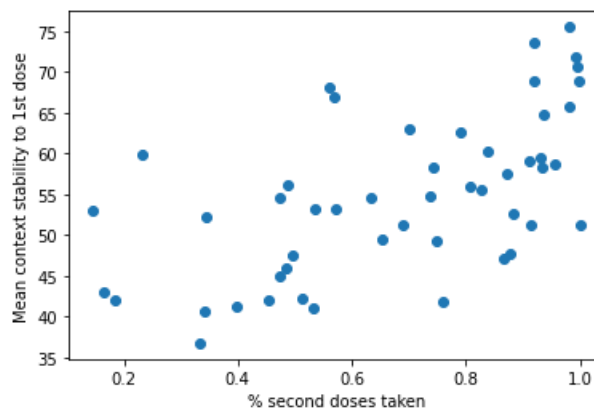
Figure 31
Relationship between the percentage of second doses taken against context stability for the first dose of the day (r=0.59, 95%CI=[0.38, 0.75]).

9.4.7 Post-hoc exploration of the observed effects in the analysis of RQ3 and RQ4

The descriptive statistics presented in Table 41 may give some insight into the observed effects found in the analysis of RQ3. The mean number of minutes between the 1st and 2nd doses taken was 359.9 minutes (SD=214.0). Among those that took more than 90% of their second dose of the day (n=14), the average time between doses was 415 minutes (SD=221.5 minutes) and not all people with very few average minutes between 1st and 2nd doses were highly adherent to their 2nd dose (see Figure 32). This indicates heterogeneity in the medication-taking sessions participants were choosing throughout the day; some participants appeared to take a single medication in the first session of the day and two in the second, whereas others appeared to take two in the first and one in the second, and some may have taken medications over three separate sessions.
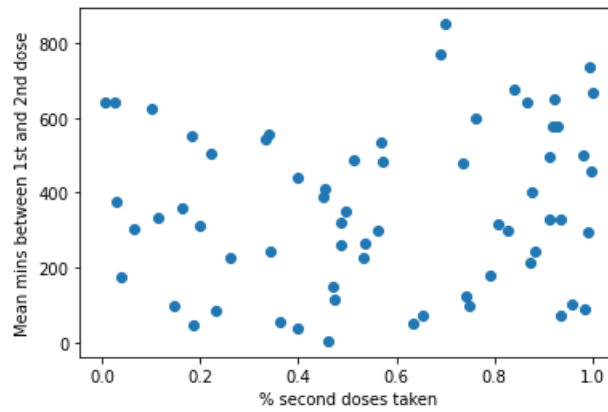
Figure 32

The relationship between the % second doses taken and the mean number of minutes between 1st and 2nd doses (r=0.12, 95%CI=[-0.13, 0.35]).

This is exemplified by the following cases, presented in Table 43. All three participants had greater than 90% adherence to their 2nd medication of the day but took it separately from the 1st dose; on average, participants took their 2nd medication between approximately five and 12 hours after the first. All participants had high context stability for the 1st medication of the day but comparably lower context stability for the 2nd medication of the day, despite taking a high percentage of their 2nd medications.

Table 43

Example cases that achieved high adherence to 2nd medication of the day without taking the dose with a contextually stable first dose of the day. The top panel presents summary statistics for the three cases. The bottom panel presents plots of the times of day at which each of the 1st, 2nd and 3rd inhalations were taken, by day of their involvement in the RCT.

Case	% 2nd doses	Mean minutes between 1st and 2nd dose (SD)	Mean context stability to 1st medication	Mean context stability to 2nd medication
A	99.5%	736.3 (158.9)	70.6%	60.6%
B	92.0%	576.6 (205.4)	73.5%	52.8%
C	99.1%	297.1 (107.8)	71.9%	57.9%

A
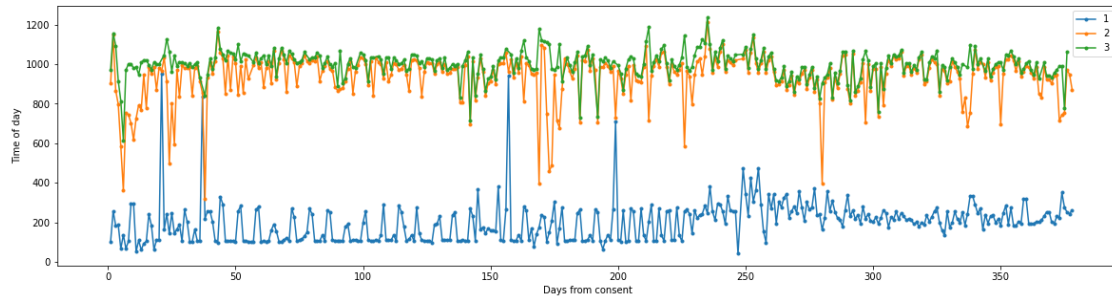
B
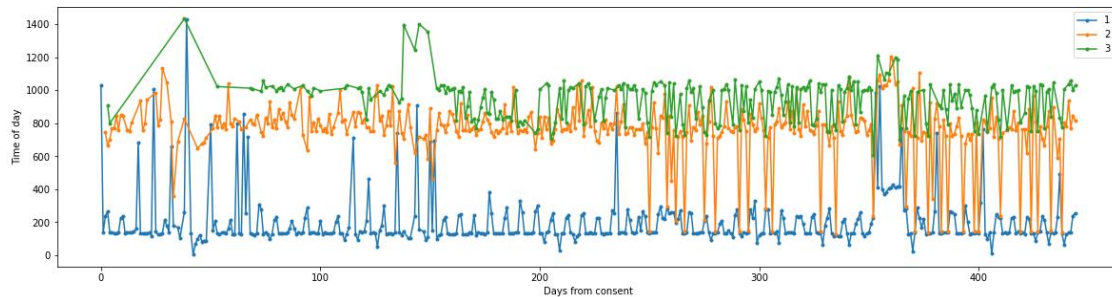
C
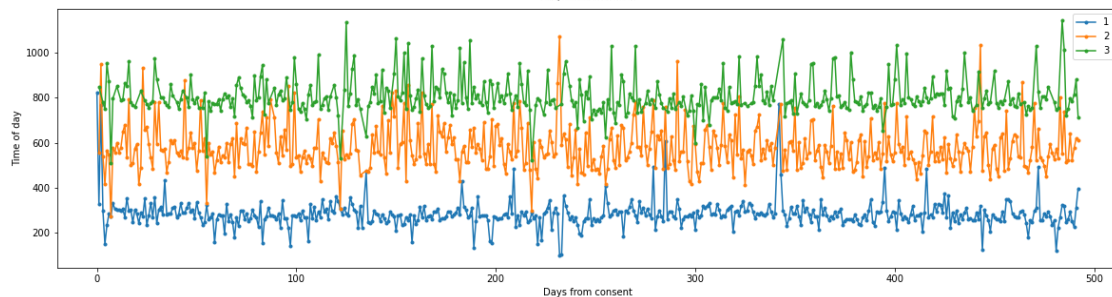
9.5 Discussion

9.5.1 Summary of Chapter 9 findings

Thesis Aim 2, Objective 4 was explored in Chapter 9 with a series of analyses which investigated the utility of an objective metric of context stability metric in answering a series of clinically applied research questions. The first research question (RQ1) asked after how many days should those most at risk of failing to maintain medication adherence be identified. The findings indicated that, once people with CF in this trial did initiate adherence (defined as 3 consecutive days of taking their first medication of the day) clinicians should expect the 25% of participants most likely to fail to maintain to have done so by day 35 (SE=3.32 days). The second most at-risk quartile of participants (25-50%)

failed to maintain adherence between days 35 (SE=3.32 days) and 126 (SE=10.2 days), a period of approximately 91 days. During this time, the Kaplan-Meier curve visibly reduced in steepness, indicating a reduction in the rate of failure to maintain adherence among the remaining participants. Furthermore, for those without any additional support, failure to maintain was expected for 25% of participants within 31 days (SE=3.21), and 50% were estimated to fail to maintain within 109 days (SE=18.16). The second research question (RQ2) asked whether the likelihood of maintaining behaviour in the future was higher for those with greater context stability. The Cox regression analysis indicted that context stability was a highly significant predictor of failure to maintain medication adherence (p<.001); high context stability reduced likelihood of failure to maintain by approximately 5% (HR=0.95 [0.93, 0.96]), after adjusting for other covariates in the model. The third research question (RQ3) asked if participants that were prescribed three doses of medication per day were more likely to take a 2nd dose of medication if they took it immediately after their 1st dose; non-significant effects were observed for the association between the percentage of 2nd doses of medication taken and the mean and standard deviation of the number of minutes between the first and second dose. Finally, RQ4 asked if participants were more likely to take a 2nd dose of medication if the 1st dose was more contextually stable; higher context stability for the 1st dose of medication was associated with higher adherence to the 2nd dose (r=0.59, 95%CI=[0.38, 0.75]). The findings for each of these research questions are discussed, below.

9.5.2 Strengths and limitations of RQ1 and RQ2 findings

The findings from RQ1 indicate that the riskiest period for failure to maintain adherence to the first medication of the day is in the first 109 to 126 days after starting to use a device that can record and monitor adherence behaviour. Unlike previous chapters, this analysis used data from the beginning of the trial. The key conceptual difference is that use of data from the start of the trial enabled an exploration of behaviour following a change in the usual medication-taking scenario, triggered by the receipt of a new device to administer medication and the awareness that adherence behaviour was being recorded. This might be considered a critical change in circumstance which could, for some participants, function as a trigger to initiate a change in behaviour (e.g. see discussion in Verplanken et al., 2018 and white coat effect, see Podsadecki et al., 2008). For clinicians and intervention designers, it could be useful to know for how long an increase in adherence caused by this change remains in effect. The findings most usefully highlight an initially steep but plateauing Kaplan-Meier curve. This describes that, as time went on, the likelihood of failing to maintain medication adherence became smaller and occurred somewhere in the region of one to four months after receipt of initiation of use of the device.

The results from RQ2 complement the identification of a critical period by identifying characteristics of behaviour which might indicate higher risk of failure to maintain within that time; context stability affected the rate of failure to maintain such that those with higher context stability were less likely to fail (p<.001, HR= 5%). Whilst a small effect, this is consistent with findings from previous chapters in which context stability significantly explained 8.2% of the variance in future behaviour (see Chapter 8). RQ2 findings provide supporting evidence of the protective attributes afforded by habit formation and/or established habits in action; behaviours performed in more contextually stable environments are theoretically more likely to be or become habitual and were found to be more likely to be maintained.

For the purposes of these analyses, initiation of and failure to maintain medication adherence was defined as three consecutive 3 days of adherence, and non-adherence, respectively. Whilst the six-month threshold suggested by the Transtheoretical Model of behaviour change (Prochaska & Di Clemente, 1982) has been widely used to define the boundary of 'maintenance', the present analysis took a different perspective which was more suited to the research question. This is a novel definition and future work could look to explore how this metric of maintenance relates to other characteristics of behavioural repetition and definitions of maintenance. Prochaska and Di Clemente's definition still requires some definition of *what* people are doing at 6 months and the definition used in the present analysis may be a useful place to begin; the analysis estimated that approximately 40% of the participants had not missed three consecutive days of medications at 180 days (~6 months) after initiating the behaviour but how this relates to, for example, how many other medications the participants were able to manage or the clinical effects on lung health are yet to be explored.

9.5.2 Strengths and limitations of RQ3 and RQ4 findings

Together, RQ3 and RQ4 findings suggested that people who demonstrated higher adherence to their 2nd dose of medication did not necessarily achieve this by initiating the 1st dose in a stable context and taking the 2nd dose in the same session of behaviour. The mean number of minutes between doses was 359.9 minutes (SD=214.0) and, exemplified by three case studies, those with highest adherence to their 2nd dose on days when a 1st dose was taken did not fall at the lower end of this distribution.

These findings are contrary to what was expected but a few factors may have converged to produce the results of this analysis. The sample was comprised of participants that were prescribed three doses of medication per day; this was a design choice made first to ensure that more than one dose of medication was prescribed to participants included in the sample whilst ensuring comparability

between participants in terms of the expected behaviour and second to maximise the sample available for analysis (i.e. because three prescribed doses of was the most frequent regimen observed in the full trial sample because it is common for people with CF to be prescribed 2 doses of antibiotics with one dose of DNAse). However, this design choice resulted in two features of the analysis sample which may have affected the results: first, a mixture of intervention arm (n=35, 52.3%) and control arm participants were included in the analysis but neither group was explicitly instructed on how to distribute their inhalations throughout the day. Some clinicians in the trial were known to prefer their patients to wait 30 minutes between a dose of antibiotic and DNAse which may explain why three medication sessions are apparent in plots for participants B and C (Table 43). Furthermore, Participant A appears to prefer to take one dose in the morning and two doses in the evening. These example plots represent three of the 14 high adherers (i.e., whose adherence to 2nd dose of the day exceeded 90%) identified in the sample selected for this analysis. As such, it appears that the evidence presented in the analysis of RQ3 and RQ4 may be limited more by absence of participants performing two doses in the morning and one in the evening in the first instance, meaning it was not possible to assess the effects of this compared to alternatives using this dataset. As discussed in the introduction to this Chapter, there is theoretical rationale to assume that taking the 2nd dose immediately after a contextually triggered first dose should ensure higher adherence to the second dose. To adequately examine this hypothesis it may be necessary first to assess the acceptability of aiming to complete more than one inhalation in the first session of the day to understand why many of the participants in this sample opted not to do so, followed by an interventional study in which some participants are explicitly asked to take two doses in the morning and one in the evening, and the other participants are asked to do the reverse. N-of-1 methods (Kwasnicka & Naughton, 2020) supported by electronic data capture and context stability metrics such as those developed in this thesis may be usefully employed to explore this idea in more detail.

Furthermore, higher context stability for the 2nd dose was not significantly associated with higher adherence to the 2nd dose (r=0.17, [-0.11, 0.43]), but higher context stability for the 1st dose of medication was associated with higher adherence to the 2nd dose, with a large effect size (r=0.59, 95%CI=[0.38, 0.75]). Given the small sample, these findings should be interpreted with caution but they may be an indication of reduced sensitivity of the objective context stability metric for behaviours performed later in the day. Using the three examples presented in Table 43, these were individuals with high adherence to their 2nd dose but their plotted data noticeably exhibits greater variability in the times that their evening doses of medication were initiated, which was also reflected in their context stability scores for the 2nd dose, compared to the first. One possible explanation for the individuals having high adherence to their 2nd dose of medication is that they

were performing that dose with habits of equal strength to their 1st dose of the day, cued by another routine, but that the evening routine happened at more variable times than that of the morning. To date there is little research and mixed findings as to whether behaviours are more frequent and more habitual in the morning compared to the evening and this may in some part relate to the quality and focus of the existing literature. Some studies have demonstrated more frequent behaviour in the morning; for example Maher et al., (2021) found that self-reported occasions of physical activity were most frequent in the morning compared to afternoon or evening, but the authors do not comment on whether context stability was higher for the morning (see also Domke et al., 2019 who demonstrated more frequent action plan enactment in the morning compared to later in the day). Other studies have more explicitly tested the relationship between morning vs evening performance of behaviour, frequency of behaviour and habit strength; Fournier et al., (2017) demonstrated that behaviours performed when waking in the morning achieved automaticity (i.e. became habitual) more quickly than those performed prior to bedtime. However, this finding was attributed to cortisol levels and the outcome was assessed using the Self-Report Behavioural Automaticity Index (SRBAI; Gardner et al., 2012). These findings do not directly relate to the findings of RQ3 and RQ4 as they do not shed light on the variability in temporal initiation of behaviours in the mornings compared to the evenings. The most comparable existing research with respect to the present findings was conducted by Phillips et al., (2021). They compared, in people prescribed two doses of medication per day, the frequency and variability in times they took their morning medications compared to their evening. They found that morning medications were taken more frequently, but variability in the times that morning and evening medications were taken were not statistically different. The authors comment that a trend indicated slightly higher variability for the evening pill and, because of their small sample (n=51 participants). Furthermore, closer inspection of the data included in this study shows that the values for variance are extremely small in both groups and thus the findings may be resultant of a homogeneous sample of highly adherent and habitual patients; on average participants were more than 80% adherent to their medications and average variance in the morning was 6.26 minutes (SD=4.12) and 6.54 minutes (SD=6.07) in the evening. Arguably, these figures indicate that both morning and evening behaviours were performed habitually in response to specific routines for most participants but it is possible that the difference in frequency may relate to the absence of cues in the evening on some days and more consistent cue encounters in the mornings. Due to unlabelled inhalations and few participants with comparable regimen, there is no way to conduct a more in-depth analysis of this issue in the ACtiF dataset but this preliminary finding is an important avenue for future research to explore.

9.6 Conclusions

In summary, the objective metric of context stability developed in Chapters 6, 7 and 8 was utilised in a series of clinically applied research questions to address Thesis Aim 2, Objective 4. The discussion highlights that the conclusions from the findings, particularly for RQ3 and RQ4, are limited in generalisability but the analysis usefully highlighted some important considerations and directions for future research. These are discussed in Chapter 10.

Chapter 10. Discussion

The primary aim of this thesis was to explore the role, conceptualisation and measurement of habit, habitual behaviour and context stability in adherence to medication for people living with long-term conditions. The final chapter of the thesis includes a discussion and reflections on the thesis findings along with an appraisal of the strengths and limitations of the work presented and recommendations for future research.

10.1 Summary of Thesis Aims, findings and contributions to habit research

There were two core aims of this thesis: i) to examine the conceptualisation of 'habit' and the implementation and utility of 'habit formation' in interventions to support the maintenance of medication adherence, and ii) to explore whether the time of day at which people take their medications can be used to derive an objective metric of context stability for use in habit research.

With respect to the first aim, the review conducted in Chapter 3 found little evidence that researchers designing interventions to improve medication adherence in people with long-term conditions use contemporary habit theory. Whilst use of the term 'habit' was identified in the description of 15 different intervention studies, few referenced key contemporary theory or literature and only six explicitly used context-dependent repetition as a behaviour change technique. The review has been published and on invitation, was presented to habit researchers at the EHPS international conference in 2022. There are two key contributions to habit research resulting from the exploration of this aim and dissemination of the review findings, discussed below.

The first key contribution was the exposure of the needs of, and opportunities afforded by, habit research conducted on medication adherence behaviour. To date, a body of work conducted by Allison Phillips and colleagues that has broadly embedded itself within the behaviour change literature (e.g. Phillips et al., 2013, 2022; Phillips, 2011; Phillips et al., 2016; Phillips et al., 2021) and latterly, the work conducted by Hoo and colleagues which been mainly embedded in the domain of CF research (e.g. Hoo et al., 2017; Hoo, Gardner, et al., 2019; Hoo, Wildman, et al., 2019) have begun to establish the theoretical basis of the role and contributions of habit in medication adherence behaviour. However, most habit theory applied to interventional research has been focused in other health behaviours, most prominently exercise (e.g. see Phillips, 2020). The review drew attention to the fact that habit has the potential to support issues with medication adherence and that quality research is both feasible (afforded by the availability of devices that can objectively measure the behaviour) and needed (indicated by the scarcity of studies identified for the review).

The second key contribution from the review is lessons learned during the conduct of BCT coding in 'habit interventions'; the review discussion highlights that even among studies using 'context-dependent repetition' as a BCT, it was never used alone, was always with some combination of behavioural regulation BCTs and was often difficult to discern. This was reflective of work conducted by Gardner and Rebar (2019); BCTs used with context-dependent repetition were not unanimous between studies and coding was challenging in some cases due to ambiguity in BCT definitions. For example, coding context-dependent repetition and action planning together was often unnecessary as the former was necessarily a subtype of the latter. Similarly, Michie et al's (2013) BCT taxonomy necessarily required 'Behavioural Practice/Rehearsal' to be coded when context dependent-repetition was coded but this was often unnecessary as usually the participant was not practicing. I presented these among other reflections in Bratislava at the EHPS 2022 Habit Special Interest Group meeting. My learning points from the conduct of this review served as a starting point for a discussion on what exactly constitutes a 'habit formation' intervention, and whether researchers can articulate this in an accurate way using existing taxonomies, or if more is needed to enable transparent and reproducible dissemination of habit intervention content.

One characteristic of the research conducted in the later Chapters of the thesis highlights the need to clarify the bounds of what constitutes a 'habit intervention'. Sixty-seven percent of the original trial sample was included the analysis of RQ1 and RQ2 in Chapter 10, meaning around a third did not initiate maintenance as it was defined, and/or contribute 15 inhalation records for the duration of their engagement with the trial. This was similar in other chapters, in which a portion of the sample were not included in analyses because they did not engage in adherence behaviour enough for context stability to be calculated. Conceptual models of habit, such as those described in Gardner et al., (2020) and Gardner & Lally (2018) outline a series of prerequisite stages which must be achieved in order for a habit to form. For example, both Gardner and Lally's (2018) generic model of habit formation, and Arden et al's (2021) model of habit formation for the CFHealthHub intervention to improve medication adherence both require that an individual engages in intention formation and action initiation stages before they can engage with habit formation. The extent to which habit interventions should include BCTs to support these prerequisite stages of habit formation is under debate but in the meantime, it is important for researchers to acknowledge that not all participants may be ready to engage in context-dependent repetition and either explicitly exclude those from their target population or include appropriate content to support them toward that stage. This is important as, if not recognised or acknowledged, it may lead to underestimated effectiveness of an intervention which identifies as a 'habit intervention' (i.e., delivers context dependent repetition as a BCT) because it is not appropriate for the population that receives it.

With respect to the second thesis aim (i.e. to explore whether the time of day at which people take their medications can be used to derive an objective metric of context stability for use in habit research), the work conducted in Chapter 4 and Chapters 6 to 9 demonstrated that a novel and objective metric of context stability was able to explain small but significant variation in the adherence behaviours of people with CF, with comparable effects to that observed in external research. The key contributions of this work lie in the progression of the discourse surrounding the objective measurement of habit and context stability and are best discussed in relation to the strengths of specific findings. These are presented below, followed by my reflection on the scope of the contributions from these chapters in terms of conceptual and methodological limitations and recommendations for future research.

10.2 Strengths of this research

The derivation of an objective measure of habit could be construed as a simple matter of taking some observations of the time at which a participant initiates a behaviour of interest and aggregating it into a single score which describes the variance of the available data (i.e. as in Phillips et al.'s 2021 conceptualisation of objective 'habit strength'). Or, derivation of an objective measure of habit could be construed as a matter of capturing variance in the time of behavioural initiation and multiplying it by the frequency of the behaviour (i.e. as in the objective measure of habit inspired by the BFCS which was developed by Hoo, Wildman, et al., 2019).

These two studies and the methods described therein served as the initial inspiration for this thesis; their existence evidences that researchers in this area are looking to explore the issue of measuring habit using objective data, and both of these studies were conducted with data derived from medical devices that tracked the time and date at which medication adherence was initiated. In my early investigation and critique of these two key studies, I found there were some obvious differences in approach, which when looking at the broader literature, became even more apparent. Considering these differences, I outlined three questions that needed to be explored before any significant progress in this area could be made: 1) Why, when implementing objective measures, do some people use context stability alone as a measure of 'habit strength' (e.g. as in Phillips et al., 2021) and others multiply context stability by frequency (e.g. as in Hoo, Wildman et al., 2019); 2) What is the impact of multiplying frequency by context stability and accepting comparability between conclusions drawn from studies using these two different but ostensible measures of 'habit'?; and 3) Should we accept the raw time of day as the closest proxy to the true contextual cue that we're trying to measure? These questions formed the basis of Thesis Aim 2.

Exploration of the first and second of these questions necessitated an in-depth analysis of the numerical properties of the BFCS measure and its use in the existing literature, as well as consideration of instances in which researchers require and/or use measures of context stability, alone (Thesis Aim 2, Objective 1). The findings were presented in Chapter 4; I concluded that the BFCS approach to measuring habit (or more correctly, the likelihood of individuals having a habit; Gardner, 2015) raised too many questions for this approach to be used in the development of an objective measure in this thesis. Furthermore, the critique highlighted that to address these and other key and unanswered questions at the core of contemporary habit research (see Gardner et al., 2021's 'Twenty-one unanswered questions'), an objective measure of context stability which was conceptually and numerically distinct from frequency of behaviour was required. This critique established a detailed theoretical rationale for the remainder of the thesis and was the first time these limitations of the BFCS approach had been explicitly articulated. The critique raises novel and important concerns about our understanding of the conclusions drawn in studies which have previously used the BFCS approach to measure habit; whether variance in behaviour explained by the BFCS is driven by change in frequency or change in context stability, or both is one among a series of key questions posed in Chapter 4. Importantly, if closer inspection of the variance explained by a BFCS measure is driven by differences in frequency and not context then the conclusions that habit explained behaviour in these instances becomes challenging. Chapter 4 can be used as a guide for future research to begin to unpick these uncertainties.

The exploration of the third of my initial questions began in Chapter 6; this pertained to whether, as habit researchers, we should accept the raw time of day as the closest proxy to the true contextual cue we are trying to measure. As discussed in Chapter 1 and more explicitly in Chapter 6, little is known about what exactly constitutes a contextual cue; the example given in Chapter 6 is that someone may have a habit for reading the paper with a morning coffee, but any one of drinking coffee, the fact it is morning, sitting at the kitchen table, or the feeling of bleariness, or all four features in combination, could conceivably cue the news-reading behaviour (Wood, 2017). Some may therefore criticise the approach of using the objective time of day as a single indicator of context stability, claiming it is reductionist and oversimplified. A number of studies have used multifaceted metrics of context stability in which they ask participants to self-report on potential cues such as the time of day, location, social environment and even preceding routines to a behaviour and combine scores on each of these items to represent context stability (e.g. Friedrichsmeier et al., 2013; Judah et al., 2018; Kilb & Labudek, 2022; McCloskey & Johnson, 2019). However, few studies have explicitly compared the predictive utility of different types of context. These studies provide some tentative evidence to suggest that temporal stability may be among the

most predictive of behaviour and habit strength in physical activity; for example, Furman et al., (2021) asked their participants to report their moderate to vigorous physical activity (MVPA) engagement before and during COVID-19, as well as the context stability of their engagement in terms of consistency of the time of day, interaction partners, type of activity, surrounding events and location of exercise. They found that participants who engaged in MVPA at consistent times of day were least likely to decrease MVPA during the COVID-19 pandemic. The authors concluded that this demonstrated the time of day was both an effective cue to behaviour and the most resilient to disruption. In another study, Kaushal & Rhodes, (2015) found self-reported temporal consistency of behaviour to be predictive of habit formation. However, the findings of these two studies may need to be considered with caution as their metrics of 'temporal consistency' may not truly reflect consistency in the time of day per se (i.e., 'clock time') but may represent where the behaviour occurs between other behaviours that routinely happen at similar times of day, or in a broader sense of the time of day such as the morning, afternoon or evening. In their study, Furman et al., (2021) asked participants to respond to the statement 'Each time I exercised … it was the same time of day' on a scale from 1 = 'not at all true', to 5 = 'very true'. The phrasing of this question may lead not people to respond in terms of clock time but in more general terms (i.e., 'in the morning after breakfast and that happens at about the same time every day'). Similarly, Kaushal and Rhodes (2015) asked participants to self-report 'temporal stability' but defined this in terms of performing the behaviour at a 'particular time or after a particular activity'. One other study compared temporal stability measured objectively using timestamps to other forms of self-reported context stability; Maher et al., (2021) found a positive association between temporal context stability and MVPA, such that people who exercised at the same time of day were more active compared to people performing MVPA at more variable times. However, temporal context was summarised quite broadly into morning, afternoon, and evening in this study. These differences in terms of conceptualisation of 'temporal context' are important to consider as the literature moves forward; we must consider how the evidence provided by self-report measures such as in Furman et al., (2021) and Kaushal & Rhodes, (2015) translates to and compares with that of Maher et al., (2021) and to that of the present thesis.

Perhaps more convincing support for the use of 'clock time' to measure context stability with objective data can be gleaned from a recent study by Keller et al., (2021). This study provided interventional evidence that performing an everyday nutrition behaviour in response to (clock) time-based and routine-based cues led to increases in automaticity and enactment of behaviour in both groups; no differences between these groups were observed throughout a follow-up period of 84 days. This study is the first of its kind to demonstrate that people can- and do- effectively perform

behaviours in response to clock-times and therefore the time of day may serve as a useful starting point for the objective measurement of context stability in health behaviour.

However, the stance of the present thesis was that it is unlikely, for habits made without specific intervention and instruction, that the time of day is either ubiquitously or consciously selected as a cue to behaviour for most people, nor does it remain the extent of the context which cues a habit once established. . For the intervention arm participants at least, CFHealthHub interventionists were trained to guide participants to identify routines rather than times of day and to build these into action plans to take their medications. The strength of the present thesis is that the methods used to capture context stability from the time of day also tried to capture routines; this is a novel and conceptually advanced approach compared to existing research that has attempted to objectively measure context stability. The thesis demonstrated how RQA methods can be used as a pragmatic solution to the problem of capturing routines; used in combination with time-stamped behavioural data, RQA methods can capture whether someone repeats a behaviour within a window of time from day-to-day and from week-to-week without having to elicit self-reported responses, or without the use of more expensive and invasive tracking devices (e.g., which capture GPS information or those attached to home devices). This is the first application of RQA methods to the problem of objectively measuring context stability and the research presented in the exploration of Thesis Aim 2, Objectives 2 and 3 demonstrated that RQA can be a useful and conceptually appealing tool when applied to this issue.

The utility of the metric of context stability established in this thesis was demonstrated in a series of theoretically driven statistical tests, designed to understand the explanatory power of context stability in behaviour. The metric is unique, not only because of the features described above, but also because it is the first of its kind to be demonstrably conceptually and objectively delineated from frequency of behaviour. The issue that variance explained in behaviour because the behaviour happened before has been explicitly criticised and perhaps also overlooked; self-reported measures of context stability, to my knowledge, have undergone no empirical scrutiny in terms of the extent to which participants can accurately report on the context stability of their behaviour without this being affected by the frequency with which they perform it. The analyses presented in Chapters 7 and 8 replicate common findings from the literature but with two methodological strengths: the confidence that context stability is delineated from frequency of behaviour, and that it relates to the likelihood of a specific habit.

The latter of these strengths has also been overlooked in much of the medication adherence and habit research; particularly when self-reported measures of habit such as the SRBAI have been used

(e.g. see Badawy et al., 2020; Hoo, Gardner, et al., 2019; Wildman et al., 2021) but even in some of the work developing objective measurements of habit (Hoo, Wildman, et al., 2019; Phillips et al., 2021), researchers have tended to ask participants to self-report on the behaviour of interest in general, and not for specific performances of the behaviour. Particularly relevant to the case of medication adherence, it is theoretically plausible for people taking medications to have habits of different strengths and with different cues for their morning and evening medications; a person may have a strong habit for their morning medication but a weak or no habit for their evening, or vice versa. In such an instance, the participant will likely find it difficult to report on the context stability or strength of habit of both sessions in a single response. The analyses conducted throughout this thesis have tried to overcome this by separately analysing each repetition of behaviour within a single day.

Finally, the utility of this metric of context stability has been demonstrated with different ways of thinking about maintenance of behaviour; the metric was put to test in classic habit research paradigms in which it was used to explain and predict maintained behaviour. However, there is a need for future research to justify and delve deeper into what forms of maintenance are important to different populations, behaviours and why. This is discussed with respect to my broader recommendations for future research, below.

Combined with the findings of the machine learning-assisted review in Chapter 3, this thesis has increased the profile of habit in medication adherence research. The thesis includes two detailed discussions which lay comprehensive and rigorous theoretical foundations for the progression of research on the objective measurement of context stability and its relationship to habit, habitual behaviour and maintenance of behaviour. Exploratory empirical studies have initiated the development of an objective metric of context stability which built on these foundations. The limitations of the research presented in this thesis are discussed below, each with recommendations on how future research could address these.

10.3 Limitations and recommendations for future research

The primary limitations of this thesis relate to several aspects of generalisability; whilst the properties of BFCS measures and theoretical insights about what is needed from a measure of context stability presented in the early chapters of the thesis are applicable to a range of behaviours and populations in the study of habit, it is unclear whether the specific metric of context stability developed in this thesis will demonstrate comparable utility in explaining other health behaviours, medication adherence behaviour beyond this sample, or the first repetition of medication adherence behaviour in the day. The metric was developed using a large dataset from the ACtiF RCT

of the CFHealthHub intervention. This was a large longitudinal dataset with time-stamped adherence data. However, this is a single dataset in a single condition. Recommendations for future research are presented below in terms of further exploration of these limitations and insights from the theoretical and empirical findings of the thesis.

The dissociation of multiple repetitions of the same behaviour from one another is a conceptual strength of the research in this thesis; the approach recognised that whilst nominally the same behaviour, each repetition conducted throughout the day may be controlled by distinct habits. Steps were taken to identify what participants may have themselves considered the first inhalation of the day; as in similar, previous research in people with CF (Hoo, Wildman, et al., 2019) the start of each day was shifted from midnight to 5am in order to reflect what participants more likely considered the start of the day and to avoid capturing late (i.e. post-midnight) inhalations as early inhalations on the following day. However, analyses 3 and 4 in Chapter 9 indicated that there may be challenges in extending the foundational work conducted on the first repetition of behaviour throughout this thesis to the second, third or other inhalation of the day for two reasons.

First, higher context stability of the second inhalation of the day was not significantly associated with higher adherence to that inhalation. One possible explanation for this result might be linked to the limitations of the dataset. As a hypothetical example, a participant prescribed three doses of medication a day may have intended to take two doses in the morning and one in the evening, but on some days didn't have enough time to do two doses in the morning. They may have then skipped their 'second' and took their 'third'. On other days, they may have managed to fit both doses into their morning schedule. Because the drug that was inhaled was not labelled in this dataset it was not possible to identify instances in which this may have happened. The impact of this is that the participant's set of 'second inhalations' would consist of some in the morning and some in the evening which might not reflect what they considered to be their adherence to the second medication. This would have artificially reduced their context stability score. The only way to investigate this further is using a dataset with some mechanism of flagging instances when this occurred (e.g., by asking participants what order and when they intend to take their medications and having devices which track which medication is taken).

The second reason the metric of context stability may not extend well to repetitions of behaviour later in the day was also indicated in Chapter 9; the three plots of inhalations and respective context stability scores showed lower context stability to evening repetitions of behaviour than the morning, whilst still showing very high adherence. One explanation for this is that these plots represent people with habits for their evening behaviour but whose cues occur at less rigid clock-times. It is

possible that many people who, for example, have work or daily commitments, have fairly rigid morning routines that are driven by clock-time limits (i.e. waking up because an alarm clock sounds at 6:30am, having to drop their children off at school by 8:30am and being at work or 9am). Because of the time pressure, activities conducted between these clock-times may follow very similar routines which closely align with clock-times. However, the same person's evening routines may follow less clock-driven agendas; the same or similar pattern of behaviours may be initiated on walking through the door at 6pm but the length of time spent on them may not be as rigid (e.g. cooking dinner may take five, 20 or 40 minutes, depending on the complexity of what is being prepared). The metric of context stability developed in this thesis was designed to capture repetition of behaviour within windows of time; the method is designed to allow for routines to happen at a different time from the day before as long as it is repeated again in the future, but this is bounded by the behaviour recurring within a specific window of time. The W_rad_90 variable indicates whether behaviour did not occur within 90 minutes of another observation and the LAM_rad_15 variable indicates whether behaviour occurred within 15 minutes of another observation. It may be that, for behaviours cued by other routines in the evening, a 15-minute window is too narrow and would need to be adjusted to better reflect the nature of the cue. These limitations are relevant both for generalising the utility of the metric of context stability developed in this thesis to other repetitions of behaviour in the day and to other behaviours. Future research should investigate whether the metric of context stability developed in this thesis can be explain the same or similar variance in behaviours performed later in the day and whether this is true of behaviours other than adherence to medication. It may be that other metrics derived from RQA are more suited or sensitive to other behaviours and those performed later in the day.

10.4 Broader reflections on the thesis findings and further recommendations

Throughout the conduct of this thesis, I have reflected on how an objective measure of context stability and medication adherence in people living with long term conditions, particularly people with CF, could be useful in the study of questions at the centre of current debate in habit theory. One such debate is around the impact of behavioural complexity. Whilst not possible with the present dataset, changing when people prescribed three doses of medication take their medications throughout the day could be a useful case study for exploring the challenges of behavioural complexity. In Chapter 9 for example, I attempted to measure context stability to the first and second repetition of behaviour to explore whether linking two repetitions of behaviour together (i.e., 'behaviour stacking') was more effective for adherence to medication in people with CF than taking only one dose in the morning and two in the evening. This was framed in terms of an applied clinical research question and the effects of time of day on behaviour and the conclusions were that

experimental manipulation of morning vs evening scheduling of doses was required to draw definitive conclusions about whether this made a difference. This same study could also be useful in developing the evidence base for recently established theory about behavioural complexity and habit (Gardner, 2022; Gardner & Lally, 2022; Phillips & Mullan, 2022; Rebar et al., 2022). Recent definitions of complexity converge on the number of sub-actions required to complete a behaviour being a key characteristic which defines degree of complexity. There are opposing views on the effects of 'behavioural complexity' on the formation and enactment of habits but there is already some convincing theoretical grounds and supporting empirical evidence that the complexity of behaviour has no impact on the relationship between habit and behaviour frequency (see Gardner, 2022; Gardner & Lally, 2022). However, the extent that complexity is something that subjectively changes despite objective similarity (as argued to be a characteristic of complexity by Rebar et al., 2022) and the effects of behaviour complexity on the translation of habit into action is a matter of contention and could be studied in behaviours in which the objective complexity in terms of number of sub-actions can be easily defined and manipulated (i.e. one vs two doses in a single session). An interventional study with ecological momentary assessment of perceived complexity, in which some participants are assigned to take their first two doses in the morning and one in the evening and the others are assigned to do the opposite, could shed light on this.

Another reflection relates to how habit and health research defines maintenance of behaviour and how much emphasis habit researchers should put on trying to achieve truly habitual health behaviour. First, with regard to maintenance: 'Maintenance' or 'future behaviour' are seldom defined in the same terms from one study to the next. One of the most commonly used definitions is taken from the Transtheoretical Model of Health Behaviour change (Prochaska & DiClemente, 1982), which among other processes and 'Stages of Change' defines maintenance as evidence of sustained behaviour six-months post-initiation. This has become popular among researchers interested in maintenance as it provides an easily identifiable time cut-off. This definition was used in the rationale for exclusion of studies with less than six-month follow-ups in the review presented in Chapter 3 and used to guide the selection of data in the exploratory research conducted in Chapters 6 to 8. However, in Chapter 9 'failure to maintain adherence' was defined as failing to take medication for three consecutive days following initiation. Whilst a very different metric of maintenance, this definition falls within the scope of one of the most seminal syntheses of maintenance theory by Kwasnicka et al., (2016), who defined maintenance as 'the continuous performance of a behaviour following an initial intentional change at a level that significantly differs from the baseline performance in the intended direction.' (p.280). Their definition appears to purposely leave room for different metrics of maintenance to capture complexities between

different behaviours, contexts and individuals; it is vague in terms of length and/or amount of change (i.e., this definition could capture clinically or statistically significant change and in terms of any number of outcomes such as the number of days observed, intensity of the behaviour, and so on). The stance of the present thesis aligns with that of Kwasnicka et al., (2016) in that maintenance is perhaps best conceptualised a case of 'horses for courses'; whilst the six-month cut-off used in the earlier thesis Chapters provided useful and easily implemented bounds to work with, the definition developed in Chapter 9 was implemented to capture possible habit formation in terms of maintenance in the absence of a metric of habitual automaticity. Whilst I believe that different definitions of maintenance should continue to be accepted in the literature, it makes synthesis of research using different definitions difficult. This could be helped if future research probes the implications of each definition on other outcomes and lengthier follow-ups. For example, syntheses could more informatively combine outcomes from multiple definitions of maintenance if they converge in predicting positive clinical outcomes at one or two-year follow-up. Specifically following on from this thesis, questions future research could pursue include: i) how meaningful is continuing to engage with a behaviour after six months in terms of clinical outcomes (e.g. lung function), on adherence behaviour at one year or later, and after an intervention is withdrawn?; and ii) Does the definition of maintenance in Chapter 9 predict increased likelihood of re-engagement after an instance of failure to maintain?

Second, regarding my reflections on measuring habit, habitual behaviour and the goal of 'habit formation interventions'. A metric of context stability, such as that developed in the present thesis, is a measure of behaviour, not habit, and cannot distinguish between people who are forming habits but self-regulating behaviour in response to a stable contextual cue, and those who are performing a habitual behaviour. There is no distinguishable difference in the outward and objective appearance of behaviour between these states; the transition is driven by cumulative repetition of behaviour in response to the cue (see conceptual model by Gardner & Lally, 2018) and the number of repetitions it takes to achieve automaticity varies, likely by behaviour and individual differences (Kilb & Labudek, 2022; Lally et al., 2010). However, the goal of habit formation interventions is to achieve maintenance of behaviour and there are several arguments as to why use of an objective metric of context stability in combination with a metric of behavioural maintenance is useful to habit formation intervention designers and providers: 1) an objective metric of context stability enables tracking and assessment of the extent to which the key BCT (i.e. context-dependent repetition) is implemented, and the subsequent impact of that on maintenance of behaviour; 2) objective context stability could be used as an indicator of people who have failed to make a habit by the time an intervention is withdrawn; 3) questions remain as to whether we should expect health behaviour to

become truly habitual and studies that have tracked habit formation in health behaviour show that on average participants' scores plateau below the maximum of self-report scales of habitual behaviour (Fournier et al., 2017; Keller et al., 2021; Lally et al., 2010). One recent theory has suggested that the ability to flexibly and synergistically switch between automatic (e.g. habit) and reflective systems (e.g. self-regulation) may actually be the key to successful behavioural maintenance (see Dunton et al., 2021). To explore the theory of Dunton and colleagues in more detail, they list several requirements which includes a metric of context stability and one which can be used to measure change over time. Whilst the work presented in this thesis used average context stability over periods of observation, future research could examine how the metric developed in this thesis could be calculated and applied in rolling windows of 14 observations and combined with methods being developed elsewhere in the field to detect moments of change (e.g. see machine learning detection algorithms developed by Phillips et al., 2022).

My final reflection is on the role of reward in habit formation; I noted that only one intervention included in the review in Chapter 3 included a reward-related BCT and this topic has recently come to the fore of habit theory discourse. Early habit theory stems from the behaviourism school of thought, in which researchers used animal paradigms to understand the mechanisms of conditioning and habit formation, concluding that experiencing reward following a specific response to a stimulus was the key to explaining why behaviour happened (Hull, 1943). The role of reward has largely remained focal to the neurological and computational study of habits (see Wood & Rünger, 2016 for a review), whose research tends to function in the realms of experimental manipulation of stimuli and reward in order to study the brain and simple behaviours such as pressing a button, under controlled experimental conditions. However, the extent to which contemporary theory from health psychology, which aims to explain much less controlled and more complex behaviours, have acknowledged the role of reward in habit has varied. Some researchers have explicitly acknowledged the necessity of reward (Phillips & Mullan, 2022; Wood & Rünger, 2016), but others have either described this as a factor which can influence the speed with which habits are acquired (Rebar et al., 2022) or have omitted it all together (see definitions in Gardner 2015). Recently, Gardner and Lally (2022) have gone as far as to suggest that there is no empirical evidence to suggest that reward is necessary for habits to form in health behaviours. This remains an under-researched issue and the conclusions of the review were limited by the needs to define what BCTs are necessary and sufficient to constitute a 'habit formation intervention'.

10.5 Conclusion

A significant effort to incorporate existing knowledge and to deconstruct and critique widely accepted practices in the habit literature has been central to the conduct of the work presented in this thesis. The key contributions of this work lie in the exposure of the needs of, and opportunities afforded by, habit research conducted on medication adherence behaviour, the lessons learned during the conduct of BCT coding in 'habit interventions' and in the progression of the discourse surrounding the objective measurement of habit and context stability. The key strength is the articulation, consideration and empirical examination of potential solutions to theoretical and methodological issues in the extant habit literature. The thesis presented evidence in support of the utility of an objective metric of context stability and is a starting point for future research to continue its development by examining the scope of its generalisability in other samples, population groups and behaviours.

Word count: 60,663

References

Aarts, H. (2007). Health and goal-directed behavior: The nonconscious regulation and motivation of goals and their pursuit. *Health Psychology Review*, *1*(1), 53–82. https://doi.org/10.1080/17437190701485852

Adams, C. D., & Dickinson, A. (1981). Instrumental Responding following Reinforcer Devaluation. *The Quarterly Journal of Experimental Psychology Section B*, *33*(2b), 109–121. https://doi.org/10.1080/14640748108400816

Adriaanse, M. A., Gollwitzer, P. M., De Ridder, D. T. D., De Wit, J. B. F., & Kroese, F. M. (2011). Breaking Habits With Implementation Intentions: A Test of Underlying Processes. *Article Personality and Social Psychology Bulletin*, *37*(4), 502–513. https://doi.org/10.1177/0146167211399102

Adriaanse, M. A., Kroese, F. M., Weijers, J., Gollwitzer, P. M., & Oettingen, G. (2018). Explaining unexplainable food choices. *European Journal of Social Psychology*, *48*(1), O15–O24. https://doi.org/10.1002/EJSP.2273

Adriaanse, M. A., Prinsen, S., de Witt Huberts, J. C., de Ridder, D. T. D., & Evers, C. (2016). "I ate too much so I must have been sad": Emotions as a confabulated reason for overeating. *Appetite*, *103*, 318–323. https://doi.org/10.1016/j.appet.2016.04.028

Ajzen, I. (1991). The Theory of Planned Behavior. *ORGANIZATIONAL BEHAVIOR AND HUMAN DECISION PROCESSES*, *50*, 179–211.

Ajzen, I. (2002). Residual Effects of Past on Later Behavior: Habituation and Reasoned Action Perspectives: *Personality and Social Psychology Review*, *6*(2), 107–122. https://doi.org/10.1207/S15327957PSPR0602_02

Ajzen, I. (2011). The theory of planned behaviour: Reactions and reflections. *Psychology & Health*, *26*(9), 1113–1127. https://doi.org/10.1080/08870446.2011.613995

Aloia, M. S., Goodwin, M. S., Velicer, W. F., Arnedt, J. T., Zimmerman, M., Skrekas, J., Harris, S., & Millman, R. P. (2008). Time series analysis of treatment adherence patterns in individuals with obstructive sleep apnea. *Annals of Behavioral Medicine*, *36*(1), 44–53. https://doi.org/10.1007/s12160-008-9052-9

Amico, K. R., Mugavero, M., Krousel-Wood, M. A., Bosworth, H. B., & Merlin, J. S. (2018). Advantages to Using Social-Behavioral Models of Medication Adherence in Research and Practice. *Journal of General Internal Medicine*, *33*(2), 207–215. https://doi.org/10.1007/s11606-017-4197-5

Amin, R., Dupuis, A., Aaron, S. D., & Ratjen, F. (2010). The effect of chronic infection with Aspergillus fumigatus on lung function and hospitalization in patients with cystic fibrosis. *Chest*, *137*(1), 171–176. https://doi.org/10.1378/chest.09-1103

Anderson, L. J., Nuckols, T. K., Coles, C., Le, M. M., Schnipper, J. L., Shane, R., Jackevicius, C., Lee, J., Pevnick, J. M., & Group, M. of the P.-D. (2020). A systematic overview of systematic reviews evaluating medication adherence interventions. *American Journal of Health-System Pharmacy*, *77*(2), 138–147. https://doi.org/10.1093/ajhp/zxz284

Arden, M A, Hutchings, M., Whelan, P., Drabble, S. J., Beever, D., Bradley, J. M., Hind, D., Ainsworth, J., Maguire, C., Cantrill, H., O'Cathain, A., & Wildman, M. (2021). Development of an intervention to increase adherence to nebuliser treatment in adults with cystic fibrosis: CFHealthHub. *Pilot and Feasibility Studies*, *7*(1), 1. https://doi.org/10.1186/s40814-020-00739-2

Arden, Madelynne A., Drabble, S., O'Cathain, A., Hutchings, M., & Wildman, M. (2019). Adherence to medication in adults with Cystic Fibrosis: An investigation using objective adherence data and the Theoretical Domains Framework. *British Journal of Health Psychology*, *24*(2), 357–380. https://doi.org/10.1111/BJHP.12357

Armitage, C. J., & Conner, M. (2000). Psychology and Health Social cognition models and health behaviour: A structured review. *Psychology and Health*, *15*, 173–189. https://doi.org/10.1080/08870440008400299

Armitage, C. J., & Conner, M. (2001). Efficacy of the Theory of Planned Behaviour: A meta-analytic review. *British Journal of Social Psychology*, *40*(4), 471–499. https://doi.org/10.1348/014466601164939

Asche, C., LaFleur, J., & Conner, C. (2011). A Review of Diabetes Treatment Adherence and the Association with Clinical and Economic Outcomes. *Clinical Therapeutics*, *33*(1), 74–109. https://doi.org/10.1016/j.clinthera.2011.01.019

Badawy, S. M., Shah, R., Beg, U., & Heneghan, M. B. (2020). Habit Strength, Medication Adherence, and Habit-Based Mobile Health Interventions Across Chronic Medical Conditions: Systematic Review. In *Journal of medical Internet research* (Vol. 22, Issue 4, p. e17883). NLM (Medline). https://doi.org/10.2196/17883

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes*, *50*(2), 248–287. https://doi.org/10.1016/0749-5978(91)90022-L

Bandura, A. (1998). Health promotion from the perspective of social cognitive theory. *Psychology and Health*, *13*(4), 623–649. https://doi.org/10.1080/08870449808407422

Bangsberg, D., Perry, S., Charlebois, E., Clark, R., Robertson, M., Zolopa, A., & Moss, A. (2001). Non-adherence to highly active antiretroviral therapy predicts progression to AIDS. *AIDS*, *15*(9), 1181–1183.

Barankay, I., Reese, P. P., Putt, M. E., Russell, L. B., Loewenstein, G., Pagnotti, D., Yan, J., Zhu, J., McGilloway, R., Brennan, T., Finnerty, D., Hoffer, K., Chadha, S., & Volpp, K. G. (2020). Effect of Patient Financial Incentives on Statin Adherence and Lipid Control: A Randomized Clinical Trial. *JAMA Network Open*, *3*(10), e2019429. https://doi.org/10.1001/jamanetworkopen.2020.19429

Bauer, M., Glenn, T., Alda, M., Bauer, R., Grof, P., Marsh, W., Monteith, S., Munoz, R., Rasgon, N., 9^, K. S., & Whybrow, P. C. (2019). Trajectories of adherence to mood stabilizers in patients with bipolar disorder. *Int J Bipolar Disord*, *7*, 19. https://doi.org/10.1186/s40345-019-0154-z

Bell, S. C., Mall, M. A., Gutierrez, H., Macek, M., Madge, S., Davies, J. C., Burgel, P. R., Tullis, E., Castaños, C., Castellani, C., Byrnes, C. A., Cathcart, F., Chotirmall, S. H., Cosgriff, R., Eichler, I., Fajac, I., Goss, C. H., Drevinek, P., Farrell, P. M., … Ratjen, F. (2020). The future of cystic fibrosis care: a global perspective. In *The Lancet Respiratory Medicine* (Vol. 8, Issue 1, pp. 65–124). Lancet Publishing Group. https://doi.org/10.1016/S2213-2600(19)30337-6

Bem, D. (1972). Self-perception theory. *Advances in Experimental Social Psychology*, *6*.

Bernard, R. S., & Cohen, L. L. (2004). Increasing Adherence to Cystic Fibrosis Treatment: A Systematic Review of Behavioral Techniques. In *Pediatric Pulmonology* (Vol. 37, Issue 1, pp. 8–16). https://doi.org/10.1002/ppul.10397

Bishay, L., & Sawicki, G. (2016). Strategies to optimize treatment adherence in adolescent patients with cystic fibrosis. *Adolescent Health, Medicine and Therapeutics*, *Volume 7*, 117–124. https://doi.org/10.2147/ahmt.s95637

Blaschke, T. F., Osterberg, L., Vrijens, B., & Urquhart, J. (2012). Adherence to Medications: Insights Arising from Studies on the Unreliable Link Between Prescribed and Actual Drug Dosing Histories. *Annual Review of Pharmacology and Toxicology*, *52*(1), 275–301. https://doi.org/10.1146/annurev-pharmtox-011711-113247

Bolón-Canedo, V., Sánchez-Maroño, N., & Alonso-Betanzos, A. (2012). A review of feature selection methods on synthetic data. *Knowledge and Information Systems* , *34*(3), 483–519. https://doi.org/10.1007/S10115-012-0487-8

Botvinick, M., & Plaut, D. C. (2004). Doing Without Schema Hierarchies: A Recurrent Connectionist Approach to Normal and Impaired Routine Sequential Action. *Psychological Review*, *111*(2), 395–429. https://doi.org/10.1037/0033-295X.111.2.395

Briesacher, B. A., Andrade, S. E., Fouayzi, H., & Chan, K. A. (2008). Comparison of drug adherence rates among patients with seven different medical conditions. *Pharmacotherapy*, *28*(4), 437–443. https://doi.org/10.1592/phco.28.4.437

Britto, M. T., Kotagal, U. R., Atherton, H. D., Wilmott, R. W., Hornung, R. W., & Tsevat, J. (2002). Impact of recent pulmonary exacerbations on quality of life in patients with cystic fibrosis. *Chest*, *121*(1), 64–72. https://doi.org/10.1378/chest.121.1.64

Brown, D., Morrissey, S., Hagger, M., & Hamilton, K. (2019). A Qualitative Investigation into the Representations of Habit by Lay People. *Preprint*. https://doi.org/10.31234/osf.io/wf4ah

Brown, L. D., Cai, T. T., & DasGupta, A. (2001). Interval Estimation for a Binomial Proportion. *Statistical Science*, *16*(2). https://doi.org/10.1214/ss/1009213286

Brown, M. T., Bussell, J., Dutta, S., Davis, K., Strong, S., & Mathew, S. (2016). Medication Adherence: Truth and Consequences. *The American Journal of the Medical Sciences*, *351*(4), 387–399. https://doi.org/10.1016/j.amjms.2016.01.010

Brown, M. T., & Bussell, J. K. (2011). Medication adherence: WHO cares? In *Mayo Clinic Proceedings* (Vol. 86, Issue 4, pp. 304–314). Elsevier Ltd. https://doi.org/10.4065/mcp.2010.0575

Bruin, M. de, Black, N., Javornik, N., Viechtbauer, W., Eisma, M. C., Hartman-Boyce, J., Williams, A. J., West, R., Michie, S., & Johnston, M. (2020). Underreporting of the active content of behavioural interventions: a systematic review and meta-analysis of randomised trials of smoking cessation interventions. *Health Psychology Review*, *15*(2), 195–213. https://doi.org/10.1080/17437199.2019.1709098

Burgel, P. R., Bellis, G., Olesen, H. V., Viviani, L., Zolin, A., Blasi, F., Elborn, J. S., Bell, S., Castellani, C., Dembski, B., Drevinek, P., Heijerman, H., Innes, A., Lindblad, A., Madge, S., De Rijcke, K., & Solé, A. (2015). Future trends in cystic fibrosis demography in 34 European countries. *European Respiratory Journal*, *46*(1), 133–141. https://doi.org/10.1183/09031936.00196314

Burrows, J. A., Bunting, J. P., Masel, P. J., & Bell, S. C. (2002). Nebulised dornase alpha: adherence in adults with cystic fibrosis. *Journal of Cystic Fibrosis : Official Journal of the European Cystic Fibrosis Society*, *1*(4), 255–259. https://doi.org/10.1016/s1569-1993(02)00095-4

Candy, B., Vickerstaff, V., Jones, L., & King, M. (2018). Description of complex interventions: analysis of changes in reporting in randomised trials since 2002. *Trials*, *19*(1). https://doi.org/10.1186/S13063-018-2503-0

Carey, R. N., Connell, L. E., Johnston, M., Rothman, A. J., De Bruin, M., Kelly, M. P., & Michie, S. (2018). Behavior Change Techniques and Their Mechanisms of Action: A Synthesis of Links Described in Published Intervention Literature. *Annals of Behavioral Medicine*, *53*(8), 693–707. https://doi.org/10.1093/abm/kay078

Carver, C. ., & Scheier, M. (1982). Control theory: A useful conceptual framework for personality-social , clinical and health psychology. *Psychological Bulletin*, *92*(1), 111–135.

CFHealthHub. (2023). *Trial Documentation - CFHealthHub Data Observatory*. https://www.cfhealthhub.com/trial-documentation/

Chalmers, J. D. (2020). Cystic fibrosis lung disease and bronchiectasis. In *The Lancet Respiratory Medicine* (Vol. 8, Issue 1, pp. 12–14). Lancet Publishing Group. https://doi.org/10.1016/S2213-2600(19)30335-2

Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. *Computers & Electrical Engineering*, *40*(1), 16–28. https://doi.org/10.1016/J.COMPELECENG.2013.11.024

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, *16*, 321–357. https://doi.org/10.1613/jair.953

Chowdhury, R., Khan, H., Heydon, E., Shroufi, A., Fahimi, S., Moore, C., Stricker, B., Mendis, S., Hofman, A., Mant, J., & Franco, O. H. (2013). Adherence to cardiovascular therapy: a meta-analysis of prevalence and clinical consequences. *European Heart Journal*, *34*(38), 2940–2948. https://doi.org/10.1093/eurheartj/eht295

Conn, V. S., Enriquez, M., Ruppar, T. M., & Chan, K. C. (2016). Meta-analyses of theory use in medication adherence intervention research. *American Journal of Health Behavior*, *40*(2), 155–171. https://doi.org/10.5993/AJHB.40.2.1

Conn, V. S., & Ruppar, T. M. (2017a). Medication adherence outcomes of 771 intervention trials: Systematic review and meta-analysis. *Preventive Medicine*, *99*, 269–276. https://doi.org/10.1016/j.ypmed.2017.03.008

Conn, V. S., Ruppar, T. M., & Chase, J. A. D. (2016). Blood pressure outcomes of medication adherence interventions: systematic review and meta-analysis. *Journal of Behavioral Medicine*, *39*(6), 1065–1075. https://doi.org/10.1007/s10865-016-9730-1

Conn, V. S., Ruppar, T. M., Enriquez, M., & Cooper, P. (2016). Medication adherence interventions that target subjects with adherence problems: Systematic review and meta-analysis. *Research in Social and Administrative Pharmacy*, *12*(2), 218–246. https://doi.org/10.1016/j.sapharm.2015.06.001

Connell, L. E., Carey, R. N., De Bruin, M., Rothman, A. J., Johnston, M., Kelly, M. P., & Michie, S. (2018). Links between Behavior Change Techniques and Mechanisms of Action: An Expert Consensus Study. *Annals of Behavioral Medicine*, *53*(8), 708–720. https://doi.org/10.1093/abm/kay082

Conner, M, & Norman, P. (2005). *Predicting Health Behaviour* (M Conner & P. Norman (eds.); 2nd ed.). Open University Press.

Conner, Mark, & Armitage, C. J. (1998). Extending the Theory of Planned Behavior: A Review and Avenues for Further Research. *Journal of Applied Social Psychology*, *28*(15), 1429–1464. https://doi.org/10.1111/J.1559-1816.1998.TB01685.X

Curioni, C. C., & Lourenço, P. M. (2005). Long-term weight loss after diet and exercise: A systematic review. In *International Journal of Obesity* (Vol. 29, Issue 10, pp. 1168–1174). Int J Obes (Lond). https://doi.org/10.1038/sj.ijo.0803015

Currie, G. L., Angel-Scott, H. N., Colvin, L., Cramond, F., Hair, K., Khandoker, L., Liao, J., Macleod, M., McCann, S. K., Morland, R., Sherratt, N., Stewart, R., Tanriver-Ayder, E., Thomas, J., Wang, Q.,

Wodarski, R., Xiong, R., Rice, A. S. C., & Sena, E. S. (2019). Animal models of chemotherapy-induced peripheral neuropathy: A machine-assisted systematic review and meta-analysis. *PLOS Biology*, *17*(5). https://doi.org/10.1371/JOURNAL.PBIO.3000243

Cutler, R. L., Fernandez-Llimos, F., Frommer, M., Benrimoj, C., & Garcia-Cardenas, V. (2018). Economic impact of medication non-adherence by disease groups: a systematic review. *BMJ Open*, *8*(1), e016982–e016982. https://doi.org/10.1136/bmjopen-2017-016982

Cyctic Fibrosis Trust. (2022). *UK Cystic Fibrosis Registry 2021 Annual Data Report*.

Cystic Fibrosis Trust. (2023). *Triple combination therapy Kaftrio (Trikafta in the US)*. https://www.cysticfibrosis.org.uk/what-is-cystic-fibrosis/cystic-fibrosis-care/treatments-and-medication/kaftrio

Daniels, T., Goodacre, L., Sutton, C., Pollard, K., Conway, S., & Peckham, D. (2011). Accurate assessment of adherence: self-report and clinician report vs electronic monitoring of nebulizers. *Chest*, *140*(2), 425–432. https://doi.org/https://dx.doi.org/10.1378/chest.09-3074

Daniels, T., Mills, N., & Whitaker, P. (2013). Nebuliser systems for drug delivery in cystic fibrosis. In *Cochrane Database of Systematic Reviews* (Vol. 2013, Issue 4). John Wiley and Sons Ltd. https://doi.org/10.1002/14651858.CD007639.pub2

Danner, U. N., Aarts, H., & De Vries, N. K. (2008). Habit vs. intention in the prediction of future behaviour: The role of frequency, context stability and mental accessibility of past behaviour. *British Journal of Social Psychology*, *47*(2), 245–265. https://doi.org/10.1348/014466607X230876

De Bruin, M., Hospers, H. J., van Breukelen, G. J. P., Kok, G., Koevoets, W. M., & Prins, J. M. (2010). Electronic monitoring-based counseling to enhance adherence among HIV-infected patients: A randomized controlled trial. *Health Psychology*, *29*(4), 421–428. https://doi.org/10.1037/a0020335

Dean, A. J., Kneebone, S., Tull, F., Lauren, N., & Smith, L. D. G. (2021). 'Stickiness' of water-saving behaviours: What factors influence whether behaviours are maintained or given up? *Resources, Conservation and Recycling*, *169*. https://doi.org/10.1016/J.RESCONREC.2021.105531

Demonceau, J., Ruppar, T., Kristanto, P., Hughes, D. A., Fargher, E., Kardas, P., De Geest, S., Dobbels, F., Lewek, P., Urquhart, J., & Vrijens, B. (2013). Identification and assessment of adherence-enhancing interventions in studies assessing medication adherence through electronically compiled drug dosing histories: A systematic literature review and meta-analysis. In *Drugs* (Vol. 73, Issue 6, pp. 545–562). https://doi.org/10.1007/s40265-013-0041-3

Department of Health. (2012). *Long Term Conditions Compendium of Information: Third Edition*. http://www.dh.gov.uk/publications

Dickinson, A., Nicholas, D. J., & Adams, C. D. (1983). The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. *The Quarterly Journal of Experimental Psychology Section B*, *35*(1), 35–51. https://doi.org/10.1080/14640748308400912

Diefenbacher, S., Lally, P., & Gardner, B. (2022). Habit formation in context: Context-specific and context-free measures for tracking fruit consumption habit formation and behaviour. *British Journal of Health Psychology*. https://doi.org/10.1111/BJHP.12637

DiMatteo, M. R. (2004). Variations in patients' adherence to medical recommendations: A quantitative review of 50 years of research. In *Medical Care* (Vol. 42, Issue 3, pp. 200–209). https://doi.org/10.1097/01.mlr.0000114908.90348.f9

Dimatteo, M. R., Reiter, R. C., & Gambone, J. C. (1994). Enhancing Medication Adherence Through Communication and Informed Collaborative Choice. *Health Communication*, *6*(4), 253–265. https://doi.org/10.1207/s15327027hc0604_2

Domke, A., Keller, J., Fleig, L., Knoll, N., & Schwarzer, R. (2019). What makes a good action plan? Characteristics and enactment of fruit and vegetable plans. *Appetite*, *142*, 104351. https://doi.org/10.1016/J.APPET.2019.104351

Drabble, S. J., O'Cathain, A., Arden, M. A., Hutchings, M., Beever, D., & Wildman, M. (2019). When Is Forgetting Not Forgetting? A Discursive Analysis of Differences in Forgetting Talk Between Adults With Cystic Fibrosis With Different Levels of Adherence to Nebulizer Treatments. *Qualitative Health Research*, *29*(14), 2119–2131. https://doi.org/10.1177/1049732319856580

Duhigg, C. (2012). *The Power of Habit: Why We Do what We Do and how to Change* . Random House.

Dunton, G. F., Rothman, A. J., Leventhal, A. M., & Intille, S. S. (2021). How intensive longitudinal data can stimulate advances in health behavior maintenance theories and interventions. *Translational Behavioral Medicine*, *11*(1), 281–286. https://doi.org/10.1093/TBM/IBZ165

Eakin, M. N., Bilderback, A., Boyle, M. P., Mogayzel, P. J., & Riekert, K. A. (2011). Longitudinal association between medication adherence and lung health in people with cystic fibrosis. *Journal of Cystic Fibrosis*, *10*(4), 258–264. https://doi.org/10.1016/j.jcf.2011.03.005

Edmondson, D., Falzon, L., Sundquist, K. J., Julian, J., Meli, L., Sumner, J. A., & Kronish, I. M. (2018). A systematic review of the inclusion of mechanisms of action in NIH-funded intervention trials to improve medication adherence. *Behaviour Research and Therapy*, *101*, 12–19. https://doi.org/10.1016/j.brat.2017.10.001

Elborn, J. S. (2016). Cystic fibrosis. In *The Lancet* (Vol. 388, Issue 10059, pp. 2519–2531). Lancet Publishing Group. https://doi.org/10.1016/S0140-6736(16)00576-6

Esther, C. R., Muhlebach, M. S., Ehre, C., Hill, D. B., Wolfgang, M. C., Kesimer, M., Ramsey, K. A., Markovetz, M. R., Garbarine, I. C., Gregory Forest, M., Seim, I., Zorn, B., Morrison, C. B., Delion, M. F., Thelin, W. R., Villalon, D., Sabater, J. R., Turkovic, L., Ranganathan, S., … Boucher, R. C. (2019). Mucus accumulation in the lungs precedes structural changes and infection in children with cystic fibrosis. *Science Translational Medicine*, *11*(486). https://doi.org/10.1126/scitranslmed.aav3488

Farmer, A. J., Oke, J., Hardeman, W., Tucker, L., Sutton, S., Kinmonth, A.-L., Griffin, S., & Holman, R. R. (2016). The effect of a brief action planning intervention on adherence to double-blind study medication, compared to a standard trial protocol, in the Atorvastatin in Factorial with Omega EE90 Risk Reduction in Diabetes (AFORRD) clinical trial: A cluster randomi. *Diabetes Research and Clinical Practice*, *120*, 56–64. https://doi.org/10.1016/j.diabres.2016.07.004

Fisher, A. J., & Boswell, J. F. (2016). Enhancing the Personalization of Psychotherapy With Dynamic Assessment and Modeling. *Assessment*, *23*(4), 496–506. https://doi.org/10.1177/1073191116638735

Fleetwood, S. (2019). A definition of habit for socio-economics. *Review of Social Economy*, *79*(2), 131–165. https://doi.org/10.1080/00346764.2019.1630668

Flume, P. A., Mogayzel, P. J., Robinson, K. A., Goss, C. H., Rosenblatt, R. L., Kuhn, R. J., Marshall, B. C., Bujan, J., Downs, A., Finder, J., Goss, C., Gutierrez, H., Hazle, L., Lester, M., Quittell, L., Sabadosa, K., Vender, R. L., White, T. B., Willey-Courand, D. B., … Waybright, S. (2009). Cystic fibrosis pulmonary guidelines: Treatment of pulmonary exacerbations. In *American Journal of Respiratory and Critical Care Medicine* (Vol. 180, Issue 9, pp. 802–808).

https://doi.org/10.1164/rccm.200812-1845PP

Flume, P. A., O'Sullivan, B. P., Robinson, K. A., Goss, C. H., Mogayzel, P. J., Willey-Courand, D. B., Bujan, J., Finder, J., Lester, M., Quittell, L., Rosenblatt, R., Vender, R. L., Hazle, L., Sabadosa, K., & Marshall, B. (2007). Cystic fibrosis pulmonary guidelines: Chronic medications for maintenance of lung health. *American Journal of Respiratory and Critical Care Medicine*, *176*(10), 957–969. https://doi.org/10.1164/rccm.200705-664OC

Fournier, M., d'Arripe-Longueville, F., & Radel, R. (2017). Testing the effect of text messaging cues to promote physical activity habits: a worksite-based exploratory intervention. *Scandinavian Journal of Medicine & Science in Sports*, *27*(10), 1157–1165. https://doi.org/10.1111/sms.12730

Fournier, Marion, D'Arripe-Longueville, F., Rovere, C., Easthope, C. S., Schwabe, L., El Methni, J., & Radel, R. (2017). Effects of circadian cortisol on the development of a health habit. *Health Psychology*, *36*(11), 1059–1064. https://doi.org/10.1037/hea0000510

Franklin, J., Shrank, W., Pakes, J., Sanfelix-Gimeno. G, Matlin, O., Brennan, T., & Choudhry, N. (2013). Group-based Trajectory Models: A New Approach to Classifying and Predicting Long-Term Medication Adherence on JSTOR. *Medical Care*, *51*(9), 789–796.

Friedrichsmeier, T., Matthies, E., & Klöckner, C. A. (2013). Explaining stability in travel mode choice: An empirical comparison of two concepts of habit. *Transportation Research Part F: Traffic Psychology and Behaviour*, *16*, 1–13. https://doi.org/10.1016/J.TRF.2012.08.008

Furman, C. R., Volz, S. C., & Rothman, A. J. (2021a). Contextual disruption and exercise: mapping changes to exercise routines and engagement during the COVID-19 pandemic. *Psychology & Health*. https://doi.org/10.1080/08870446.2021.2008393

Galla, B. M., & Duckworth, A. L. (2015). More than resisting temptation: Beneficial habits mediate the relationship between self-control and positive life outcomes. *Journal of Personality and Social Psychology*, *109*(3), 508–525. https://doi.org/10.1037/PSPP0000026

Gantman, A. P., Adriaanse, M. A., Gollwitzer, P. M., & Oettingen, G. (2017). Why did I do that? Explaining actions activated outside of awareness. *Psychonomic Bulletin & Review 2017 24:5*, *24*(5), 1563–1572. https://doi.org/10.3758/S13423-017-1260-5

Garber, M. C., Nau, D. P., Erickson, S. R., Aikens, J. E., & Lawrence, J. B. (2004). The concordance of self-report with other measures of medication adherence: A summary of the literature. *Medical Care*, *42*(7), 649–652. https://doi.org/10.1097/01.mlr.0000129496.05898.02

Gardner, B. (2015a). A review and analysis of the use of "habit" in understanding, predicting and influencing health-related behaviour. *Health Psychology Review*, *9*(3), 277–295. https://doi.org/10.1080/17437199.2013.876238

Gardner, B, Abraham, C., Lally, P., & de Bruijn, G.-J. (2012). Towards parsimony in habit measurement: testing the convergent and predictive validity of an automaticity subscale of the Self-Report Habit Index. *The International Journal of Behavioral Nutrition and Physical Activity*, *9*, 102. https://doi.org/10.1186/1479-5868-9-102

Gardner, B, & Lally, P. (2018a). Modelling habit formation and its determinants. In Bas Verplanken (Ed.), *The Psychology of Habit* (pp. 207–230). Springer Nature Switzerland AG.

Gardner, B, & Rebar, A. L. (2019). Habit Formation and Behavior Change. In *Oxford Research Encyclopedia of Psychology*. Oxford University Press. https://doi.org/10.1093/acrefore/9780190236557.013.129

Gardner, B, Rebar, A. L., & Lally, P. (2020a). Habit Interventions. In M. S. Hagger, L. Cameron, K. Hamilton, N. Hankonen, & T. Lintunen (Eds.), *The Handbook of Behavior Change* (pp. 599–616). Cambridge University Press. https://doi.org/10.1017/9781108677318.041

Gardner, B, Rebar, A., & Lally, P. (2021). How does habit form? Guidelines for tracking real-world habit formation. *Preprint*.

Gardner, Benjamin. (2012). Habit as automaticity, not frequency . *The European Health Psychologist*, 32–36. https://www.researchgate.net/publication/230576965

Gardner, Benjamin. (2015). A review and analysis of the use of 'habit' in understanding, predicting and influencing health-related behaviour. *Health Psychology Review*, *9*(3), 277–295. https://doi.org/10.1080/17437199.2013.876238

Gardner, Benjamin. (2022). Habit and behavioural complexity: habitual instigation and execution as predictors of simple and complex behaviours. *Current Research in Behavioral Sciences*, *3*, 100081. https://doi.org/10.1016/J.CRBEHA.2022.100081

Gardner, Benjamin, Arden, M. A., Brown, D., Eves, F. F., Green, J., Hamilton, K., Hankonen, N., Inauen, J., Keller, J., Kwasnicka, D., Labudek, S., Marien, H., Masaryk, R., McCleary, N., Mullan, B. A., Neter, E., Orbell, S., Potthoff, S., & Lally, P. (2021). Developing habit-based health behaviour change interventions: twenty-one questions to guide future research. *Psychology and Health*. https://doi.org/10.1080/08870446.2021.2003362

Gardner, Benjamin, de Bruijn, G.-J., & Lally, P. (2011). A Systematic Review and Meta-analysis of Applications of the Self-Report Habit Index to Nutrition and Physical Activity Behaviours. *Annals of Behavioral Medicine*, *42*(2), 174–187. https://doi.org/10.1007/s12160-011-9282-0

Gardner, Benjamin, & Lally, P. (2013). Does intrinsic motivation strengthen physical activity habit? Modeling relationships between self-determination, past behaviour, and habit strength. *Journal of Behavioral Medicine*, *36*(5), 488–497. https://doi.org/10.1007/s10865-012-9442-0

Gardner, Benjamin, & Lally, P. (2022). Habit and habitual behaviour. *Health Psychology Review*. https://doi.org/10.1080/17437199.2022.2105249

Gardner, Benjamin, Lally, P., & Rebar, A. L. (2020). Does habit weaken the relationship between intention and behaviour? Revisiting the habit-intention interaction hypothesis. *Social and Personality Psychology Compass*, *14*(8), e12553. https://doi.org/10.1111/SPC3.12553

Gardner, Benjamin, Phillips, L. A., & Judah, G. (2016). Habitual instigation and habitual execution: Definition, measurement, and effects on behaviour frequency. *British Journal of Health Psychology*, *21*(3), 613–630. https://doi.org/10.1111/BJHP.12189

Gardner, Benjamin, & Tang, V. (2014). Reflecting on non-reflective action: An exploratory think-aloud study of self-report habit measures. *British Journal of Health Psychology*, *19*(2), 258–273. https://doi.org/10.1111/BJHP.12060

Glass, T. R., Battegay, M., Cavassini, M., De Geest, S., Furrer, H., Vernazza, P. L., Hirschel, B., Bernasconi, E., Rickenbach, M., Günthard, H. F., & Bucher, H. C. (2010). Longitudinal Analysis of Patterns and Predictors of Changes in Self-Reported Adherence to Antiretroviral Therapy: Swiss HIV Cohort Study. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, *54*(2), 197–203. https://doi.org/10.1097/QAI.0b013e3181ca48bf

Gregoriano, C., Dieterle, T., Breitenstein, A. L., Dürr, S., Baum, A., Giezendanner, S., Maier, S., Leuppi-Taegtmeyer, A., Arnet, I., Hersberger, K. E., & Leuppi, J. D. (2019). Does a tailored intervention to promote adherence in patients with chronic lung disease affect exacerbations? A randomized controlled trial. *Respiratory Research*, *20*(1), 273. https://doi.org/10.1186/s12931-

019-1219-3

Guyon, A., Gunn, S., Nikravesh, M., & Zadeh, L. (2006). *Feature extraction book: Foundations and applications* (A. Guyon, S. Gunn, M. Nikravesh, & L. Zadeh (eds.)). Springer. http://clopinet.com/fextract-book/

Hagger, M., Cameron, L., Hamilton, K., Hankonen, N., & Lintunen, T. (2021). Changing Behavior : A Theory-and Evidence-Based Approach. In M. Hagger, L. Cameron, K. Hamilton, N. Hankonen, & T. Lintunen (Eds.), *The Handbook of Behaviour Change* (pp. 1–14). Cambridge University Press. Cambridge Handbooks in Psychology. https://doi.org/10.1017/9781108677318.001

Hagger, M. S. (2019). Habit and physical activity: Theoretical advances, practical implications, and agenda for future research. *Psychology of Sport and Exercise*, *42*, 118–129. https://doi.org/10.1016/J.PSYCHSPORT.2018.12.007

Hagger, M. S., Moyers, S., McAnally, K., & McKinley, L. E. (2020). Known knowns and known unknowns on behavior change interventions and mechanisms of action. *Health Psychology Review*, *14*(1), 199–212. https://doi.org/10.1080/17437199.2020.1719184

Hagger, M. S., Rebar, A. L., Mullan, B., Lipp, O. V., & Chatzisarantis, N. L. D. (2015). The subjective experience of habit captured by self-report indexes may lead to inaccuracies in the measurement of habitual action. *Health Psychology Review*, *9*(3), 296–302. https://doi.org/10.1080/17437199.2014.959728

Hayes, A. F. (2022). *Introduction to mediation, moderation, and conditional process analysis : a regression-based approach* (2nd ed.). Guilford Publications.

Haynes, R. B., Ackloo, E., Sahota, N., McDonald, H. P., & Yao, X. (2008). Interventions for enhancing medication adherence. *Cochrane Database of Systematic Reviews*, *2*. https://doi.org/10.1002/14651858.CD000011.pub3

Haynes, R. B., McDonald, H. P., & Garg, A. X. (2002). Helping patients follow prescribed treatment: Clinical applications. *Journal of the American Medical Association*, *288*(22), 2880–2883. https://doi.org/10.1001/jama.288.22.2880

Haynes, R. B., McKibbon, K. A., & Kanani, R. (1996). Systematic review of randomised trials of interventions to assist patients to follow prescriptions for medications. In *Lancet* (Vol. 348, Issue 9024, pp. 383–386). Lancet Publishing Group. https://doi.org/10.1016/S0140-6736(96)01073-2

Haynes, R., Gibson, E. ., Hackett, B. ., Sackett, D. ., Taylor, D. ., Roberts, R. ., & Johnson, A. . (1976). Improvement of medication compliance in uncontrolled hypertension. *The Lancet*, *307*(7972), 1265–1268. https://doi.org/10.1016/s0140-6736(76)91737-2

Heijerman, H., Westerman, E., Conway, S., & Touw, D. (2009). Inhaled medication and inhalation devices for lung disease in patients with cystic fibrosis: A European consensus. In *Journal of Cystic Fibrosis* (Vol. 8, Issue 5, pp. 295–315). https://doi.org/10.1016/j.jcf.2009.04.005

Heino, M. T. J., Knittle, K., Noone, C., Hasselman, F., & Hankonen, N. (2021). Studying Behaviour Change Mechanisms under Complexity. *Behavioral Sciences*, *11*(5), 77. https://doi.org/10.3390/BS11050077

Hewer, S. L. (2012). Inhaled antibiotics in cystic fibrosis: what's new? *Journal of the Royal Society of Medicine*, *105*(2_suppl), 19–24. https://doi.org/10.1258/jrsm.2012.12s004

Higgins, J. P., Savovic, J., Page, M., Sterne, J. A., & RoB2 Development Group. (2019). *Revised Cochrane risk-of-bias tool for randomized trials (RoB 2)*.

https://sites.google.com/site/riskofbiastool/welcome/rob-2-0-tool/current-version-of-rob-2

Higgins, J., Savović, J., Page, M., Elbers, R., & Sterne, J. (2022). Chapter 8: Assessing risk of bias in a randomized trial. In J. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. Page, & V. Welch (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions* (version 6.). Cochrane. www.training.cochrane.org/handbook.

Hind, D., Drabble, S. J., Arden, M. A., Mandefield, L., Waterhouse, S., Maguire, C., Cantrill, H., Robinson, L., Beever, D., Scott, A. J., Keating, S., Hutchings, M., Bradley, J., Nightingale, J., Allenby, M. I., Dewar, J., Whelan, P., Ainsworth, J., Walters, S. J., … Wildman, M. J. (2019). Supporting medication adherence for adults with cystic fibrosis: A randomised feasibility study. *BMC Pulmonary Medicine*, *19*(1), 77. https://doi.org/10.1186/s12890-019-0834-6

Ho, P. M., Magid, D. J., Shetterly, S. M., Olson, K. L., Maddox, T. M., Peterson, P. N., Masoudi, F. A., & Rumsfeld, J. S. (2008). Medication nonadherence is associated with a broad range of adverse outcomes in patients with coronary artery disease. *American Heart Journal*, *155*(4), 772–779. https://doi.org/10.1016/j.ahj.2007.12.011

Hommel, K. A., McGrady, M. E., Peugh, J., Zacur, G., Loreaux, K., Saeed, S., Williams, E., & Denson, L. A. (2017). Longitudinal Patterns of Medication Nonadherence and Associated Health Care Costs. *Inflammatory Bowel Diseases*, *23*(9), 1577–1583. https://doi.org/10.1097/MIB.0000000000001165

Hoo, Z. H., Boote, J., Wildman, M. J., Campbell, M. J., & Gardner, B. (2017). Determinants of objective adherence to nebulised medications among adults with cystic fibrosis: an exploratory mixed methods study comparing low and high adherers. *Health Psychology & Behavioural Medicine*, *5*(1), 299–316. https://doi.org/10.1080/21642850.2017.1338958

Hoo, Zhe H., Campbell, M. J., Curley, R., & Wildman, M. J. (2017). An empirical method to cluster objective nebulizer adherence data among adults with cystic fibrosis. *Patient Preference and Adherence*, *11*, 631–642. https://doi.org/10.2147/PPA.S131497

Hoo, Zhe Hui, Coates, E., Maguire, C., Cantrill, H., Shafi, N., Nash, E. F., McGowan, A., Bourke, S. J., Flight, W. G., Daniels, T. V., Nightingale, J. A., Allenby, M. I., Curley, R., & Wildman, M. J. (2018). Pragmatic criteria to define chronic pseudomonas aeruginosa infection among adults with cystic fibrosis. *European Journal of Clinical Microbiology & Infectious Diseases : Official Publication of the European Society of Clinical Microbiology*, *37*(11), 2219–2222. https://doi.org/10.1007/S10096-018-3358-8

Hoo, Zhe Hui, Gardner, B., Arden, M. A., Waterhouse, S., Walters, S. J., Campbell, M. J., Hind, D., Maguire, C., Dewar, J., & Wildman, M. J. (2019). Role of habit in treatment adherence among adults with cystic fibrosis. *Thorax*, *74*(2), 197–199. https://doi.org/10.1136/THORAXJNL-2017-211453

Hoo, Zhe Hui, Wildman, M. J., Campbell, M. J., Walters, S. J., & Gardner, B. (2019). A pragmatic behavior-based habit index for adherence to nebulized treatments among adults with cystic fibrosis. *Patient Preference and Adherence*, *13*, 294. https://doi.org/10.2147/PPA.S186417

Hoo, Zhe Hui, Wildman, M. J., Curley, R., Walters, S. J., & Campbell, M. J. (2018). Rescue therapy within the UK Cystic Fibrosis Registry: An exploration of predictors of intravenous antibiotic use amongst adults with CF. *Respirology*, *23*(2), 190–197. https://doi.org/10.1111/resp.13174

Horne, R., & Weinman, J. (1999). Patients' beliefs about prescribed medicines and their role in adherence to treatment in chronic physical illness. *Journal of Psychosomatic Research*, *47*(6), 555–567. https://doi.org/10.1016/S0022-3999(99)00057-4

Horne, R., Weinman, J., & Hankins, M. (1999a). The beliefs about medicines questionnaire: The development and evaluation of a new method for assessing the cognitive representation of medication. *Psychology & Health*, *14*(1), 1–24. https://doi.org/10.1080/08870449908407311

Hull, C. (1943). *Principles of behavior: An introduction to behavior theory*. Appleton-Century.

Jackevicius, C. A., Mamdani, M., & Tu, J. V. (2002). Adherence with statin therapy in elderly patients with and without acute coronary syndromes. *Journal of the American Medical Association*, *288*(4), 462–467. https://doi.org/10.1001/jama.288.4.462

Jeffery, R. A., Navarro, T., Wilczynski, N. L., Iserman, E. C., Keepanasseril, A., Sivaramalingam, B., Agoritsas, T., & Haynes, R. B. (2014). Adherence measurement and patient recruitment methods are poor in intervention trials to improve patient adherence. In *Journal of Clinical Epidemiology* (Vol. 67, Issue 10, pp. 1076–1082). Elsevier USA. https://doi.org/10.1016/j.jclinepi.2014.06.008

Jenkins, B. N., Hunter, J. F., Richardson, M. J., Conner, T. S., & Pressman, S. D. (2019). Affect Variability and Predictability: Using Recurrence Quantification Analysis to Better Understand How the Dynamics of Affect Relate to Health. *Emotion*, *20*(3), 391–402. https://doi.org/10.1037/EMO0000556

Ji, M. F., & Wood, W. (2007). Purchase and Consumption Habits: Not Necessarily What You Intend. *Journal of Consumer Psychology*, *17*(4), 261–276. https://doi.org/10.1016/S1057-7408(07)70037-2

Judah, G., Gardner, B., Kenward, M. G., DeStavola, B., & Aunger, R. (2018). Exploratory study of the impact of perceived reward on habit formation. *BMC Psychology*, *6*(1), 1–12. https://doi.org/10.1186/S40359-018-0270-Z/TABLES/6

Kaushal, N., & Rhodes, R. E. (2015). Exercise habit formation in new gym members: a longitudinal study. *Journal of Behavioral Medicine*, *38*(4), 652–663. https://doi.org/10.1007/S10865-015-9640-7/TABLES/3

Keller, J., Kwasnicka, D., Klaiber, P., Sichert, L., Lally, P., & Fleig, L. (2021). Habit formation following routine-based versus time-based cue planning: A randomized controlled trial. *British Journal of Health Psychology*. https://doi.org/10.1111/bjhp.12504

Kilb, M., & Labudek, S. (2022). Effects of behavioral performance, intrinsic reward value, and context stability on the formation of a higher-order nutrition habit: an intensive longitudinal diary study. *International Journal of Behavioral Nutrition and Physical Activity*, *19*(1), 1–14. https://doi.org/10.1186/S12966-022-01343-8/TABLES/3

Kini, V., & Michael Ho, P. (2018). Interventions to Improve Medication Adherence: A Review. *JAMA*, *320*(23), 2461–2473. https://doi.org/10.1001/JAMA.2018.19271

Knafl, G. J., Delucchi, K., Bova, C., Fennie, K., Ding, K., & Williams, A. (2010). A systematic approach for analyzing electronically monitored adherence data. In B. Ekwall & M. Cronquist (Eds.), *Micro Electro Mechanical Systems*. Nova Science Publishers, Inc.

Knafl, G. J., Schoenthaler, A., & Ogedegbe, G. (2012). Secondary analysis of electronically monitored medication adherence data for a cohort of hypertensive African-Americans. *Patient Preference and Adherence*, *6*, 207–219. https://doi.org/10.2147/PPA.S30582

Kok, G., Gottlieb, N. H., Peters, G.-J. Y., Mullen, P. D., Parcel, G. S., Ruiter, R. A. C., Fernández, M. E., Markham, C., & Kay Bartholomew, L. (2016). A taxonomy of behaviour change methods: an intervention mapping approach. *Health Psychology Review*, *10*(3), 297–312. https://doi.org/10.1080/17437199.2015.1077155

Kor, K., & Mullan, B. A. (2011). Sleep hygiene behaviours: An application of the theory of planned behaviour and the investigation of perceived autonomy support, past behaviour and response inhibition. *Psychology & Health*, *26*(9), 1208–1224. https://doi.org/10.1080/08870446.2010.551210

Kwasnicka, D., Dombrowski, S. U., White, M., & Sniehotta, F. (2016b). Theoretical explanations for maintenance of behaviour change: a systematic review of behaviour theories. *Health Psychology Review*, *10*(3), 277–296. https://doi.org/10.1080/17437199.2016.1151372

Kwasnicka, D., Dombrowski, S. U., White, M., & Sniehotta, F. F. (2017). N-of-1 study of weight loss maintenance assessing predictors of physical activity, adherence to weight loss plan and weight change. *Psychology and Health*, *32*(6), 686–708. https://doi.org/10.1080/08870446.2017.1293057

Kwasnicka, D., Konrad, B., Kronish, I. M., & Davidson, K. (2018). Using N-of-1 Methods to Explore Habit Formation. In Bas Verplanken (Ed.), *The Psychology of Habit* (pp. 231–246). Springer Nature Switzerland AG.

Kwasnicka, D., & Naughton, F. (2020). N-of-1 methods: A practical guide to exploring trajectories of behaviour change and designing precision behaviour change interventions. *Psychology of Sport and Exercise*, *47*, 101570. https://doi.org/10.1016/j.psychsport.2019.101570

Labrecque, J. S., & Wood, W. (2015). What measures of habit strength to use? Comment on Gardner (2015). *Https://Doi.Org/10.1080/17437199.2014.992030*, *9*(3), 303–310. https://doi.org/10.1080/17437199.2014.992030

Lally, P., & Gardner, B. (2013). Promoting habit formation. *Health Psychology Review*, *7*(sup1), S137–S158. https://doi.org/10.1080/17437199.2011.603640

Lally, P., van Jaarsveld, C. H. M., Potts, H. W. W., & Wardle, J. (2010). How are habits formed: Modelling habit formation in the real world. *European Journal of Social Psychology*, *40*(6), 998–1009. https://doi.org/10.1002/ejsp.674

Latchford, G., Duff, A., Quinn, J., Conway, S., & Conner, M. (2009). Adherence to nebulised antibiotics in cystic fibrosis. *Patient Education and Counseling*, *75*(1), 141–144. https://doi.org/10.1016/j.pec.2008.08.027

Lee, J. A., Cho, A., Huang, E. N., Xu, Y., Quach, H., Hu, J., & Wong, A. P. (2021). Gene therapy for cystic fibrosis: new tools for precision medicine. *Journal of Translational Medicine 2021 19:1*, *19*(1), 1–15. https://doi.org/10.1186/S12967-021-03099-4

Lee, W. C., Balu, S., Cobden, D., Joshi, A. V., & Pashos, C. L. (2006). Prevalence and economic consequences of medication adherence in diabetes: A systematic literature review. In *Managed Care Interface* (Vol. 19, Issue 7, pp. 31–41).

Lehmann, A., Aslani, P., Ahmed, R., Celio, J., Gauchet, A., Bedouch, P., Bugnon, O., Allenet, B., & Schneider, M. P. (2013). Assessing medication adherence: options to consider. *International Journal of Clinical Pharmacy*, *36*(1), 55–69. https://doi.org/10.1007/s11096-013-9865-x

Lin, A. H. A., Kendrick, J. G., Wilcox, P. G., & Quon, B. S. (2017). Patient knowledge and pulmonary medication adherence in adult patients with cystic fibrosis. *Patient Preference and Adherence*, *11*, 691–698. https://doi.org/10.2147/PPA.S129088

Lin, C.-Y., Yaseri, M., Pakpour, A. H., Malm, D., Broström, A., Fridlund, B., Burri, A., & Webb, T. L. (2017). Can a Multifaceted Intervention Including Motivational Interviewing Improve Medication Adherence, Quality of Life, and Mortality Rates in Older Patients Undergoing Coronary Artery Bypass Surgery? A Multicenter, Randomized Controlled Trial with 18-Month

Fo. *Drugs & Aging*, *34*(2), 143–156. https://doi.org/10.1007/s40266-016-0429-3

Maher, J. P., Rebar, A. L., & Dunton, G. F. (2021). The influence of context stability on physical activity and sedentary behaviour habit and behaviour: An ecological momentary assessment study. *British Journal of Health Psychology*, *26*(3), 861–881. https://doi.org/10.1111/BJHP.12509

Mäkelä, M. J., Backer, V., Hedegaard, M., & Larsson, K. (2013). Adherence to inhaled therapies, health outcomes and costs in patients with asthma and COPD. *Respiratory Medicine*, *107*(10), 1481–1490. https://doi.org/10.1016/j.rmed.2013.04.005

Mantzari, E., Vogt, F., Shemilt, I., Wei, Y., Higgins, J. P. T., & Marteau, T. M. (2015). Personal financial incentives for changing habitual health-related behaviors: A systematic review and meta-analysis. In *Preventive Medicine* (Vol. 75, pp. 75–85). Academic Press Inc. https://doi.org/10.1016/j.ypmed.2015.03.001

Marhe, R., Waters, A. J., Van De Wetering, B. J. M., & Franken, I. H. A. (2013). Implicit and explicit drug-related cognitions during detoxification treatment are associated with drug relapse: An ecological momentary assessment study. *Journal of Consulting and Clinical Psychology*, *81*(1), 1–12. https://doi.org/10.1037/a0030754

Markowitz, M., Karve, S., Panish, J., Candrilli, S. D., & Alphs, L. (2013). Antipsychotic adherence patterns and health care utilization and costs among patients discharged after a schizophrenia-related hospitalization. *BMC Psychiatry*, *13*(1), 246. https://doi.org/10.1186/1471-244X-13-246

Marshall, I. J., & Wallace, B. C. (2019). Toward systematic review automation: a practical guide to using machine learning tools in research synthesis. *Systematic Reviews*, *8*(1), 163. https://doi.org/10.1186/s13643-019-1074-9

Marwan, N., Carmen Romano, M., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports*, *438*(5–6), 237–329. https://doi.org/10.1016/j.physrep.2006.11.001

Mazar, A., & Wood, W. (2019). Defining Habit in Psychology. In B Verplanken (Ed.), *The Psychology of Habit* (pp. 13–30). Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-319-97529-0

McCloskey, K., & Johnson, B. T. (2019). Habits, Quick and Easy: Perceived Complexity Moderates the Associations of Contextual Stability and Rewards With Behavioral Automaticity. *Frontiers in Psychology*, *10*, 1556. https://doi.org/10.3389/FPSYG.2019.01556/BIBTEX

McEachan, R. R. C., Conner, M., Taylor, N. J., & Lawton, R. J. (2011). Prospective prediction of health-related behaviours with the theory of planned behaviour: A meta-analysis. In *Health Psychology Review* (Vol. 5, Issue 2, pp. 97–144). Taylor & Francis Group . https://doi.org/10.1080/17437199.2010.521684

McKenzie, J. E., & Brennan, S. E. (2019). Synthesizing and presenting findings using other methods. In *Cochrane Handbook for Systematic Reviews of Interventions* (pp. 321–347). Wiley. https://doi.org/10.1002/9781119536604.ch12

Michell, J. (2008). Is Psychometrics Pathological Science? *Measurement: Interdisciplinary Research and Perspectives*, *6*(1–2), 7–24. https://doi.org/10.1080/15366360802035489

Michie, S, Atkins, L., & West, R. (2014a). *The Behaviour Change Wheel: A guide to designing interventions*. Silverback Publishing. http://www.behaviourchangewheel.com/

Michie, Susan, & Prestwich, A. (2010). Are interventions theory-based? Development of a theory

coding scheme. *Health Psychology : Official Journal of the Division of Health Psychology, American Psychological Association*, *29*(1), 1–8. https://doi.org/10.1037/a0016939

Michie, Susan, Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., Eccles, M. P., Cane, J., & Wood, C. E. (2013). The Behavior Change Technique Taxonomy (v1) of 93 Hierarchically Clustered Techniques: Building an International Consensus for the Reporting of Behavior Change Interventions. *Annals of Behavioral Medicine*, *46*(1), 81–95. https://doi.org/10.1007/s12160-013-9486-6

Michie, Susan, 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

Milam, J., Richardson, J. L., McCutchan, A., Stoyanoff, S., Weiss, J., Kemper, C., Larsen, R. A., Hollander, H., Weismuller, P., & Bolan, R. (2005). Effect of a Brief Antiretroviral Adherence Intervention Delivered by HIV Care Providers. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, *40*(3), 356–363. https://doi.org/10.1097/01.qai.0000159710.98960.81

Modi, A. C., Cassedy, A. E., Quittner, A. L., Accurso, F., Sontag, M., Koenig, J. M., & Ittenbach, R. F. (2010). Trajectories of Adherence to Airway Clearance Therapy for Patients with Cystic Fibrosis. *Journal of Pediatric Psychology*, *35*(9), 1028–1037. https://doi.org/10.1093/jpepsy/jsq015

Modi, A. C., Lim, C. S., Yu, N., Geller, D., Wagner, M. H., & Quittner, A. L. (2006). A multi-method assessment of treatment adherence for children with cystic fibrosis. *Journal of Cystic Fibrosis*, *5*(3), 177–185. https://doi.org/10.1016/j.jcf.2006.03.002

Mogayzel, P. J., Naureckas, E. T., Robinson, K. A., Mueller, G., Hadjiliadis, D., Hoag, J. B., Lubsch, L., Hazle, L., Sabadosa, K., & Marshall, B. (2013). Cystic fibrosis pulmonary guidelines: Chronic medications for maintenance of lung health. In *American Journal of Respiratory and Critical Care Medicine* (Vol. 187, Issue 7, pp. 680–689). American Thoracic Society. https://doi.org/10.1164/rccm.201207-1160OE

Mohideen, A., Bouvin, C., Judah, G., Picariello, F., & Gardner, B. (2023). Feasibility and acceptability of a personalised script-elicitation method for improving evening sleep hygiene habits. *Health Psychology & Behavioural Medicine*, *11*(1). https://doi.org/10.1080/21642850.2022.2162904

Munro, S., Lewin, S., Swart, T., & Volmink, J. (2007). A review of health behaviour theories: How useful are these for developing interventions to promote long-term medication adherence for TB and HIV/AIDS? In *BMC Public Health* (Vol. 7, Issue 1, p. 104). BioMed Central. https://doi.org/10.1186/1471-2458-7-104

Naab, T. K., & Schnauber, A. (2016). Habitual Initiation of Media Use and a Response-Frequency Measure for Its Examination. *Media Psychology*, *19*(1), 126–155. https://doi.org/10.1080/15213269.2014.951055

National Institute for Health and Care Excellence. (2017). *Cystic fibrosis: diagnosis and management*. NICE. https://www.nice.org.uk/guidance/ng78/chapter/Recommendations#pulmonary-monitoring-assessment-and-management

Neal, D. T., Wood, W., & Drolet, A. (2013). How do people adhere to goals when willpower is low? The profits (and pitfalls) of strong habits. *Journal of Personality and Social Psychology*, *104*(6), 959–975. https://doi.org/10.1037/A0032626

Neal, D. T., Wood, W., Labrecque, J. S., & Lally, P. (2012). How do habits guide behavior? Perceived and actual triggers of habits in daily life. *Journal of Experimental Social Psychology*, *48*(2), 492–498. https://doi.org/10.1016/J.JESP.2011.10.011

Neal, D. T., Wood, W., Wu, M., & Kurlander, D. (2011). The pull of the past: When do habits persist despite conflict with motives? *Personality and Social Psychology Bulletin*, *37*(11), 1428–1437. https://doi.org/10.1177/0146167211419863

Nguyen, T. T. D., Thia, L. P., Hoo, A. F., Bush, A., Aurora, P., Wade, A., Chudleigh, J., Lum, S., & Stocks, J. (2014). Evolution of lung function during the first year of life in newborn screened cystic fibrosis infants. *Thorax*, *69*(10), 910–917. https://doi.org/10.1136/thoraxjnl-2013-204023

Nielsen, L., Riddle, M., King, J. W., Aklin, W. M., Chen, W., Clark, D., Collier, E., Czajkowski, S., Esposito, L., Ferrer, R., Green, P., Hunter, C., Kehl, K., King, R., Onken, L., Simmons, J. M., Stoeckel, L., Stoney, C., Tully, L., & Weber, W. (2018). The NIH Science of Behavior Change Program: Transforming the science through a focus on mechanisms of change. *Behaviour Research and Therapy*, *101*, 3–11. https://doi.org/10.1016/j.brat.2017.07.002

Nieuwlaat, R., Wilczynski, N., Navarro, T., Hobson, N., Jeffery, R., Keepanasseril, A., Agoritsas, T., Mistry, N., Iorio, A., Jack, S., Sivaramalingam, B., Iserman, E., Mustafa, R. A., Jedraszewski, D., Cotoi, C., & Haynes, R. B. (2014). Interventions for enhancing medication adherence. *Cochrane Database of Systematic Reviews*, *11*. https://doi.org/10.1002/14651858.CD000011.pub4

Nilsen, W. J., Haverkos, L., Nebeling, L., & Taylor, M. V. (2010). Maintenance of long-term behavior change. *American Journal of Health Behavior*, *34*(6), 643–646. https://doi.org/10.5993/AJHB.34.6.1

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231–259. https://doi.org/10.1037/0033-295X.84.3.231

Nixon, G. M., Armstrong, D. S., Carzino, R., Carlin, J. B., Olinsky, A., Robertson, C. F., & Grimwood, K. (2001). Clinical outcome after early Pseudomonas aeruginosa infection in cystic fibrosis. *Journal of Pediatrics*, *138*(5), 699–704. https://doi.org/10.1067/mpd.2001.112897

Norman, P, & Connor, M. (2015). Predicting and changing health behaviour: Future directions. In P Norman & M. Connor (Eds.), *Predicting and changing health behaviour: Research and practice with social cognition models* (3rd ed., pp. 390–430). Open University Press.

Norman, Paul, & Cooper, Y. (2011). The theory of planned behaviour and breast self-examination: Assessing the impact of past behaviour, context stability and habit strength. *Psychology & Health*, *26*(9), 1156–1172. https://doi.org/10.1080/08870446.2010.481718

O'Dwyer, S., Greene, G., MacHale, E., Cushen, B., Sulaiman, I., Boland, F., Bosnic-Anticevich, S., Mokoka, M. C., Reilly, R. B., Taylor, T., Ryder, S. A., & Costello, R. W. (2020). Personalized Biofeedback on Inhaler Adherence and Technique by Community Pharmacists: A Cluster Randomized Clinical Trial. *The Journal of Allergy and Clinical Immunology: In Practice*, *8*(2), 635–644. https://doi.org/10.1016/j.jaip.2019.09.008

O'Mara-Eves, A., Thomas, J., McNaught, J., Miwa, M., & Ananiadou, S. (2015). Using text mining for study identification in systematic reviews: a systematic review of current approaches. *Systematic Reviews*, *4*(1). https://doi.org/10.1186/s13643-015-0031-5

O'Sullivan, B. P., & Freedman, S. D. (2009). Cystic fibrosis. *The Lancet*, *373*, 1891–1904. https://doi.org/10.1016/S0140

Orbell, S., & Verplanken, B. (2010). The automatic component of habit in health behavior: Habit as cue-contingent automaticity. *Health Psychology*, *29*(4), 374–383. https://doi.org/10.1037/a0019596

Orbell, S., & Verplanken, B. (2015). The strength of habit. *Health Psychology Review*, *9*(3), 311–317. https://doi.org/10.1080/17437199.2014.992031

Orbell, S., & Verplanken, B. (2020). Changing behaviour using habit theory. In M. Hagger, L. Cameron, K. Hamilton, N. Hankonen, & T. Lintunen (Eds.), *The Handbook of Behaviour Change* (pp. 178–192). Cambridge University Press.

Ory, M. G., Smith, M. L., Mier, N., & Wernicke, M. M. (2010). The science of sustaining health behavior change: The health maintenance consortium. *American Journal of Health Behavior*, *34*(6), 647–659. https://doi.org/10.5993/AJHB.34.6.2

Osterberg, L., & Blaschke, T. (2005). Adherence to Medication. *New England Journal of Medicine*, *353*(5), 487–497. https://doi.org/10.1056/nejmra050100

Ouellette, J. A., & Wood, W. (1998). Habit and Intention in Everyday Life: The Multiple Processes by Which Past Behavior Predicts Future Behavior. *Psychological Bulletin*, *124*(1), 54–74. https://doi.org/10.1037/0033-2909.124.1.54

Pakpour, A. H., Gholami, M., Esmaeili, R., Naghibi, S. A., Updegraff, J. A., Molloy, G. J., & Dombrowski, S. U. (2015). A randomized controlled multimodal behavioral intervention trial for improving antiepileptic drug adherence. *Epilepsy & Behavior*, *52*, 133–142. https://doi.org/10.1016/j.yebeh.2015.08.036

Phillips, A., & Gardner, B. (2016). Habitual exercise instigation (vs. execution) predicts healthy adults' exercise frequency. *Health Psychology*, *35*(1), 69–77. https://psycnet.apa.org/buy/2015-30036-001

Phillips, A., Leventhal, H., & Leventhal, E. A. (2013). Assessing theoretical predictors of long-term medication adherence: Patients' treatment-related beliefs, experiential feedback and habit development. *Psychology & Health*, *28*(10), 1135–1151. https://doi.org/10.1080/08870446.2013.793798

Phillips, A., Pironet, A., & Vrijens, B. (2022). Evaluating Objective Metrics of Habit Strength for Taking Medications. *Society of Behavioural Medicine 2022*.

Phillips, L. A. (2011). *Using habit theory and commonsense self-regulation theory to improve medication adherence assessment.*

Phillips, L. A. (2020). Challenging assumptions about habit: A response to Hagger (2019). *Psychology of Sport and Exercise*, *47*. https://doi.org/10.1016/j.psychsport.2019.03.005

Phillips, L. A., Burns, E., & Leventhal, H. (2021). Time-of-Day Differences in Treatment-Related Habit Strength and Adherence. *Annals of Behavioral Medicine*, *55*(3), 280–285. https://doi.org/10.1093/ABM/KAAA042

Phillips, L. A., Chamberland, P. É., Hekler, E. B., Abrams, J., & Eisenberg, M. H. (2016). Intrinsic rewards predict exercise via behavioral intentions for initiators but via habit strength for maintainers. *Sport, Exercise, and Performance Psychology*, *5*(4), 352–364. https://doi.org/10.1037/spy0000071

Phillips, L. A., Cohen, J., Burns, E., Abrams, J., & Renninger, S. (2016). Self-management of chronic illness: the role of 'habit' versus reflective factors in exercise and medication adherence. *Journal of Behavioral Medicine*, *39*(6), 1076–1091. https://doi.org/10.1007/S10865-016-9732-Z/TABLES/7

Phillips, L. A., & Mullan, B. A. (2022). Ramifications of behavioural complexity for habit conceptualisation, promotion, and measurement. *Health Psychology Review*. https://doi.org/10.1080/17437199.2022.2060849

Pimm, R., Vandelanotte, C., Rhodes, R. E., Short, C., Duncan, M. J., & Rebar, A. L. (2015). Cue

Consistency Associated with Physical Activity Automaticity and Behavior. *Http://Dx.Doi.Org/10.1080/08964289.2015.1017549*, *42*(4), 248–253. https://doi.org/10.1080/08964289.2015.1017549

Podsadecki, T. J., Vrijens, B. C., Tousset, E. P., Rode, R. A., & Hanna, G. J. (2008). "White Coat Compliance" Limits the Reliability of Therapeutic Drug Monitoring in HIV-1—Infected Patients. *HIV Clinical Trials*, *9*(4), 238–246. https://doi.org/10.1310/HCT0904-238

Prestwich, A., Sniehotta, F. F., Whittington, C., Dombrowski, S. U., Rogers, L., & Michie, S. (2014). Does theory influence the effectiveness of health behavior interventions? Meta-analysis. *Health Psychology*, *33*(5), 465–474. https://doi.org/10.1037/A0032853

Prochaska, J. O., & Di Clemente, C. C. (1982). Transtheoretical therapy: Toward a more integrative model of change. *Psychotherapy*, *19*(3), 276–288. https://doi.org/10.1037/h0088437

Pugatsch, T., Shoseyov, D., Cohen-Cymberknoh, M., Hayut, B., Armoni, S., Griese, M., & Kerem, E. (2016). Adherence pattern to study drugs in clinical trials by patients with cystic fibrosis. *Pediatric Pulmonology*, *51*(2), 143–146. https://doi.org/10.1002/ppul.23344

Quittner, A. L., Eakin, M. N., Alpern, A. N., Ridge, A. K., McLean, K. A., Bilderback, A., Criado, K. K., Chung, S. E., & Riekert, K. A. (2019). Clustered randomized controlled trial of a clinic-based problem-solving intervention to improve adherence in adolescents with cystic fibrosis. *Journal of Cystic Fibrosis*, *18*(6), 879–885. https://doi.org/10.1016/j.jcf.2019.05.004

Quittner, A. L., Zhang, J., Marynchenko, M., Chopra, P. A., Signorovitch, J., Yushkina, Y., & Riekert, K. A. (2014). Pulmonary medication adherence and health-care use in cystic fibrosis. *Chest*, *146*(1), 142–151. https://doi.org/10.1378/chest.13-1926

Rand, M., Goyder, E., Norman, P., & Womack, R. (2020). Why do new members stop attending health and fitness venues? The importance of developing frequent and stable attendance behaviour. *Psychology of Sport and Exercise*, *51*, 101771. https://doi.org/10.1016/J.PSYCHSPORT.2020.101771

Ratjen, F., Bell, S. C., Rowe, S. M., Goss, C. H., Quittner, A. L., & Bush, A. (2015). Cystic fibrosis. In *Nature reviews. Disease primers* (Vol. 1, Issue 1, p. 15010). Nature Publishing Group. https://doi.org/10.1038/nrdp.2015.10

Rawald, T., Sips, M., & Marwan, N. (2017). PyRQA - Conducting Recurrence Quantification Analysis on Very Long Time Series Efficiently. *Computers & Geosciences*, *104*, 101–108. https://doi.org/10.1016/j.cageo.2016.11.016

Rebar, A., Gardner, B., Rhodes, R., & Verplanken, B. (2019). The Measurement of Habit. In B Verplanken (Ed.), *The Psychology of Habit* (pp. 31–49). Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-319-97529-0

Rebar, A. L., Dimmock, J. A., Jackson, B., Rhodes, R. E., Kates, A., Starling, J., & Vandelanotte, C. (2016). A systematic review of the effects of non-conscious regulatory processes in physical activity. *Health Psychology Review*, *10*(4), 395–407. https://doi.org/10.1080/17437199.2016.1183505

Rebar, A. L., Rhodes, R. E., & Verplanken, B. (2022). Habits and behavioral complexity – dynamic and distinct constructs. *Health Psychology Review*. https://doi.org/10.1080/17437199.2022.2098163

Reddy, A., Huseman, T. L., Canamucio, A., Marcus, S. C., Asch, D. A., Volpp, K., & Long, J. A. (2017). Patient and Partner Feedback Reports to Improve Statin Medication Adherence: A Randomized Control Trial. *Journal of General Internal Medicine*, *32*(3), 256–261.

https://doi.org/10.1007/s11606-016-3858-0

Remeseiro, B., & Bolon-Canedo, V. (2019). A review of feature selection methods in medical applications. *Computers in Biology and Medicine*, *112*, 103375. https://doi.org/10.1016/J.COMPBIOMED.2019.103375

Riordan, J. R., Rommens, J. M., Kerem, B. S., Alon, N. O. A., Rozmahel, R., Grzelczak, Z., Zielenski, J., Lok, S. I., Plavsic, N., Chou, J. L., Drumm, M. L., Iannuzzi, M. C., Collins, F. S., & Tsui, L. C. (1989). Identification of the cystic fibrosis gene: Cloning and characterization of complementary DNA. *Science*, *245*(4922), 1066–1073. https://doi.org/10.1126/science.2475911

Robinson, L., Arden, M. A., Dawson, S., Walters, S. J., Wildman, M. J., & Stevenson, M. (2022). A machine-learning assisted review of the use of habit formation in medication adherence interventions for long-term conditions. *Health Psychology Review*. https://doi.org/10.1080/17437199.2022.2034516

Rommens, J. M., Iannuzzi, M. C., Kerem, B. S., Drumm, M. L., Melmer, G., Dean, M., Rozmahel, R., Cole, J. L., Kennedy, D., Hidaka, N., Zsiga, M., Buchwald, M., Riordan, J. R., Tsui, L. C., & Collins, F. S. (1989). Identification of the cystic fibrosis gene: Chromosome walking and jumping. *Science*, *245*(4922), 1059–1065. https://doi.org/10.1126/science.2772657

Rosenstock, I. M., Strecher, V. J., & Becker, M. H. (1988). Social Learning Theory and the Health Belief Model. *Health Education Quarterly*, *15*(2), 175–183.

Rothman, A. (2000). Toward a theory based analysis of behavioural maintenance. *Health Psychology*, *19*(1), 64–69.

Rothman, A. J., Sheeran, P., & Wood, W. (2009). Reflective and Automatic Processes in the Initiation and Maintenance of Dietary Change. *Annals of Behavioral Medicine*, *38*(suppl_1), s4–s17. https://doi.org/10.1007/S12160-009-9118-3

Russell, C. L., Hathaway, D., Remy, L. M., Aholt, D., Clark, D., Miller, C., Ashbaugh, C., Wakefield, M., Ye, S., Staggs, V. S., Ellis, R. J., & Goggin, K. (2020). Improving medication adherence and outcomes in adult kidney transplant patients using a personal systems approach: SystemCHANGE^{TM} results of the MAGIC randomized clinical trial. *American Journal of Transplantation*, *20*(1), 125–136. https://doi.org/10.1111/ajt.15528

Saag, K. G., Bhatia, S., Mugavero, M. J., & Singh, J. A. (2018). Taking an Interdisciplinary Approach to Understanding and Improving Medication Adherence. *Journal of General Internal Medicine*, *33*(2), 136–138. https://doi.org/10.1007/s11606-017-4215-7

Sanders, D. B., Bittner, R. C. L., Rosenfeld, M., Hoffman, L. R., Redding, G. J., & Goss, C. H. (2010). Failure to recover to baseline pulmonary function after cystic fibrosis pulmonary exacerbation. *American Journal of Respiratory and Critical Care Medicine*, *182*(5), 627–632. https://doi.org/10.1164/rccm.200909-1421OC

Sanders, D. B., Zhao, Q., Li, Z., & Farrell, P. M. (2017). Poor recovery from cystic fibrosis pulmonary exacerbations is associated with poor long-term outcomes. *Pediatric Pulmonology*, *52*(10), 1268–1275. https://doi.org/10.1002/ppul.23765

Sawicki, G. S., Sellers, D. E., & Robinson, W. M. (2009). High treatment burden in adults with cystic fibrosis: Challenges to disease self-management. *Journal of Cystic Fibrosis*, *8*(2), 91–96. https://doi.org/10.1016/j.jcf.2008.09.007

Schnauber-Stockmann, A., & Naab, T. K. (2018). The process of forming a mobile media habit: results of a longitudinal study in a real-world setting. *Media Psychology*, *22*(5), 714–742. https://doi.org/10.1080/15213269.2018.1513850

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*(1), 1–66. https://doi.org/10.1037/0033-295X.84.1.1

Sheeran, P., & Conner, M. (2019). Degree of reasoned action predicts increased intentional control and reduced habitual control over health behaviors. *Social Science & Medicine*, *228*, 68–74. https://doi.org/10.1016/J.SOCSCIMED.2019.03.015

Shemilt, I., Anneliese, A., Thomas, J., Lorenc, T., Khouja, C., Raine, G., Sutcliffe, K., D'Souza, P., Wright, K., & Sowden, A. (2021). Using automation to produce a 'living map' of the COVID-19 research literature. *Journal of EAHIL*, *17*(2), 11–15.

Shemilt, I., Noel-Storr, A., Thomas, J., Featherstone, R., & Mavergames, C. (2021). *Machine Learning Reduced Workload for the Cochrane COVID-19 Study Register: Development and Evaluation of the Cochrane COVID-19 Study Classifier*. https://assets.researchsquare.com/files/rs-689189/v1/789144f5-66e5-4303-81af-63f1c45bd793.pdf?c=1626298662

Shiffman, S., Hufford, M., Hickcox, M., Paty, J. A., Gnys, M., & Kassel, J. D. (1997). Remember that? A comparison of real-time versus retrospective recall of smoking lapses. *Journal of Consulting and Clinical Psychology*, *65*(2), 292–300. https://doi.org/10.1037/0022-006X.65.2.292.a

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*(2), 127–190. https://doi.org/10.1037/0033-295X.84.2.127

Sniehotta, F. F., & Presseau, J. (2012). The Habitual Use of the Self-report Habit Index. *Annals of Behavioral Medicine*, *43*(1), 139–140. https://doi.org/10.1007/S12160-011-9305-X

Solem, C. T., Vera-Llonch, M., Liu, S., Botteman, M., & Castiglione, B. (2016). Impact of pulmonary exacerbations and lung function on generic health-related quality of life in patients with cystic fibrosis. *Health and Quality of Life Outcomes*, *14*(1), 1–9. https://doi.org/10.1186/s12955-016-0465-z

Stacy, J. N., Schwartz, S. M., Ershoff, D., & Shreve, M. S. (2009). Incorporating Tailored Interactive Patient Solutions Using Interactive Voice Response Technology to Improve Statin Adherence: Results of a Randomized Clinical Trial in a Managed Care Setting. *Population Health Management*, *12*(5), 241–254. https://doi.org/10.1089/pop.2008.0046

Stecher, C., Berardi, V., Fowers, R., Christ, J., Chung, Y., & Huberty, J. (2021). Identifying App-Based Meditation Habits and the Associated Mental Health Benefits: Longitudinal Observational Study. *J Med Internet Res* , e27282. https://doi.org/10.2196/27282

Stirratt, M. J., Curtis, J. R., Danila, M. I., Hansen, R., Miller, M. J., & Gakumo, C. A. (2018). Advancing the Science and Practice of Medication Adherence. *Journal of General Internal Medicine*, *33*(2), 216–222. https://doi.org/10.1007/s11606-017-4198-4

Stirratt, M. J., Dunbar-Jacob, J., Crane, H. M., Simoni, J. M., Czajkowski, S., Hilliard, M. E., Aikens, J. E., Hunter, C. M., Velligan, D. I., Huntley, K., Ogedegbe, G., Rand, C. S., Schron, E., & Nilsen, W. J. (2015). Self-report measures of medication adherence behavior: recommendations on optimal use. *Translational Behavioral Medicine*, *5*(4), 470–482. https://doi.org/10.1007/s13142-015-0315-2

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*(3), 220–247. https://doi.org/10.1207/s15327957pspr0803_1

Sturgess, J., & Imrie, J. (1982). Quantitative evaluation of the development of tracheal submucosal glands in infants with cystic fibrosis and control infants. *American Journal of Pathology*, *106*(3),

303–311.

Tan, H., Cai, Q., Agarwal, S., Stephenson, J. J., & Kamat, S. (2011). Impact of adherence to disease-modifying therapies on clinical and economic outcomes among patients with multiple sclerosis. *Advances in Therapy*, *28*(1), 51–61. https://doi.org/10.1007/s12325-010-0093-7

Tang, F., Zhu, G., Jiao, Z., Ma, C., Chen, N., & Wang, B. (2014). The effects of medication education and behavioral intervention on Chinese patients with epilepsy. *Epilepsy & Behavior*, *37*, 157–164. https://doi.org/10.1016/j.yebeh.2014.05.017

Tappe, K. A., & Glanz, K. (2013). Measurement of exercise habits and prediction of leisure-time activity in established exercise. *Psychology, Health & Medicine*, *18*(5), 601–611. https://doi.org/10.1080/13548506.2013.764458

Tappenden, P., Sadler, S., & Wildman, M. (2017). An Early Health Economic Analysis of the Potential Cost Effectiveness of an Adherence Intervention to Improve Outcomes for Patients with Cystic Fibrosis. *PharmacoEconomics*, *35*(6), 647–659. https://doi.org/https://dx.doi.org/10.1007/s40273-017-0500-x

Thomas, G. O., Poortinga, W., & Sautkina, E. (2016). Habit Discontinuity, Self-Activation, and the Diminishing Influence of Context Change: Evidence from the UK Understanding Society Survey. *PLOS ONE*, *11*(4), e0153490. https://doi.org/10.1371/journal.pone.0153490

Thornton, J., Elliott, R., Tully, M. P., Dodd, M., & Webb, A. K. (2004). Long term clinical outcome of home and hospital intravenous antibiotic treatment in adults with cystic fibrosis. *Thorax*, *59*, 242–246. https://doi.org/10.1136/thx.2003.005876

Tibshirani, R. (1997). THE LASSO METHOD FOR VARIABLE SELECTION IN THE COX MODEL. *STATISTICS IN MEDICINE*, *16*, 385–395. https://doi.org/10.1002/(SICI)1097-0258(19970228)16:4

Triandis, H. C. (1977). Cross-Cultural Social and Personality Psychology. *Psychology and Social Psychology Bulletin*, *3*, 143–158.

Tricomi, E., Balleine, B. W., & O'Doherty, J. P. (2009). A specific role for posterior dorsolateral striatum in human habit learning. *European Journal of Neuroscience*, *29*(11), 2225–2232. https://doi.org/10.1111/j.1460-9568.2009.06796.x

Tsui, L. C., Buchwald, M., Barker, D., Braman, J. C., Knowlton, R., Schumm, J. W., Eiberg, H., Mohr, J., Kennedy, D., Plavsic, N., Zsiga, M., Markiewicz, D., Akots, G., Brown, V., Helms, C., Gravius, T., Parker, C., Rediker, K., & Donis-Keller, H. (1985). Cystic fibrosis locus defined by a genetically linked polymorphic DNA marker. *Science*, *230*(4729), 1054–1057. https://doi.org/10.1126/science.2997931

Tuldrà, A., Fumaz, C. R., Ferrer, M. J., Bayés, R., Arnó, A., Balagué, M., Bonjoch, A., Jou, A., Negredo, E., Paredes, R., Ruiz, L., Romeu, J., Sirera, G., Tural, C., Burger, D., & Clotet, B. (2000). Prospective Randomized Two-Arm Controlled Study To Determine the Efficacy of a Specific Intervention To Improve Long-Term Adherence to Highly Active Antiretroviral Therapy. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, *25*(3), 221–228. https://doi.org/10.1097/00126334-200011010-00003

Turcios, N. L. (2020). Cystic Fibrosis Lung Disease: An Overview. *Respiratory Care*, *65*(2), 233–251. https://doi.org/10.4187/RESPCARE.06697

Valenstein, M., Ganoczy, D., McCarthy, J. F., Myra Kim, H., Lee, T. A., & Blow, F. C. (2006). Antipsychotic adherence over time among patients receiving treatment for schizophrenia: a retrospective review. *The Journal of Clinical Psychiatry*, *67*(10), 1542–1550. https://doi.org/10.4088/jcp.v67n1008

van Boven, J. F. M., Chavannes, N. H., van der Molen, T., Rutten-van Molken, M. P. M. H., Postma, M. J., & Vegter, S. (2014). Clinical and economic impact of non-adherence in COPD: a systematic review. *Respiratory Medicine*, *108*(1), 103–113. https://doi.org/https://dx.doi.org/10.1016/j.rmed.2013.08.044

Verplanken, B, Myrbakk, V., & Rudi, E. (2005). Measuring Habit. In T. Betsch & S. Haberstroh (Eds.), *The Routines of Decision Making* (pp. 231–247). Psychology Press.

Verplanken, B, Roy, D., & Whitmarsh, L. (2018). Cracks in the Wall: Habit Discontinuities as Vehicles for Behaviour Change. In B Verplanken (Ed.), *The Psychology of Habit* (pp. 189–205). Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-319-97529-0_11

Verplanken, Bas. (2006). Beyond frequency: Habit as mental construct. *British Journal of Social Psychology*, *45*, 639–656. https://doi.org/10.1348/014466605X49122

Verplanken, Bas, & Orbell, S. (2003). Reflections on Past Behavior: A Self-Report Index of Habit Strength. *Journal of Applied Social Psychology*, *33*(6), 1313–1330. https://doi.org/10.1111/J.1559-1816.2003.TB01951.X

Verplanken, Bas, Walker, I., Davis, A., & Jurasek, M. (2008). Context change and travel mode choice: Combining the habit discontinuity and self-activation hypotheses. *Journal of Environmental Psychology*, *28*(2), 121–127. https://doi.org/10.1016/J.JENVP.2007.10.005

Vink, N. M., Klungel, O. H., Stolk, R. P., & Denig, P. (2009). Comparison of various measures for assessing medication refill adherence using prescription data. *Pharmacoepidemiology and Drug Safety*, *18*(2), 159–165. https://doi.org/10.1002/pds.1698

Vrijens, B., Antoniou, S., Burnier, M., de la Sierra, A., & Volpe, M. (2017). Current Situation of Medication Adherence in Hypertension. *Frontiers in Pharmacology*, *8*(MAR), 100. https://doi.org/10.3389/fphar.2017.00100

Vrijens, B., De Geest, S., Hughes, D. A., Przemyslaw, K., Demonceau, J., Ruppar, T., Dobbels, F., Fargher, E., Morrison, V., Lewek, P., Matyjaszczyk, M., Mshelia, C., Clyne, W., Aronson, J. K., & Urquhart, J. (2012a). A new taxonomy for describing and defining adherence to medications. *British Journal of Clinical Pharmacology*, *73*(5), 691–705. https://doi.org/10.1111/j.1365-2125.2012.04167.x

Wagener, J. S., Rasouliyan, L., VanDevanter, D. R., Pasta, D. J., Regelmann, W. E., Morgan, W. J., & Konstan, M. W. (2013). Oral, inhaled, and intravenous antibiotic choice for treating pulmonary exacerbations in cystic fibrosis. *Pediatric Pulmonology*, *48*(7), 666–673. https://doi.org/10.1002/ppul.22652

Wagener, J. S., Williams, M. J., Millar, S. J., Morgan, W. J., Pasta, D. J., & Konstan, M. W. (2018). Pulmonary exacerbations and acute declines in lung function in patients with cystic fibrosis. *Journal of Cystic Fibrosis*, *17*(4), 496–502. https://doi.org/10.1016/j.jcf.2018.02.003

Wark, P. A. B., Tooze, M., Cheese, L., Whitehead, B., Gibson, P. G., Wark, K. F., & McDonald, V. M. (2012). Viral infections trigger exacerbations of cystic fibrosis in adults and children. In *European Respiratory Journal* (Vol. 40, Issue 2, pp. 510–512). https://doi.org/10.1183/09031936.00202311

Webb, T. L., Sheeran, P., & Luszczynska, A. (2009). Planning to break unwanted habits: Habit strength moderates implementation intention effects on behaviour change. *British Journal of Social Psychology*, *48*(3), 507–523. https://doi.org/10.1348/014466608X370591

Weber, C. L., & Zbilut, J. (2005). Recurrence quantification analysis of nonlinear dynamical systems. In M. J. Riley & G. C. Van Orden (Eds.), *Tutorials in Contemporary Nonlinear Methods for the*

Behavioural Sciences (pp. 26–94). http://www.nsf.gov/sbe/bcs/pac/nmbs/nmbs.jsp

Weiner, J. R., Toy, E. L., Sacco, P., & Duh, M. S. (2008). Costs, quality of life and treatment compliance associated with antibiotic therapies in patients with cystic fibrosis: A review of the literature. In *Expert Opinion on Pharmacotherapy* (Vol. 9, Issue 5, pp. 751–766). https://doi.org/10.1517/14656566.9.5.751

Wiecek, E., Tonin, F. S., Torres-Robles, A., Benrimoj, S. I., Fernandez-Llimos, F., & Garcia-Cardenas, V. (2019). Temporal effectiveness of interventions to improve medication adherence: A network meta-analysis. *PLoS ONE*, *14*(3). https://doi.org/10.1371/journal.pone.0213432

Wildman, M. J., & Hoo, Z. H. (2014). Moving cystic fibrosis care from rescue to prevention by embedding adherence measurement in routine care. *Paediatric Respiratory Reviews*, *15*(S1), 16–18. https://doi.org/10.1016/j.prrv.2014.04.007

Wildman, M. J., O'Cathain, A., Hind, D., Maguire, C., Arden, M. A., Hutchings, M., Bradley, J., Walters, S. J., Whelan, P., Ainsworth, J., Tappenden, P., Buchan, I., Elliott, R., Nicholl, J., Elborn, S., Michie, S., Mandefield, L., Sutton, L., Hoo, Z. H., … Sasso, A. (2021). An intervention to support adherence to inhaled medication in adults with cystic fibrosis: the ACtiF research programme including RCT. *Programme Grants for Applied Research*, *9*(11), 1–146. https://doi.org/10.3310/PGFAR09110

Wildman, M. J., O'Cathain, A., Maguire, C., Arden, M. A., Hutchings, M., Bradley, J., Walters, S. J., Whelan, P., Ainsworth, J., Buchan, I., Mandefield, L., Sutton, L., Tappenden, P., Elliott, R. A., Hoo, Z. H., Drabble, S. J., & Beever, D. (2022). Self-management intervention to reduce pulmonary exacerbations by supporting treatment adherence in adults with cystic fibrosis: a randomised controlled trial. *Thorax*, *77*(5), 461–469. https://doi.org/10.1136/THORAXJNL-2021-217594

Wildman, M, Hutchings, M., Arden, M., O'Cathain, A., Drabble, S., Tappenden, P., Beever, D., Whelan, P., Bradley, J., Cantrill, H., Maguire, C., & Hind, D. (2018). *Development and evaluation of an intervention to support adherence to treatment in adults with Cystic Fibrosis : a randomised controlled trial and parallel process evaluation protocol.* https://doi.org/10.15131/SHEF.DATA.7338815

Williams, A. B., Amico, K. R., Bova, C., & Womack, J. A. (2013). A proposal for quality standards for measuring medication adherence in research. In *AIDS and Behavior* (Vol. 17, Issue 1, pp. 284–297). https://doi.org/10.1007/s10461-012-0172-7

Wohn, D. Y., & Ahmadi, M. (2019). Motivations and habits of micro-news consumption on mobile social media. *Telematics and Informatics*, *44*, 101262. https://doi.org/10.1016/J.TELE.2019.101262

Wood, W. (2017). Habit in personality and social psychology. *Personality and Social Psychology Review*, *21*(4), 389–403. https://journals.sagepub.com/doi/pdf/10.1177/1088868317720362?casa_token=Ufu3Dj5q_4I AAAAA:7QBtzYxbz_OLgd4gqWnKRoHOuHKWZIcJD4ZKIcScxOSaB7sNxL3Ks2nL1Lql1x5Nxb6Nmt Sgmlo

Wood, Wendy, & Neal, D. (2007). A new look at habits and the habit-goal interface. *Psychological Review*, *114*(4), 843–863. https://doi.org/10.1037/0033-295X.114.4.843

Wood, Wendy, & Neal, D. T. (2009). The habitual consumer. *Journal of Consumer Psychology*, *19*(4), 579–592. https://doi.org/10.1016/J.JCPS.2009.08.003

Wood, Wendy, & Rünger, D. (2016). Psychology of habit. *Annual Review of Psychology*, *67*, 289–314.

https://doi.org/10.1146/annurev-psych-122414-033417

Wood, Wendy, Tam, L., & Witt, M. G. (2005). Changing circumstances, disrupting habits. *Journal of Personality and Social Psychology*, *88*(6), 918–933. https://doi.org/10.1037/0022-3514.88.6.918

World Health Organisation. (2001). *Adherence to long term therapies: policy for action*. https://apps.who.int/iris/handle/10665/66984

World Health Organisation. (2003). *Adherence to long-term therapies: Evidence for action*. https://www.who.int/chp/knowledge/publications/adherence_report/%0Aen/

Yeaw, J., Benner, J. S., Walt, J. G., Sian, S., & Smith, D. B. (2009). Comparing Adherence and Persistence Across 6 Chronic Medication Classes. *Journal of Managed Care Pharmacy*, *15*(9), 728–740. https://doi.org/10.18553/jmcp.2009.15.9.728

Zullig, L. L., Mendys, P., & Bosworth, H. B. (2017). Medication adherence: A practical measurement selection guide using case studies. In *Patient Education and Counseling* (Vol. 100, Issue 7, pp. 1410–1414). Elsevier Ireland Ltd. https://doi.org/10.1016/j.pec.2017.02.001

Appendices

Appendix A. Chapter 4 machine-learning assisted review search strategy

MEDLINE Search strategy (searched 13th May 2020)- 2953 hits

1. (MeSH) exp Medication Adherence/

2. (MeSH) exp Patient Compliance/

3. ((medication or treatment) adj3 (adheren* or complian*)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

4. 1 or 2 or 3

5. randomi?ed.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

6. placebo.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

7. (MeSH) exp Randomized Controlled Trial/

8. 5 or 6 or 7

9. (habit* or routine* or automatic*).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

10. (behavio?r* adj3 (intervention or treatment)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

11. ('behavio?r-change' or 'behavio?r change').mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

12. ((complex or multicomponent or multi-component or multifunction* or multilevel) adj (intervention or treatment or programme)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

13. ((adheren* or complian*) adj3 (intervention or treatment or support or improve* or promote or foster)).mp. [mp=title, abstract, original title, name of substance word, subject heading word,

floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

14. 9 or 10 or 11 or 12 or 13

15. 4 and 8 and 14

16. limit 15 to yr="2015 -Current"

Appendix B. Chapter 4 machine learning-assisted review: overview of the implementation of an Active Learning technique during title and abstract screening

An active learning approach (O'Mara-Eves et al., 2015) was used to assist with the title and abstract screening phase. In the present study active learning describes an iterative process, in which a machine classifier is trained on incrementally larger datasets of labelled studies, and applies the resulting classifier to the remaining unlabelled studies with an output of studies ranked by relevance for full-text screening inclusion. To begin, the first 10% of studies identified from the database searches were manually labelled 'include' or 'exclude' against the pre-specified criteria. This set of studies was used to train a machine-learning classifier based on a standard logistic regression model. First, a 'bag of words' approach was adopted by concatenating titles and abstracts for each study. Stop words, i.e. common English words that carry little meaning for classification (e.g. 'a', 'the'), were removed as were any words occurring in fewer than 10% of the studies in the training set. TF-IDF scores of the remaining terms were then calculated (TF-IDF scores are a numerical value intended to reflect a word's relative importance to the study description in relation to how frequently it appears within the entire corpus of words across all studies; words that appear frequently in every document have less discriminant power and so are assigned lower values than words that appear multiple times in one document but infrequently in other documents, and therefore have greater discriminant power). The resulting feature vectors were then used to train a logistic regression model, using k-fold cross-validation with 5 folds, to predict likelihood of inclusion for full-text screening. Initial model specification was optimised by maximising overall model performance (accuracy) and identification of relevant documents (relevant document recall). Model optimisation was evaluated on the performance of a logistic regression model which gave highest overall model accuracy and recall for articles for inclusion. Following model optimisation on the 10% labelled dataset, the next 5% of studies were labelled (include/exclude), ordered from highest to lowest on predicted likelihood of inclusion. This process was iterated, retraining the machine learning classifier after each 5% increment of labelling using the reranked study list for each iteration, until identification of additional studies for inclusion was assessed to be unlikely. All remaining, unlabelled studies were excluded at this stage.

Class imbalance was handled using synthetic minority oversampling technique (SMOTE; Chawla et al., 2002). This technique randomly synthesises minority class data points (here represented by studies assigned to the 'include' category) along axes between existing data points.

The whole implementation was carried out in Python using Scikit-learn.

The screening process and number of texts included after each iteration are summarised in Table B1.

Table B1

Summary of active learning screening process

Total labelled n(%)	Labelled include n	Labelled exclude n	Remaining unlabelled n	Additional studies identified for inclusion for full text screening n	Studies which were included after full-text screening n	Model Accuracy Mean(SD)[a]	Model Recall[b] Mean (SD)[a]
597 (10%)	79	518	5376	n/a	0	0.83(0.02)	0.59(0.07)
896 (15%)	262	634	5077	183	7	0.79(0.25)	0.84(0.06)
1195 (20%)	424	771	4778	162	4	0.75(0.03)	0.86(0.05)
1494 (25%)	554	940	4479	130	1	0.72(0.03)	0.79(0.03)
1793 (30%)	647	1146	4180	93	2	0.69(0.02)	0.72(0.10)
2092 (35%)	715	1377	3881	68	1	0.71(0.05)	0.69(0.05)
2391 (40%)	771	1620	3582	56	0	0.71(0.02)	0.68(0.06)
2690 (45%)	803	1886	3284	32	1	0.72(0.03)	0.66(0.01)
2989 (50%)	818	2171	2984	15	0	0.74(0.02)	0.68(0.03)
3288 (55%)[t]	**826**	**2462**	**2685**	**8**	**1**	**0.76(0.02)**	**0.71(0.04)**

[a]Mean(SD) calculated for model summaries across 5 folds

[b]Recall describes the proportion of studies identified for inclusion that were labelled include in the training dataset.

[t]Title and abstract screening was ceased after 55% of texts were labelled.

Screening ceased after manually labelling 55% of the titles and abstracts. During the final 5% of studies screened, only an additional 8/299 studies were included for full-text screening. Furthermore, only 17 of the 826 studies included during title and abstract screening were included in the final review (precision ~2.5%) and 15/17 of those were found within 35% of the documents screened. The study found in the last screening batch was 10th out of 299 studies screened in this batch. This analysis indicates that the rate at which documents relevant for inclusion were being identified was dropping rapidly and the reviewers believed there to be adequate justification to cease title and abstract screening at this stage.

Appendix C. Chapter 4 machine learning-assisted review: BCTs coded by intervention and frequencies

The following presents the behaviour change techniques (BCTs) coded for each intervention arm of each study included in the review presented in Chapter 4 (Table C1), followed by the frequency of each BCT coded (Table C2).

Table C1

BCTs coded by intervention arm

Study and intervention name	BCTs coded§
Barankay et al (Simple sweepstake)	2.2 Feedback on behaviour
	5.5 Anticipated regret
	7.1 Prompts/cues
	10.1 Material incentive (behaviour)
	10.2 Material reward (behaviour)*
Barankay et al	2.2 Feedback on behaviour
	5.5 Anticipated regret
(Deadline sweepstake / "habit formation")	7.1 Prompts/ cues
	10.1 Material incentive (behaviour)
	10.2 Material reward (behaviour)*
	14.6 Situation specific reward*
Barankay et al	2.2 Feedback on behaviour;
	7.1 Prompts/cues;
(Sweepstake plus deposit contract/"hybrid")	10.1 Material incentive (behaviour);
	10.2 Material reward (behaviour)
	14.1 Behaviour cost
De Bruin et al	1.1 Goal setting (behaviour)
	1.2 Problem solving;
	1.4 Action planning;
	1.5 Review behaviour goals;
	1.6 discrepancy between current behaviour and goal;
	2.2 Feedback on behaviour;
	2.3 Self monitoring of behaviour;
	5.3 Information about social and environmental consequences;
	10.4 Social reward;
	12.5 Adding objects to the environment
Farmer et al	1.4 Action planning;
	7.1 Prompts/cues
Gregoriano et al	1.2 Problem solving;
	2.2 Feedback on behaviour;
	7.1 Prompts/cues;
	8.1 Behavioural practice/rehearsal;
	8.3 Habit formation
Haynes et al	1.2 Problem solving;
	1.3 Goal setting (outcome);
	1.4 Action planning;
	1.6 Discrepancy between current behaviour and goal;
	2.2 Feedback on behaviour;
	2.3 Self-monitoring of behaviour;
	2.4 Self-monitoring of outcome of behaviour;
	2.6 Biofeedback;
	3.1 Social support (unspecified);
	7.1 Prompts/cues;
	8.1 Behavioural practice/ rehearsal;
	8.3 Habit formation;
	10.4 Social reward;
	10.10 Reward (outcome);

Lin et al	12.1 Restructuring the physical environment
	1.2 Problem solving;
	1.4 Action planning;
	2.2 Feedback on behaviour;
	3.1 Social support (unspecified);
	3.3 Social support (emotional);
	5.1 Information about health consequences;
	7.1 Prompts/cues;
	9.2 Pros and cons;
	9.3 Comparative imagining of future outcomes;
	15.2 Mental rehearsal of successful performance
	15.4 Self talk
Milam et al	1.2 Problem solving
	1.4 Action planning
	7.1 Prompts/cues
	8.1 Behavioural practice/ rehearsal
	8.3 Habit formation
O'Dwyer et al	2.2 Feedback on behaviour
Biofeedback intervention	4.1 Instruction on how to perform a behaviour
	8.1 Behavioural practice/rehearsal
O'Dwyer et al	4.1 Instruction on how to perform a behaviour
Inhaler technique education only	6.1 Demonstration of the behaviour
	8.1 Behavioural practice/rehearsal
Pakpour et al	**1.2 Problem solving**
	1.4 Action planning
	2.3 Self-monitoring of behaviour
	5.1 Information about health consequences
	9.1 Credible source
	13.4 Valued self-identity
	15.4 Self-talk
Reddy et al	2.2. Feedback on behaviour;
Individual feedback	7.1 Prompts/cues;
	10.4 Social reward;
	12.5 Adding objects to the environment
Reddy et al	2.1 Monitoring of behaviour by others without feedback
Partner feedback	2.2. Feedback on behaviour
	7.1 Prompts/cues
	10.4 Social reward
	12.5 Adding objects to the environment
Russell et al	1.1 Goal setting (behaviour)
	1.4 Action planning
	1.5 Review behaviour goals
	1.6 Discrepancy between current behaviour and goal
	2.2 Feedback on behaviour
	4.4 Behavioural experiments
	8.1 Behavioural practice/rehearsal
	8.3 Habit formation
	15.3 Focus on past success
Stacy et al	1.2 Problem solving
	3.1 Social support (unspecified)
	9.2 Pros/cons
	5.3 Information about health consequences
Tang et al	1.4 Action planning
	7.1 Prompts/cues
	8.1 Behavioural practice/rehearsal
	8.3 Habit formation
	9.1 Credible source
Tuldra et al	1.2 Problem solving
	5.1 Information about health consequences
	9.1 Credible source
Wildman et al	**1.1 Goal setting (behaviour)**
	1.2 Problem solving
	1.4 Action planning

1.5 Review behavioural goals
1.6 Discrepancy between current behaviour and goal
2.2 Feedback on behaviour
2.3 Self-monitoring of behaviour
3.2 social support (practical)
4.1 Instruction on how to perform the behaviour
5.1 Information about health consequences
5.2 Salience of consequences
6.1 Demonstration of the behaviour
7.1 Prompts/cues
8.1 Behavioural practice/rehearsal
8.3 Habit formation
8.7 Graded tasks
9.1 Credible source
10.4 Social reward
12.1 Restructure the physical environment
12.5 Adding objects to the environment
15.3 Focus on past success
15.4 Self-talk
16.3 Vicarious consequences

§Intervention techniques in bold were self-coded by authors using the Michie and colleagues (2013) BCT taxonomy
*Participants were entered into a sweepstake so rewards only received by some people, some of the time

Table C2

Frequencies of coded BCTs

BCT	Frequency (%) coded
1.1 Goal setting (behaviour)	3 (17%)
1.2 Problem solving	9(50%)
1.3 Goal setting (outcome)	1 (6%)
1.4 Action planning	9(50%)
1.5 Review behavioural goals	3(17%)
1.6 Discrepancy between current behaviour and goal	4(22%)
2.1 Monitoring of behaviour by others without feedback	1(6%)
2.2 Feedback on behaviour	12(67%)
2.3 Self-monitoring of behaviour	4(22%)
2.4 Self-monitoring of outcome of behaviour	1(6%)
2.6 Biofeedback	1(6%)
3.1 Social support (unspecified);	2(11%)
3.2 Social support (practical)	1(6%)
3.3 Social support (emotional);	1(6%)
4.1 Instruction on how to perform the behaviour	3(17%)
4.4 Behavioural experiments	1(6%)
5.1 Information about health consequences	4(22%)
5.2 Salience of consequences	1(6%)
5.3 Information about social and environmental consequences	2(11%)
5.5 Anticipated regret	2(11%)
6.1 Demonstration of the behaviour	2(11%)
7.1 Prompts/cues	11(61%)
8.1 Behavioural practice/rehearsal	8(44%)
8.3 Habit formation	6(33%)
8.7 Graded tasks	1(6%)

9.1 Credible source	3(17%)
9.2 Pros and cons	1(6%)
9.3 Comparative imagining of future outcomes	1(6%)
10.1 Material incentive (behaviour)	3(17%)
10.2 Material reward (behaviour)	3(17%)
10.10 Reward (behaviour)	1(6%)
12.1 Restructuring the physical environment	2(11%)
12.5 Adding objects to the environment	4(22%)
14.1 Behaviour cost	1(6%)
14.6 Situation specific reward	1(6%)
15.2 Mental rehearsal of successful performance	1(6%)
15.3 Focus on past success	2(11%)
15.4 Self-talk	3(17%)
16.3 Vicarious consequences	1(6%)

Appendix D. Chapter 4 machine learning-assisted review: Effectiveness of all interventions.

Intervention effects on adherence for all included studies

The primary adherence outcome measures, summaries and results of interest for the 18 intervention comparisons are reported in Table D1.

Adherence data were derived from a variety of sources. Twelve studies (67% of intervention comparisons) utilised data from electronic devices for pill dispensing (e.g Medication Event Monitoring System, MEMS, or variant of this device) or electronic devices for inhaled medications (e.g. PARI eTrack™, INCA®). Two intervention comparisons (11%) used pill counts or pharmacy refill data, two used self-report medication adherence scales (MARS, MMAS-4; Morisky, Green & Levine, 1986) and two asked participants to self-report the number of doses taken over a specified time period.

Primary adherence outcomes were extracted for a minimum follow-up period of six months; 17 of 18 interventions were favoured in these comparisons (94%, 95%CI=74 to 99%; p<.001). Four of five studies reporting additional results at follow-up points >6months (range 32 weeks to 18 months) favoured the intervention (80%; 95%CI= 38 to 96%, p=0.375). Of all summaries that were taken at >6 months, 7 of 8 interventions were favoured (88%, 95%CI= 53 to 98%; p=0.07).

These aggregated findings must be interpreted with caution as even within studies with similar data collection methods, data were summarised into various different outcome summary measures. Among studies which compared the effects of an intervention immediately after its termination, as well as the persistence of those effects after a period of intervention, withdrawal showed a trend for reduced effect sizes, although not all reduced to non-significant effects (Barankay et al., 2020). Among all 12 intervention comparisons using data captured on electronic devices, effect sizes ranged from d=-0.06 to 1.71, indicating disparate findings between these studies using the most objective measures of adherence, nearest to 6 months. This complexity precluded meta-analysis.

Table D1

Adherence outcomes by intervention

First author (Year)	Intervention and comparator description	Intervention duration to assessment timepoint details	Data collection	Adherence summary and timepoint measured	Adjustments to estimates	n (I,C) analysed	Summary of findings (nearest to 6 months)	Direction of effect (nearest to 6 months) and standardised effect size*	>6-month follow-up adherence findings
Barankay (2020)	Simple sweepstakes vs control	Intervention delivered over 6 month period	Electronic pill bottle	Proportion of 180 days post-intervention in which the pill bottle registered an opening	n/a	I=199, C=201	Significant difference: Unadjusted mean difference= 0.15 (0.11 to 0.19), p<.001	Favours intervention **Unadjusted d (95%CI)= 0.68(0.48, 0.88)**	
Barankay (2020)	Deadline sweepstakes vs control	Intervention delivered over 6 month period	Electronic pill bottle	Proportion of 180 days post-intervention in which the pill bottle registered an opening	n/a	I=204, C=201	Significant difference: Unadjusted mean difference= 0.17 (0.13 to 0.21), p<.001	Favours intervention **Unadjusted d (95%CI)= 0.77(0.57, 0.96)**	
Barankay (2020)	Sweepstakes plus deposit contract vs control	Intervention delivered over 6 month period	Electronic pill bottle	Proportion of 180 days post-intervention in which the pill bottle registered an opening	n/a	I=201, C=201	Significant difference: Unadjusted mean difference= 0.19 (0.15 to 0.23) p<.001	Favours intervention **Unadjusted d (95%CI)= 0.86 (0.66, 1.07)**	
De Bruin (2010)	Electronic Monitoring-Based Counselling vs usual care	Intervention delivered months 1 to 5-6	MEMS-cap	Timing adherence (proportion of doses taken within specified time interval	Baseline timing adherence	I=66 C=67	Significant difference: Adjusted mean difference= 7.4, [95%CI 3.50-11.30%]. F(1, 129) = 14.11, p<.001	Favours intervention **Adjusted d (95%CI)= 0.65 (0.30, 0.99)**	(Primary outcome measured at 8-9 months)

each day) over
8-9 months

Study	Intervention	Delivery	Measure device	Outcome measure	Adjustment	N	Results	Effect size	Follow-up
Farmer (2016)	Brief action planning intervention vs usual care	Intervention sessions delivered up to 32 weeks	eMEMS device	Proportion of days within the 18-32 week period for which all prescribed medications were taken	Adjusted for clustering only	I=263 C=331	No significant difference: intervention mean(95%CI)= 83.4% (80.8 to 86.0%) vs control mean(95%CI)= 81.1% (78.8 to 83.4%), adjusted mean difference= 2.3 (95% CI, -1.2 to 5.8), p=.19	Favours intervention **Adjusted d (95%CI)= 0.11 (-0.06, 0.27)**	No significant difference between 32-52 weeks: adjusted mean difference= -2.3 [95%CI -7.8 to 3.0]. **Adjusted d (95%CI)= -0.07 (-.24, 0.01)**Does not favour intervention.
Gregoriano (2019)	Daily alarm clock and support phone calls vs control	Intervention delivered over 6 months	Electronic device or POEMs where appropriate for dry powder medications	Mean number of days (max 200) that taking-adherence was in target adherence range (80-100%)	n/a	I (Puff inhaler)= 57, I (Dry powders) = 41 C(Puff inhaler)= 60, C (Dry powders) = 49	Significant difference for puff inhalers and dry powder capsules: Puff inhalers, Intervention mean(SD)= 81.6(14.2), control mean(SD) =60.1(30.3), p<.001; Dry powders Intervention mean(SD)= 89.6(9.8), Control mean(SD) =80.2(21.3), p=.01	Favours intervention Puff inhalers: **Unadjusted d (95%CI)= 0.90 (0.52, 1.28)** **Dry powders: Unadjusted d (95%CI)= 0.55 (0.13, 0.97)**	
Haynes et (1976)	Behavioural intervention vs usual care	Intervention delivered over full 6 months	Unused pill count	Change in adherence at 6-months, adherence calculated as proportion of pills prescribed that are taken	n/a	I=20 C= 18	Significant difference: Intervention mean (SE)=65.8(8.2) vs control mean(SE) = 43.2 (10.1), Baseline adjusted difference = 22.8, p=.025	Favours intervention **Unadjusted Hedges g (95%CI)= 0.56 (-0.09, 1.21)**	

in the month of follow-up

Study	Intervention	Delivery	Measure	Outcome	Covariates	N	6-month results	Effect size	12- and 18-month results
Lin (2017)	Multifaceted intervention including motivational interviewing vs usual care	Intervention delivered over 6 months	Medication Adherence Rating Scale (MARS)	Self-report adherence at 6-months	Age, sex, Charlson comorbidity index, and body mass index	I=144 C=144	Significant difference: Intervention mean(SD)=13.67 (2.8)vs control mean=7.69(2.7). After adjustment B= 3.97 (SE= 0.22, p<.01)	Favours intervention **Adjusted d (95%CI)= 2.13 (1.84, 2.42)**	Significant differences observed at 12 months. Intervention mean=13.61(2.82) control mean(SD)= 7.71(2.79) (after adjustment B = 3.83, SE=0.23, p<.01). Favours intervention. **Adjusted d (95%CI)= 1.96 (1.68, 2.24)** Significant difference at 18 months. Intervention mean(SD)= 13.70(2.75), control mean(SD)= 7.63(2.88) (B=4.24, p<.01)** Favours intervention. **Adjusted d (95%CI)= 1.92 (1.64, 2.20)**

Milam (2005)	Brief adherence intervention vs safe sex intervention (no adherence component)	Intervention delivered until month 10-11	Self-report of number of pills taken over past 7 days	Percent of patients whose 7-day adherence >95 % measured at 17-18 months	Income, ethnicity, employment status, AIDS diagnosis, HAART regimen (vs. non-HAART), and number of pills per day, adherence at baseline and clustering	I=149 C=288	No significant effect: Intervention n(%)=128 (85.9%) vs control n(%)=201(69.8%). Adjusted OR = 2.05, [95% CI: 0.92 to 4.56], p=.077	Favours intervention **(Unadjusted) RR (95% CI)=1.23 (1.11, 1.36)**	(Primary outcome measured at 17-18 months)
O'Dwyer (2020)	Biofeedback vs usual care	Intervention delivered in months 1-2	INCA® electronic device	Actual adherence calculated as percentage of correctly taken doses of those expected from the prescription, during month 6	Clustering by pharmacy, baseline differences in adherence, age, and sex.	I=71 C=22	Significant difference: (Per protocol: Intervention mean(SD)= 60.8(30.4), Control mean(SD)= 33.2(30.7)). Adjusted mean difference= 24.61(SE=7.07)(95%CI 10.07 to 38.89), p=.002	Favours intervention **Adjusted d (95%CI)= 0.69(0.21, 1.18)**	
O'Dwyer (2020)	Demonstration vs usual care	Intervention delivered in months 1-2	INCA® electronic device	Actual adherence calculated as percentage of correctly taken doses of those expected from the prescription, during month 6	Clustering by pharmacy, baseline differences in adherence, age, and sex.	I=54 C=22	No significant difference: (Per protocol: Intervention mean(SD)= 44.2 (30.7), Control mean(SD)= 33.2(30.7)). Adjusted mean difference= 16.59(SE=8.42) (95%CI, -0.53 to 34.51), p=.057	Favours intervention **Adjusted d (95%CI)= 0.43(-0.07, 0.93)**	

Study	Intervention	Delivery	Measure	Outcome	Covariates	N	Results	Effect	12-month
Pakpour (2015)	Multimodal behavioural intervention vs usual care	Intervention completed at 3 weeks	Medication Adherence Report Scale (MARS)	Self-report adherence at 6-months	Age, gender, number of medications, and time since seizure.	I=137 C=134	Significant difference: Intervention 18.61(2.86) vs Control mean(SD)= 15.98(3.65), adjusted B(SE)= 1.73(0.47), p<.001	Favours intervention **Unadjusted d (95%CI)= 0.45 (0.21, 0.69)**	
Reddy et (2017)	Individual feedback vs usual care	Intervention delivered weeks 1-13	GlowCap electronic monitoring device	Adherence calculated for weeks 13-26	n/a	I= 36 C= 36	No significant difference: Individual feedback= 60% (95% CI 0.49 to 0.72) vs control =54% (95%CI= 0.43 to 0.66), p = 0.75	Favours intervention **Unadjusted d (95%CI)= 0.18(-0.29, 0.64)**	
Reddy (2017)	Individual and partner feedback vs usual care	Intervention delivered weeks 1-13	GlowCap electronic monitoring device	Adherence calculated for weeks 13-26	n/a	I=54 C=36	No significant difference: Partner feedback =52% (95%CI = 0.42 to 0.61) vs Control= 54% (95%CI= 0.43 to 0.66), p = 0.95	Does not favour intervention **Unadjusted d (95%CI)= -0.06(-0.48, 0.36)**	
Russell (2020)	SystemCHANGE intervention vs attention control	Intervention delivered over 6 months	MEMS-cap	Adherence rate defined as doses taken on time/total doses at 6 months	Ethnicity, marital status, perceived health score, and perceived social support	I=45 C= 44	Significant difference: Intervention mean(SD)=0.81(0.25) vs control mean(SD)= 0.64(0.24), p<.001. Adjusted difference B= 0.2 (95% CI= 0.12 to 0.27; SE = 0.039, p<.001)	Favours intervention **Unadjusted d (95%CI)= 0.69(0.27, 1.12)**	Significant difference at 12 months: Intervention mean(SD)=0.65(0.37) vs control mean(SD)= 0.53(0.29), p=.004. Adjusted B=0.16 (95% CI= 0.06 to 0.26); SE = 0.05) p<.001. Favours intervention. **Unadjusted d (95%CI)= 0.36 (-0.06, 0.78)**
Stacy (2009)	Interactive voice response technology intervention vs enhanced care control	Intervention delivered over 6 month period	Pharmacy claims database to determine if prescriptions filled	6-month point prevalence persistency defined as a subject being in possession of a	Number of chronic medications in 3-month period prior to index statin	I=253 C=244	Significant effect: Intervention % persistent =70.4% vs control % persistent=60.7%, adjusted OR= 1.64 (90% CI 1.19– 2.27), p<.05	Favours intervention **RR(95% CI)= 1.16 (1.02, 1.32)**	

				statin at the end of the 180-day observation period	and intention to continue taking statin once cholesterol level drops to a healthy level				
Tang (2014)	Education plus behavioural intervention vs education only	Intervention delivered over full 6 months	Morisky Medication Adherence Scale (MMAS-4)	Number of individuals whose self-report adherence improved at 6-months.	n/a	I=56 C=53	No significant effect: Intervention n improved (%)= 36(64.3%) vs Control n improved (%)=33 (62.3%), p=0.827	Favours intervention **RR(95% CI)= 1.03 (0.78, 1.37)**	
Tuldrà (2000)	Psychoeducative intervention vs usual care	Unclear on duration of intervention	Self-report of the number of pills taken in the last month	Percentage of patients achieving adherence >=95% at week 24 (summary for week is number of pills over the past 4 weeks)	n/a	I=55 C=61	No significant effect: Intervention n(%)=28(51%) vs control n(%)= 24(39%) p=NS	Favours intervention **RR(95% CI)= 1.29 (0.86, 1.94)**	No significant effect at 48 weeks: Intervention (n)% = 32 (58) vs control n(%)= 25 (41%), p=.064. Favours intervention. **RR (95% CI)= 1.42 (0.98, 2.07)**
Wildman (2021)	CFHealthHub intervention vs usual care	Intervention delivered over 12-month period	PARI eTrack™ electronic device	Normative adherence (adjustments made for ideal treatment for effectiveness) at 12 months	Treatment arm, time in weeks, baseline adherence (first two weeks), and past-year IV days	I=295 C=293	Significant effect: intervention mean(SD)= 52.9%(31.4) vs control mean(SD)= 34.9%(31.7), adjusted mean difference= 9.5% (95%CI, 8.6-10.4)	Favours intervention **Adjusted d (95%CI)= 1.71 (1.52, 1.90)**	(Primary outcome measured at 12 months)

I= Intervention; C= Control; MEMS= Medication event monitoring system; POEMs= Polymedication Electronic Monitoring System; HAART= Highly active anti-retroviral therapy; SE= standard error; SD= standard deviation; CI= confidence interval

*The effect size was estimated for each RCT based on the raw observed data; if this was not available on the adjusted difference.

**The effect estimate for this follow-up point was not included in binomial probability analysis of findings >6 months due to non-independence from 12-month estimate for the same intervention

Appendix E. Complete list of the 127 variables derived in Chapter 7

range_time	LAM_rad_35	VENT_rad_25
iqr_time	LAM_rad_40	VENT_rad_30
std_time	LAM_rad_45	VENT_rad_35
time_entropy	LAM_rad_50	VENT_rad_40
diffs_entropy	LAM_rad_55	VENT_rad_45
average_diff	LAM_rad_60	VENT_rad_50
std_diff	LAM_rad_70	VENT_rad_55
RR_rad_5	LAM_rad_80	VENT_rad_60
RR_rad_10	LAM_rad_90	VENT_rad_70
RR_rad_15	TT_rad_5	VENT_rad_80
RR_rad_20	TT_rad_10	VENT_rad_90
RR_rad_25	TT_rad_15	W_rad_5
RR_rad_30	TT_rad_20	W_rad_10
RR_rad_35	TT_rad_25	W_rad_15
RR_rad_40	TT_rad_30	W_rad_20
RR_rad_45	TT_rad_35	W_rad_25
RR_rad_50	TT_rad_40	W_rad_30
RR_rad_55	TT_rad_45	W_rad_35
RR_rad_60	TT_rad_50	W_rad_40
RR_rad_70	TT_rad_55	W_rad_45
RR_rad_80	TT_rad_60	W_rad_50
RR_rad_90	TT_rad_70	W_rad_55
DET_rad_5	TT_rad_80	W_rad_60
DET_rad_10	TT_rad_90	W_rad_70
DET_rad_15	VMAX_rad_5	W_rad_80
DET_rad_20	VMAX_rad_10	W_rad_90
DET_rad_25	VMAX_rad_15	WENT_rad_5
DET_rad_30	VMAX_rad_20	WENT_rad_10
DET_rad_35	VMAX_rad_25	WENT_rad_15
DET_rad_40	VMAX_rad_30	WENT_rad_20
DET_rad_45	VMAX_rad_35	WENT_rad_25
DET_rad_50	VMAX_rad_40	WENT_rad_30
DET_rad_55	VMAX_rad_45	WENT_rad_35
DET_rad_60	VMAX_rad_50	WENT_rad_40
DET_rad_70	VMAX_rad_55	WENT_rad_45
DET_rad_80	VMAX_rad_60	WENT_rad_50
DET_rad_90	VMAX_rad_70	WENT_rad_55
LAM_rad_5	VMAX_rad_80	WENT_rad_60
LAM_rad_10	VMAX_rad_90	WENT_rad_70
LAM_rad_15	VENT_rad_5	WENT_rad_80
LAM_rad_20	VENT_rad_10	WENT_rad_90
LAM_rad_25	VENT_rad_15	
LAM_rad_30	VENT_rad_20	

Appendix F. Comparisons between full-trial and analysis-subsample characteristics

Appendix F contains a series of tables that compare the full-trial sample characteristics with characteristics of each sub-samples used in each analysis presented throughout the thesis.

Table F1

Comparison of analysis sample baseline characteristics to full trial sample Chapter 7 and Chapter 8 Part 1, interval 7 data

Baseline Characteristic	Full trial sample Mean (SD)	Median [LQ, UQ]	n(%)	Analysis sample Mean (SD)	Median [LQ, UQ]	n(%)
Intervention arm			305 (50.2%)			273 (50.9%)
Age	30.7 (10.7)	28.1 [22.6, 36.1]		31.3 (11.0)	28.8 [23.1,36.6]	
Female			310 (51.1%)			273 (50.9%)
IV days (home)	13.2 (21.2)	0 [0, 19]		13.5 (21.4)	0 [0,20.8]	
IV days (hospital)	12.7 (21.6)	0 [0,15.5]		12.6 (22.2)	0 [0,15]	
FEV1%	59.5 (23.1)	59.1 [40.0, 77.9]		58.7 (22.7)	57.0 [39.5, 76.6]	
IMD quintile	3.04 (1.36)	3 [2,4]		3.03 (1.36)	3 [2,4]	
Number of prescribed doses	3.26 (1.60)	3 [2, 4.29]		3.39 (1.61)	3.00 [2.32, 4.46]	
Max days in trial	463.8 (123.2)	486 [430.5, 538]		475.5 (109.0)	489 [438, 542.8]	
Numerator adjusted adherence during trial	48.3 (32.8)	48.0 [16.4, 79.0]		54.8 (30.0)	57.4 [28.3, 81.6]	
CFHH Pseudomonas status			349 (57.9%)			309 (57.6%)
Necessity	3.56 (0.74)	3.6 [3.10, 4.00]		3.63 (0.72)	3.7 [3.1, 4.0]	
Concerns	2.08 (0.54)	2.1 [1.7,2.4]		2.07 (0.55)	2.1 [1.6, 2.4]	

IV= intravenous antibiotics; FEV1%= forced expiratory volume in first second (per cent); IMD quintile= Index of Multiple Deprivation quintiles; CFHH Pseudomonas status: chronic=diagnostic criteria fulfilled for chronic pseudomonas status using the CFHealthHub criteria tool (Hoo, Coates, et al., 2018).

Table F2

Comparison of analysis sample baseline characteristics to full trial sample Chapter 7 Part 1, interval 10 data

Baseline Characteristic	Full trial sample			Analysis sample		
	Mean (SD)	Median [LQ, UQ]	n(%)	Mean (SD)	Median [LQ, UQ]	n(%)
Intervention arm			305 (50.2%)			282 (52%)
Age	30.7 (10.7)	28.1 [22.6, 36.1]		31.1 (10.9)	28.8 [23.0, 35.4]	
Female			310 (51.1%)			280 (51.7%)
IV days (home)	13.2 (21.2)	0 [0, 19]		13.2 (21.2)	0 [0, 19]	
IV days (hospital)	12.7 (21.6)	0 [0,15.5]		12.6 (22.0)	0 [0, 0, 15]	
FEV1%	59.5 (23.1)	59.1 [40.0, 77.9]		58.8 (22.7)	57.7 [39.5, 76.6]	
IMD quintile	3.04 (1.36)	3 [2,4]		3.04 (1.36)	3 [2,4]	
Number of prescribed doses	3.26 (1.60)	3 [2, 4.29]		3.36 (1.60)	3 [2.24, 4.44]	
Max days in trial	463.8 (123.2)	486 [430.5, 538]		476.2 (108.2)	489 [438, 542]	
Numerator adjusted adherence during trial	48.3 (32.8)	48.0 [16.4, 79.0]		53.6 (30.4)	55.0 [26.2, 81.0]	
CFHH Pseudomonas status			349 (57.9%)			316 (58.6%)
Necessity	3.56 (0.74)	3.6 [3.10, 4.00]		3.61 (0.73)	3.7 [3.1, 4]	
Concerns	2.08 (0.54)	2.1 [1.7,2.4]		2.07 (0.55)	2.1 [1.6, 2.4]	

IV= intravenous antibiotics; FEV1%= forced expiratory volume in first second (per cent); IMD quintile= Index of Multiple Deprivation quintiles; CFHH Pseudomonas status: chronic=diagnostic criteria fulfilled for chronic pseudomonas status using the CFHealthHub criteria tool (Hoo, Coates, et al., 2018).

Table F3

Comparison of analysis sample baseline characteristics to full trial sample Chapter 7 Part 1, interval 12 data

Baseline Characteristic	Full trial sample			Analysis sample		
	Mean (SD)	Median [LQ, UQ]	n(%)	Mean (SD)	Median [LQ, UQ]	n(%)
Intervention arm			305 (50.2%)			285 (52.2%)
Age	30.7 (10.7)	28.1 [22.6, 36.1]		31.1 (10.9)	28.8 [23, 36.3]	
Female			310 (51.1%)			284 (52.0%)
IV days (home)	13.2 (21.2)	0 [0, 19]		13.2 (21.2)	0 [0,19.8]	
IV days (hospital)	12.7 (21.6)	0 [0,15.5]		12.6 (21.9)	0 [0,15]	
FEV1%	59.5 (23.1)	59.1 [40.0, 77.9]		58.9 (22.7)	57.7 [39.5, 76.8]	
IMD quintile	3.04 (1.36)	3 [2,4]		3.03 (1.36)	3 [2,4]	
Number of prescribed doses	3.26 (1.60)	3 [2, 4.29]		3.35 (1.60)	3 [2.22, 4.41]	
Max days in trial	463.8 (123.2)	486 [430.5, 538]		476.6 (107.9)	489 [438, 542]	
Numerator adjusted adherence during trial	48.3 (32.8)	48.0 [16.4, 79.0]		53.3 (30.5)	54.0 [25.2, 80.9]	
CFHH Pseudomonas status			349 (57.9%)			320 (58.9%)
Necessity	3.56 (0.74)	3.6 [3.10, 4.00]		3.60 (0.74)	3.7 [3.10, 4.0]	
Concerns	2.08 (0.54)	2.1 [1.7,2.4]		2.07 (0.55)	2.1 [1.60, 2.40]	

IV= intravenous antibiotics; FEV1%= forced expiratory volume in first second (per cent); IMD quintile= Index of Multiple Deprivation quintiles; CFHH Pseudomonas status: chronic=diagnostic criteria fulfilled for chronic pseudomonas status using the CFHealthHub criteria tool (Hoo, Coates, et al., 2018).

Table F4

Comparison of analysis sample baseline characteristics to full trial sample Chapter 7 Part 2 model 1, 2 and analysis 5 (H[1,2,5]) data

	Full trial sample			Analysis sample		
Baseline Characteristic	**Mean (SD)**	**Median [LQ, UQ]**	**n(%)**	**Mean (SD)**	**Median [LQ, UQ]**	**n(%)**
Intervention arm			305 (50.2%)			254 (51.5%)
Age	30.7 (10.7)	28.1 [22.6, 36.1]		31.1 (11.0)	28.7 [23, 36.2]	
Female			310 (51.1%)			260 (52.7%)
IV days (home)	13.2 (21.2)	0 [0, 19]		13.2 (20.5)	0 [0, 21]	
IV days (hospital)	12.7 (21.6)	0 [0,15.5]		12.1 (21.7)	0 [0, 14]	
FEV1%	59.5 (23.1)	59.1 [40.0, 77.9]		59.4 (22.5)	59.2 [40.7, 76.9]	
IMD quintile	3.04 (1.36)	3 [2,4]		3.03 (1.37)	3 [2, 4]	
Number of prescribed doses	3.26 (1.60)	3 [2, 4.29]		3.31 (1.59)	3 [2.20, 4.37]	
Max days in trial	463.8 (123.2)	486 [430.5, 538]		492.4 (89.0)	500 [446, 556]	
Numerator adjusted adherence during trial	48.3 (32.8)	48.0 [16.4, 79.0]		53.4 (30.3)	53.4 [26.7, 80.4]	
CFHH Pseudomonas status			349 (57.9%)			286 (58.2%)
Necessity	3.56 (0.74)	3.6 [3.10, 4.00]		3.59 (0.73)	3.7 [3.1, 4.0]	
Concerns	2.08 (0.54)	2.1 [1.7, 2.4]		2.08 (0.55)	2.1 [1.6, 2.4]	

IV= intravenous antibiotics; FEV1%= forced expiratory volume in first second (per cent); IMD quintile= Index of Multiple Deprivation quintiles; CFHH Pseudomonas status: chronic=diagnostic criteria fulfilled for chronic pseudomonas status using the CFHealthHub criteria tool (Hoo, Coates, et al., 2018).

Table F5

Comparison of analysis sample baseline characteristics to full trial sample Chapter 7 Part 2 model 3 (H^3) data (SRBAI)

Baseline Characteristic	Full trial sample Mean (SD)	Median [LQ, UQ]	n(%)	Analysis sample Mean (SD)	Median [LQ, UQ]	n(%)
Intervention arm			305 (50.2%)			253 (51.4%)
Age	30.7 (10.7)	28.1 [22.6, 36.1]		31.1 (11.0)	28.7 [23.0, 36.2]	
Female			310 (51.1%)			259 (52.6%)
IV days (home)	13.2 (21.2)	0 [0, 19]		13.3 (20.5)	0 [0, 21.3]	
IV days (hospital)	12.7 (21.6)	0 [0, 15.5]		12.1 (21.7)	0 [0, 14]	
FEV1%	59.5 (23.1)	59.1 [40.0, 77.9]		59.5 (22.5)	59.2 [40.9, 77.0]	
IMD quintile	3.04 (1.36)	3 [2, 4]		3.03 (1.37)	3 [2, 4]	
Number of prescribed doses	3.26 (1.60)	3 [2, 4.29]		3.32 (1.59)	3 [2.2, 4.37]	
Max days in trial	463.8 (123.2)	486 [430.5, 538]		492.4 (89.1)	499.5 [446, 556]	
Numerator adjusted adherence during trial	48.3 (32.8)	48.0 [16.4, 79.0]		53.4 (30.3)	53.3 [26.6, 80.5]	
CFHH Pseudomonas status			349 (57.9%)			285 (58.2%)
Necessity	3.56 (0.74)	3.6 [3.10, 4.00]		3.59 (0.73)	3.7 [3.1, 4]	
Concerns	2.08 (0.54)	2.1 [1.7,2.4]		2.08 (0.55)	2.1 [1.6, 2.4]	

IV= intravenous antibiotics; FEV1%= forced expiratory volume in first second (per cent); IMD quintile= Index of Multiple Deprivation quintiles; CFHH Pseudomonas status: chronic=diagnostic criteria fulfilled for chronic pseudomonas status using the CFHealthHub criteria tool (Hoo, Coates, et al., 2018).

Table F6

Comparison of analysis sample baseline characteristics to full trial sample Chapter 7 Part 2 model 4 (H^4) data (Effort)

Baseline Characteristic	Full trial sample			Analysis sample		
	Mean (SD)	Median [LQ, UQ]	n(%)	Mean (SD)	Median [LQ, UQ]	n(%)
Intervention arm			305 (50.2%)			252 (51.3%)
Age	30.7 (10.7)	28.1 [22.6, 36.1]		31.1 (11.0)	28.6 [23.0, 36.2]	
Female			310 (51.1%)			258 (52.5%)
IV days (home)	13.2 (21.2)	0 [0, 19]		13.1 (20.3)	0 [0, 21]	
IV days (hospital)	12.7 (21.6)	0 [0,15.5]		12.1 (21.7)	0 [0, 14]	
FEV1%	59.5 (23.1)	59.1 [40.0, 77.9]		59.5 (22.5)	59.4 [40.9, 77.0]	
IMD quintile	3.04 (1.36)	3 [2,4]		3.03 (1.37)	3 [2, 4]	
Number of prescribed doses	3.26 (1.60)	3 [2, 4.29]		3.33 (1.58)	3 [2.21, 4.37]	
Max days in trial	463.8 (123.2)	486 [430.5, 538]		492.2 (89.1)	499 [446, 555]	
Numerator adjusted adherence during trial	48.3 (32.8)	48.0 [16.4, 79.0]		53.5 (30.3)	53.4 [26.8, 80.6]	
CFHH Pseudomonas status			349 (57.9%)			285 (58.8%)
Necessity	3.56 (0.74)	3.6 [3.10, 4.00]		3.60 (0.73)	3.7 [3.1, 4.0]	
Concerns	2.08 (0.54)	2.1 [1.7,2.4]		2.08 (0.55)	2.1 [1.6, 2.4]	

IV= intravenous antibiotics; FEV1%= forced expiratory volume in first second (per cent); IMD quintile= Index of Multiple Deprivation quintiles; CFHH Pseudomonas status: chronic=diagnostic criteria fulfilled for chronic pseudomonas status using the CFHealthHub criteria tool (Hoo, Coates, et al., 2018).

Table F7

Comparison of analysis sample baseline characteristics to full trial sample Chapter 8 analysis 1 & 2 (H[1,2]) data

Baseline Characteristic	Full trial sample			Analysis sample		
	Mean (SD)	Median [LQ, UQ]	n(%)	Mean (SD)	Median [LQ, UQ]	n(%)
Intervention arm			305 (50.2%)			246 (52.1%)
Age	30.7 (10.7)	28.1 [22.6, 36.1]		31.3 (10.9)	29.0 [23.3, 36.6]	
Female			310 (51.1%)			246 (52.1%)
IV days (home)	13.2 (21.2)	0 [0, 19]		13.2 (20.7)	0 [0, 21]	
IV days (hospital)	12.7 (21.6)	0 [0,15.5]		12.0 (21.9)	0 [0, 14]	
FEV1%	59.5 (23.1)	59.1 [40.0, 77.9]		59.4 (22.6)	59.2 [40.0, 76.7]	
IMD quintile	3.04 (1.36)	3 [2,4]		3.02 (1.38)	3 [2, 4]	
Number of prescribed doses	3.26 (1.60)	3 [2, 4.29]		3.30 (1.57)	3 [2.18, 4.33]	
Max days in trial	463.8 (123.2)	486 [430.5, 538]		504.7 (65.4)	502.5[451.8, 558]	
Numerator adjusted adherence during trial	48.3 (32.8)	48.0 [16.4, 79.0]		54.2 (30.4)	55.4 [27.5, 81.1]	
CFHH Pseudomonas status			349 (57.9%)			273 (57.8%)
Necessity	3.56 (0.74)	3.6 [3.1, 4.0]		3.59 (0.74)	3.7 [3.1, 4.0]	
Concerns	2.08 (0.54)	2.1 [1.7,2.4]		2.07 (0.55)	2.1 [1.7, 2.4]	

IV= intravenous antibiotics; FEV1%= forced expiratory volume in first second (per cent); IMD quintile= Index of Multiple Deprivation quintiles; CFHH Pseudomonas status: chronic=diagnostic criteria fulfilled for chronic pseudomonas status using the CFHealthHub criteria tool (Hoo, Coates, et al., 2018).

Table F8

Comparison of analysis sample baseline characteristics to full trial sample Chapter 8 analysis 3 (H^3) data

Baseline Characteristic	Full trial sample			Analysis sample		
	Mean (SD)	Median [LQ, UQ]	n(%)	Mean (SD)	Median [LQ, UQ]	n(%)
Intervention arm			305 (50.2%)			239 (54.4%)
Age	30.7 (10.7)	28.1 [22.6, 36.1]		31.9 (11.1)	29.7 [23.8, 37.3]	
Female			310 (51.1%)			217 (49.4%)
IV days (home)	13.2 (21.2)	0 [0, 19]		12.8 (20.2)	0 [0, 18]	
IV days (hospital)	12.7 (21.6)	0 [0,15.5]		12.3 (22.7)	0 [0, 14]	
FEV1%	59.5 (23.1)	59.1 [40.0, 77.9]		59.1 (22.9)	59.1 [39.5, 76.4]	
IMD quintile	3.04 (1.36)	3 [2,4]		2.99 (1.37)	3 [2, 4]	
Number of prescribed doses	3.26 (1.60)	3 [2, 4.29]		3.42 (1.57)	3.06 [2.41. 4.45]	
Max days in trial	463.8 (123.2)	486 [430.5, 538]		504.2 (66.0)	502 [450, 558]	
Numerator adjusted adherence during trial	48.3 (32.8)	48.0 [16.4, 79.0]		59.3 (27.8)	64.2 [35.8, 84.2]	
CFHH Pseudomonas status			349 (57.9%)			256 (58.3%)
Necessity	3.56 (0.74)	3.6 [3.10, 4.00]		3.64 (0.73)	3.7 [3.3, 4.1]	
Concerns	2.08 (0.54)	2.1 [1.7,2.4]		2.08 (0.53)	2.1 [1.7, 2.4]	

IV= intravenous antibiotics; FEV1%= forced expiratory volume in first second (per cent); IMD quintile= Index of Multiple Deprivation quintiles; CFHH Pseudomonas status: chronic=diagnostic criteria fulfilled for chronic pseudomonas status using the CFHealthHub criteria tool (Hoo, Coates, et al., 2018).

Table F9

Comparison of analysis sample baseline characteristics to full trial sample Chapter 9 analyses of RQ 1&2

Baseline Characteristic	Full trial sample			Analysis sample		
	Mean (SD)	Median [LQ, UQ]	n(%)	Mean (SD)	Median [LQ, UQ]	n(%)
Intervention arm			305 (50.2%)			223 (54.7%)
Age	30.7 (10.7)	28.1 [22.6, 36.1]		32.2 (11.4)	30.3 [23.5, 38.1]	
Female			310 (51.1%)			203 (49.8%)
IV days (home)	13.2 (21.2)	0 [0, 19]		13.9 (22.0)	0 [0,19]	
IV days (hospital)	12.7 (21.6)	0 [0,15.5]		12.0 (21.5)	0 [0,14]	
FEV1%	59.5 (23.1)	59.1 [40.0, 77.9]		57.7 (23.2)	55.9 [37.1,75.9]	
IMD quintile	3.04 (1.36)	3 [2,4]		3.00 (1.36)	3 [2,4]	
Number of prescribed doses	3.26 (1.60)	3 [2, 4.29]		3.52 (1.63)	3.14 [2.55, 4.64]	
Max days in trial	463.8 (123.2)	486 [430.5, 538]		480.3 (108.9)	492.5 [440.8,549]	
Numerator adjusted adherence during trial	48.3 (32.8)	48.0 [16.4, 79.0]		62.0 (27.7)	69.6 [41.3,85.6]	
CFHH Pseudomonas status			349 (57.9%)			246 (60.3%)
Necessity	3.56 (0.74)	3.6 [3.10, 4.00]		3.70 (0.72)	3.70 [3.30,4.10]	
Concerns	2.08 (0.54)	2.1 [1.7,2.4]		2.04 (0.54)	2.00 [1.60,2.40]	

IV= intravenous antibiotics; FEV1%= forced expiratory volume in first second (per cent); IMD quintile= Index of Multiple Deprivation quintiles; CFHH Pseudomonas status: chronic=diagnostic criteria fulfilled for chronic pseudomonas status using the CFHealthHub criteria tool (Hoo, Coates, et al., 2018).

Table F10

Comparison of analysis sample baseline characteristics to full trial sample Chapter 9 analyses of RQ 3&4

Baseline Characteristic	Full trial sample Mean (SD)	Median [LQ, UQ]	n(%)	Analysis sample Mean (SD)	Median [LQ, UQ]	n(%)
Intervention arm			305 (50.2%)			35 (52.2%)
Age	30.7 (10.7)	28.1 [22.6, 36.1]		30.1 (11.4)	27.8 [22.0, 33.8]	
Female			310 (51.1%)			32 (48.8%)
IV days (home)	13.2 (21.2)	0 [0, 19]		7.24 (14.3)	0 [0,10]	
IV days (hospital)	12.7 (21.6)	0 [0,15.5]		6.13 (11.7)	0 [0, 8]	
FEV1%	59.5 (23.1)	59.1 [40.0, 77.9]		67.0 (20.3)	64.3 [51.8, 84.2]	
IMD quintile	3.04 (1.36)	3 [2,4]		3.03 (1.30)	3 [2,4]	
Number of prescribed doses	3.26 (1.60)	3 [2, 4.29]		3 (0)	3[3,3]	
Max days in trial	463.8 (123.2)	486 [430.5, 538]		421.0 (158.0)	458 [387.5,526]	
Numerator adjusted adherence during trial	48.3 (32.8)	48.0 [16.4, 79.0]		47.0 (34.9)	39.2 [14.8,84.0]	
CFHH Pseudomonas status			349 (57.9%)			37 (56.1%)
Necessity	3.56 (0.74)	3.6 [3.10, 4.00]		3.7 (0.73)	3.7 [3.3,4.1]	
Concerns	2.08 (0.54)	2.1 [1.7,2.4]		2.05 (0.46)	2.1 [1.7,2.4]	

IV= intravenous antibiotics; FEV1%= forced expiratory volume in first second (per cent); IMD quintile= Index of Multiple Deprivation quintiles; CFHH Pseudomonas status: chronic=diagnostic criteria fulfilled for chronic pseudomonas status using the CFHealthHub criteria tool (Hoo, Coates, et al., 2018).