



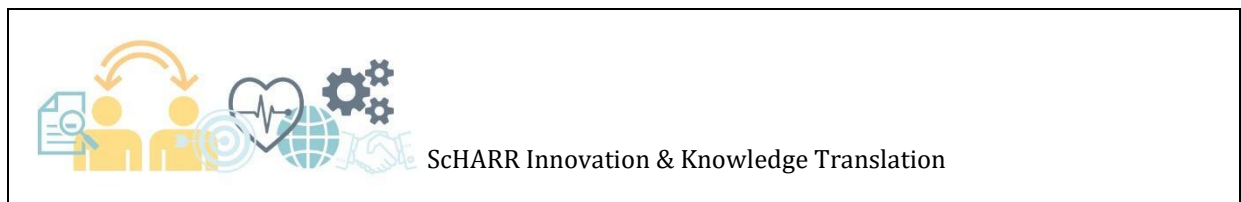
School of
Health and
Related
Research

**NIHR CLAHRC
Yorkshire and Humber**

Sheffield City Region Test Bed

‘Perfect Patient Pathway’ Final Evaluation Report

July 27th 2018



Document Status

Title	Sheffield City Region Test Bed, 'Perfect Patient Pathway' Final Evaluation Report
Authors	S. Ariss, J. Read, M. Franklin, J. Dawson, N. Nasr, E. Scott, B. Field, K. Lowrie, H. Dunn, L. Cook, R. Simmonds
Date and Version	15.08.18: V.20.1
Supplementary Material	Sheffield City Region Test Bed, 'Perfect Patient Pathway' Final Evaluation Report - Appendices 15.08.18: V.3.1

This report presents the findings of an independent evaluation of the Sheffield City Region Perfect Patient Pathway (PPP) Test Bed Programme. It has been prepared by The University of Sheffield, in collaboration with National Institute for Health Research, Collaboration for Leadership in Applied Health Research Yorkshire and Humber (NIHR CLAHRC YH) and Healthwatch Sheffield under contract to Sheffield Teaching Hospitals NHS Foundation Trust (STH). Information from other sources (e.g. PPP Test Bed Project Management Office (PMO) and market research carried out as part of the programme by innovator companies) is used with permission and sources are identified within the text. The findings and interpretations in this report are those of the authors and do not necessarily represent the views of the services or organisations involved in the delivery of the programme or those of the NHS, the NIHR or the Department of Health and Social Care.

Intellectual property rights belong to STH. However, the authors and their organisations retain licence to use this report, its contents and any other intellectual property arising from the evaluation activities for academic teaching and research purposes, including but not limited to publications and other related dissemination activities.

The authors have taken all reasonable care to ensure the accuracy and completeness of information used in the production of this report. However, they do not accept responsibility for any legal commercial or other consequences that might result from the use of any inaccurate or incomplete information that was supplied to them during the preparation of this report.

Contact Details:

Dr Steven Ariss

The University of Sheffield

School of Health and Related Research

Innovation Centre

217 Portobello

S1 4DP

s.ariss@sheffield.ac.uk

Contents

List of Tables	5
List of Figures.....	5
Glossary	5
1 Executive summary	8
1.1 Falls Prevention.....	8
2 Introduction.....	16
2.1 Purpose and Overview	16
2.2 Research or Evaluation	16
2.3 Ethical approvals	17
2.4 Evaluation scope and approach	17
2.5 Governance structure	18
2.6 Engagement with the public and service users.....	19
2.7 Core projects.....	19
2.8 Test Bed 'Plus'	19
3 Falls prevention/Strength and Balance.....	20
3.1 Engagement report provided by Healthwatch Sheffield.....	21
3.2 Evaluation Focus.....	22
3.3 Methods	22
3.4 Key Findings.....	25
3.5 Limitations.....	34
3.6 Conclusions and implications.....	37
3.7 Recommendations.....	38
4 Asthma	39
4.1 Engagement report provided by Healthwatch Sheffield.....	41
4.2 Evaluation Focus.....	42
4.3 Methods	43
4.4 Key findings.....	44
4.5 Limitations.....	53
4.6 Conclusions and implications.....	53
4.7 Recommendations.....	54
4.8 Asthma project: Secondary care asthma clinic roll out	55
5 Emergency Care Mobile app (SOS UK).....	57
5.1 Engagement report provided by Healthwatch Sheffield.....	58
5.2 Evaluation Focus and methods.....	58
5.3 Key findings.....	59
5.4 Conclusions and implications (including limitations)	61
5.5 Recommendations.....	63
6 Digital Care Home.....	64

6.1	Background	65
6.2	Engagement report provided by Healthwatch Sheffield.....	66
6.3	Methods	67
6.4	Key Findings.....	69
6.5	Limitations.....	78
6.6	Conclusions and implications.....	79
6.7	Recommendations.....	80
7	Diabetes	81
7.1	The technology and intervention	82
7.2	Engagement report provided by Healthwatch Sheffield.....	83
7.3	Evaluation Focus.....	83
7.4	Methods	84
7.5	Key findings.....	84
7.6	Limitations.....	87
7.7	Conclusions and implications.....	87
7.8	Recommendations.....	88
8	Test Bed 'Plus'	89
8.1	3Rings Project	89
8.2	Digital Health Training project.....	91
8.3	Predictive Analytics	92
9	Programme-Wide Evaluation	94
9.1	Background	95
9.2	Technology and intervention	95
9.3	Rationale for the intervention	97
9.4	Evaluation Focus.....	97
9.5	Methods	97
9.6	Key findings.....	98
9.7	Conclusions and implications.....	99
9.8	Overall conclusions and implications.....	106
9.9	Recommendations.....	109
	References	110

List of Tables

Table 1: Meeting frequency for projects	19
Table 2: Falls Prevention Project Summary	20
Table 3: Falls prevention cohort size and probability of fall, injurious fall, or transition to long-term care by age group	24
Table 4: Falls risk assessment sensitivity, specificity, and costs.....	24
Table 5: Falls prevention intervention efficacy and costs ¹¹	24
Table 6: Falls prevention Utility values and costs associated with Markov model states.....	24
Table 7: Falls prevention PROMs completed at baseline, three and 6-month follow-up.....	31
Table 8: Cost effectiveness summary	33
Table 9: Results of Community strength and Balance Assessments	35
Table 10: Asthma project summary	39
Table 11: Asthma project PROM completion rates for six and 3-month programmes	51
Table 12: Adherence to medication regimen by age group.....	52
Table 13: Asthma project adherence to medication regimen by medication type	52
Table 14: Asthma project adherence to medication regimen over time	53
Table 15: Emergency Care Mobile app (SOS UK) project summary	57
Table 16: Digital Care Home project summary	64
Table 17: Number of hospital contacts needed to avoid for DCH intervention cost-neutrality.....	76
Table 18: Emergency contact frequency.....	77
Table 19: Emergency contact rate per 12 months	77
Table 20: Whole Programme summary	94

List of Figures

Figure 1: Decision tree example – Care pathway versus no intervention	24
Figure 2: Markov model transition diagram	25
Figure 3: Recruitment rates by month for QTUG assessment.....	29
Figure 4: Recruitment Flow Diagram.....	29
Figure 5: Emergency care mobile app installation and evaluation survey link use.....	59
Figure 6: Emergency care mobile app usage and retention.....	60
Figure 7: The PPP vision	96

Glossary

A&E	Accident and Emergency
AHSN	Academic Health Science Network
CCG	Clinical Commissioning Group
CLAHRC	NIHR Collaboration for Leadership in Applied Health Research and Care
CLAHRC YH	NIHR Collaboration for Leadership in Applied Health Research and Care for Yorkshire and Humber
Combinatorial innovations	When two or more innovations are used at the same time
COPD	Chronic Obstructive Pulmonary Disease
CPD	Continued professional Development
CSRI	Client Service Receipt Inventory
DCH	Digital Care Home
DP	Data Protection
DPIA	Data Protection Impact Assessment
EAC	Equivalent Annual Cost
eFI	Electronic Frailty Index

Engagement summary	Summary report from Healthwatch Sheffield, summarising the work they did to engage members of the public and people with long term conditions in the work of the PPP Test Bed
ENT	Ear, Nose & Throat
EQ-5D-5L	Patient-reported outcome measure of health status: EuroQol Five Dimension with Five Levels questionnaire
EVPI	Expected Value of Perfect Information
FaME	Falls Management Exercise Programme
FC	Family Carer
GP	General Practitioner: a doctor based in the community who treats patients with minor or chronic illnesses and refers those with serious conditions to a hospital
HAM	Home safety Assessment and Modification
HbA1c	HbA1c is a measure of glycated haemoglobin. This is a marker that can determine an average blood glucose levels over the previous 3-months
HCA	Health Care Assistants
HCPs	Healthcare Professionals
HRA	Health Research Authority
ICECAP-A	Patient-reported outcome measure: ICEpop CAPability Measure of capability for the general adult (18+) population
ICER	Incremental Cost-Effectiveness Ratio
ICT team	(Sheffield) Integrated Community Therapy team
IG	Information Governance
IP	Intellectual Property
IQR	Interquartile Range
Kinesis QTUG™	'Quantitative Timed up and Go' technology provided by Kinesis Health Technologies Ltd
Markov model	A stochastic model describing a sequence of possible events in which the probability of each event depends only on the state attained in the previous event
MDT	Multi-disciplinary Team
MEWS	Modified Early Warning Score
MFES	Patient-reported outcome measure: Modified Falls Efficacy
MPR	Medication Possession Rate
NEWS	National Early Warning Score
NGH	Sheffield Teaching Hospitals' Northern General Hospital site
NHSE	National Health Service England
NSHI	National Services for Health Improvement is an independent service provider commissioned by Teva Pharmaceutical Industries to provide specialist nursing support for their CareTRx programme (within the Asthma PPP Test Bed project)
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
PAM-13	Patient-reported outcome measure: Patient Activation Measure 13-item
PC	Personal Computer (desk top/lap top)
PID	Project Initiation Document
PMO	Programme Management Office: central support structure, designed to provide assistance to change and delivery initiatives
PPG	Patient Participation Groups
PPI	Patient and Public Involvement
PPP Test Bed	Sheffield City Region Perfect Patient Pathway Test Bed Programme
Primary Care	Healthcare provided in the community for people making an initial approach to healthcare practitioners for advice or treatment
PROM	Patient-Reported Outcome Measure
PSA	Probabilistic Sensitivity Analysis
PSP	Patient Support Programme
PSSRU	Personal Social Services Research Unit
PwD	Person with Dementia
QR (code)	Quick-read: a visual digital code that, when scanned (e.g. with a smartphone) links to a web-based resource

QTUG	Quantitative Timed up and Go
QALY	Quality Adjusted Life Year
RaR	Rate Ratio
RCP	Royal College of Physicians
RCT	Randomised Controlled Trial
REC	Research Ethics Committee
ReQoL-10	Patient-reported outcome measure : Recovering Quality of life 10-item to assess quality of life for people with mental health conditions
RHH	Sheffield Teaching Hospitals' Royal Hallamshire Hospital site
ROI	Return On Investment
RR	Risk Ratio
ScHARR	The University of Sheffield's School of Health and Related Research
SD	Standard Deviation
Secondary Care	Specialised healthcare that requires more specialised knowledge, skill, or equipment than the primary care practitioner can provide
SOS	'Save Our Souls' – emergency call
SOS UK	A healthcare mobile app that records medical information for access in an emergency or routine consultation
SPA	(Sheffield) Single Point of Access
SRO	Senior Responsible Officer: This was the senior, accountable figure for the PPP Test Bed. They are the visible owner and leader of the PPP Test Bed Programme,
St Luke's	Hospice also providing community specialist palliative care services in Sheffield
Stata	An integrated statistics analysis software package
STH	Sheffield Teaching Hospitals NHS Foundation Trust
SWEMWBS	Short Warwick-Edinburgh Mental Wellbeing Scale to enable the monitoring of mental wellbeing in the general population
SystemOne	A software system used by some NHS services to share patient data securely across services (within Sheffield some NHS services (mostly primary care) use it, but not all)
TAG	The PPP Test Bed Advisory Group. A group of Individuals with long term conditions and relatives or carers of people with such conditions who advised the PPP Test Bed throughout its life
Test Bed Champions	Individuals with long term conditions relating to specific projects)
Teva	Teva Pharmaceutical Industries Ltd
TUG	Timed up and Go test to assess falls risk
UoS	The University of Sheffield
VAS	Visual Analogue Scale
WiFi	Wireless internet connection provided through a local router
WTP	Willingness To Pay

1 Executive summary

This report focuses on the implementation of a Test Bed site (Sheffield Perfect Patient Pathway (PPP)) within the National Health Service England (NHSE) national Test Bed programme. It is important to establish the process of implementation of individual projects and the quality of evidence that can be gained from this 'testing'. However, the main purpose of these projects within the programme are as examples to try methods and infrastructure; to establish the effectiveness of this type of programme for identifying, implementing and evaluating the use of health technology in the NHS.

Whilst the main report covers the broad range of activity that took place within the PPP Test Bed programme, this summary will only describe the substantive findings from across the three main interventions that have been evaluated within re-designed pathways: Falls prevention, Asthma and Digital Care Homes. These findings will be included here for the purpose of demonstrating key themes related to the programme-wide focus of the evaluation.

1.1 Falls Prevention

This project used the Kinesis QTUG™ (Quantitative Timed Up and Go) technology. The technology uses body worn sensors and a mobile software app (including a falls risk questionnaire) to assess frailty, mobility, and falls risk. It can be used by non-specialists with minimal training, and is wireless and portable. QTUG uses proprietary algorithms to give an objective assessment of falls risk based upon the 'Timed up and go' test. There is a NICE Medtech Innovation Briefing on QTUG <https://www.nice.org.uk/advice/mib73>.

The rationale underlying the intervention is that the digital technology replaces the manual Timed Up and Go (TUG) test so that with minimal training clinicians and health care assistants using the Kinesis QTUG™ device can assess and identify older people aged 65 and over at risk of a future fall. This could enable appropriate clinical decisions and referrals to falls-prevention interventions based on the information obtained. The key expected outcomes of the intervention are:

- Reduced falls risk
- Reduced injurious falls (resulting in reduced mortality & morbidity)
- Improved quality of life and wellbeing of people at risk of falling

In this project, the new care pathway involved identification of patients registered with one of three GP clinics in Sheffield. Patients were aged 65+, identified as 'moderately frail' on the Electronic Frailty Index (eFI) and had no clinically reported falls. Patients that are known to have fallen should be automatically referred to Integrated Community Therapy team (ICT) falls prevention services. They were invited to attend an assessment clinic at their GP practice. If they were assessed as being at high risk of having a fall, they were referred to the ICT for a fall prevention intervention. This is compared to usual care in which people are generally referred for specialist falls risk assessment and intervention after a fall has been recorded.

Falls prevention was identified as a clinical priority within workshops/engagement events held with clinicians in March 2016. Project planning started in September 2016; the project went live in December 2016. The evaluation ethics approval was gained in February 2017. This lag was a result of clarity regarding the project's aims, outcomes and implementation plans not appearing until November/December 2016 as all involved struggled to identify an agreed evaluation methodology and outcome measures. The pace and demands to set up and implement a project and evaluation in three months was challenging. Evaluation ethics documentation was submitted in December 2016, and requested amendments were addressed collaboratively.

This experience highlights a key recurrent feature of the programme. Once the interventions were designed, there was a desire to move straight to the implementation phase. However, it is not until the intervention is designed that the evaluation design can be agreed in detail, protocols written, information sheets and consent forms designed, information governance issues explored etc. Therefore, early on in the programme, timing of implementation and evaluation activities was misaligned and implementation

design was not always conducive to successful evaluation approaches. However, this situation improved as co-production processes became embedded in day-to-day operations.

The programme management and governance arrangements worked well to make rapid consensual adjustments to the project; as evidence suggested required changes, such as changing the description of the intervention from being concerned with 'falls' to 'strength, mobility and balance', reducing the length of time for the assessment, and providing vouchers to reimburse travel expenses.

180 patients were invited for assessment, and 60 attended. A total of 26 were assessed as high risk, but 7 declined referrals. Of the 20 people successfully referred to the ICT falls prevention service, 19 patients were assessed as 'high risk' of falls with the QTUG device, one was assessed as 'moderate risk' of falls but reported a history of falls, so was also referred onto the ICT. Follow-up assessments took place 3-months after referral and final follow-up assessments were carried out on discharge from the ICT falls prevention service.

The total number returning quantitative data are very small. However, descriptive findings are interesting. QTUG falls questionnaire data were available for 12 of the participants at baseline; 50% (6) reported having fallen in the last 12 months (despite not having recorded fall in their records). All were on four or more prescription medication and many had issues with mobility, their feet, dizziness blood pressure and vision. Of the ten subjects for whom we have baseline and 6-month follow-up data:

- Five showed a decrease in falls risk estimate (range -30.76 to -1.13 percentage points)
- Five showed an increase in falls risk estimate (range 0.3 to 14.7 percentage points)

Owing to small numbers of participants, no statistically significant conclusions can be drawn from the PROMs data. In order to assess the cost-effectiveness, in light of the lack of study-based data, an exploratory economic model was developed. According to this model, screening with QTUG *dominates* (produces more QALYs and cost-savings) over screening with TUG irrespective of which falls prevention intervention follows. Therefore, if falls-risk screening is perceived as desirable by decision makers or clinicians, then based on this analysis, QTUG should be considered over TUG for this assessment in older people aged 65 to 89.

Exploratory economic modelling based results suggest that a falls prevention care pathway has a higher probability of being cost-effective (high 90% to 100%) at a specified willingness to pay (WTP; e.g. £30,000 per QALY) threshold when screening and falls prevention interventions are utilised in a population aged 75 to 89 compared to no care pathway. When screening is implemented in those aged 65 to 69, compared to doing no care pathway the probability of the pathway being cost-effective is almost zero, and it is around about 50% in those aged 70 to 74. The exact results are dependent on the specific pathway implemented in terms of type of fall-risk screening and falls-prevention intervention utilised.

For the care pathway to be cost-effective, there is a need to identify those most likely to have an injurious fall using a falls-risk assessment which has high sensitivity and specificity, in order to correctly refer people to a falls prevention intervention which has a high rate of efficacy, at a reasonable cost which is off-set against downstream care cost-savings associated with when an older person has an injurious fall.

Community Strength and Balance Project: Between February 2018 and April 2018 a community strength and balance project, using the Kinesis QTUG, was established, led by Healthwatch Sheffield and the STH telehealth team. The programme planning group discussed various opportunities to assess recruitment in different settings and this aligned with the PPP Test Bed Advisory Group (TAG) suggesting using the QTUG in community groups.

Healthwatch Sheffield contacted voluntary sector groups (such as lunch clubs) and sheltered housing schemes with members over 65 years, based on their local knowledge and experience. Once the project had started, other groups also made contact with Healthwatch Sheffield after hearing about it. The Telehealth team also made contact with community groups and sheltered housing schemes.

The QTUG assessment was conducted on those people who gave signed, informed, consent. The QTUG results (high/medium and low risk of falls) were discussed with the person. A referral for a falls prevention service with the Integrated Care Team (ICT) –Therapy was offered via Single Point of Access (SPA) if the person was either:

- identified by the QTUG as at high risk of falling,
- appeared to have needs in the clinical judgement of the community nurse although not identified as high risk by the QTUG score
- reported a fall in the recent past

Assessments and referrals: 281 people had been seen by the time the project closed in April 2018.

Nine of these were not fully assessed (1 under 65 years of age, 1 equipment failure, 1 unable to walk 3 meters, 6 not carried out due to concerns by nursing staff (these were immediately referred)). Of the 272 assessments conducted 138 (50.7%) were successfully referred to the ICT for the falls prevention intervention. However, the consequence of this was that a long waiting list of eight to ten weeks soon developed for the ICT falls prevention service. This indicates that a sustainable assessment service would require improved coordination with prevention and support services and possibly increased capacity or alternative support to manage referrals.

This finding indicates that systems approaches are important, to consider unintended consequences and explore effects across organisations, service-providers and pathways.

Secondary care balance clinic: Between Dec 2017- March 2018 the QTUG assessment device was introduced into a weekly, hospital-based balance clinic. Previously, referrals made to falls prevention were not systematic and there were no formal assessments of falls risk. Initial feedback indicates that the assessment process is suitable for a busy clinical setting, aids interactions with patients, provides systematic assessment and referral routes and is well received by patients and clinicians. A total of 45 out of 84 attendees were assessed using the QTUG device; seven of these with high falls risk scores (16%) were referred to the ICT falls service.

Two clinicians provided the following feedback:

- More targeted referrals than previous random ones
- Takes around 5-7 minutes to do the test; 'hence sits well in a busy clinic'
- Trialled by two users (Consultant and Clinical Fellow)
- 'Easy to put on patient' (with the new Velcro bands - not the original bandage)
- 'Easy to set up and use with the software'
- Reported by clinicians as 'well received by patients - user friendly'
- 'Data is very well presented for explaining to the patient'

1.1.1 Asthma project

The asthma project was characterised by a combinatorial innovation and implementation model that was largely designed by the innovator company. Planning began in June 2016 and the project went live in December 2016. It was originally agreed that the evaluation would rely on retrospective analysis of routinely collected data, whilst a proposal was collaboratively developed. Discussions about what data were available for the evaluation and how to collate and manage the data, as well as associated IG issues, were very time consuming. Evaluation ethics approval for primary data collection (PROMs) was gained in April 2017. Initially the project involved a 6-month intervention (110 recruits), which was reduced to 3-months towards the end of the project to aid continued recruitment (22 recruits), and finally a secondary care pathway was trialled (4 recruits).

The CareTRx programme was developed by Teva Pharmaceutical Industries Ltd. and included the following components:

- CareTRx Sensor that fits over the top of an inhaler device (to record dosing events and lights up to remind user to take medication);
- CareTRx Clinic (to train the user, familiarise them with the intervention download the app to their phone and assess likely adherence level and ascertain how to target the PSP digital coaching);
- Linked smartphone mobile application (App): CareTR App;
- Data analytics dashboard and a data hosting cloud (for users and clinicians to view medication history);
- Patient Support Programme (PSP digital coaching) (including email, text messaging and website, linked to behaviour change approaches).

The CareTRx intervention aimed to supplement the usual care pathway by providing data on inhaler usage and treatment adherence to the patient and the healthcare professional reviewing the patient's medication.

The low recruitment rates are important to note: 132 people were recruited from 6,043 people with asthma on GP lists (2.2%). The market research report states that approximately 4,500 invitations were sent out, which equates to a 2.9% response rate.

There were a large number of barriers to the collection of comparative data to assess the effectiveness or cost effectiveness of the intervention. These included the technology not being available on the open market. Therefore, no costs for the intervention were available. Owing to the evaluation status of the study it would not have been possible to recruit a comparison group within ethical guidelines. Additionally, there were no resources for recruiting a comparison cohort, which would not have had the incentive of being offered the technology, and therefore would be expected to have even lower recruitment rates than the intervention group (2.9%). The intervention was designed specifically to assess acceptability and utility of the intervention and feasibility of the implementation. Implementation was therefore not designed at the outset with consideration of an effectiveness evaluation (e.g. collection of appropriate baseline outcome measures).

The challenges of a global company such as Teva and a large NHS organisation such as STH trying to ensure they met their obligations for data protection whilst negotiating data sharing between organisations were reported as very challenging and given as one reason why the project took so long to implement. However, once the project was running, and NHS patient adherence and usage data and patient reported outcome data was shared between the NHS, Teva and the evaluation team, governance arrangements appeared effective.

Recruitment was slower than anticipated and GP practices soon became exhausted of potential participants; requiring further recruitment of GP practices along with further identification and invitation of potential participants.

Quantitative evaluation focused on adherence to scheduled medication doses. Data were received from 83 patients (46 women and 37 men). Patients' age ranged from 18 to 77, with a mean of 41.9 (standard deviation 18.1). There are numerous difficulties interpreting the current evidence on the relationship between medication adherence and exacerbations. Adherence levels vary considerably, and those with higher adherence rates tend to have a higher symptom burden and can therefore experience more frequent exacerbations, despite better adherence. However, whilst there is some good quality evidence to support the link between adherence and exacerbations, establishing a level of 'adherence' varies between studies (e.g. >79%, >49%, >74%).

Although there is no way to estimate change or attribute adherence levels to the intervention, it is worth noting that adherence in month-one was 65%, which is higher than Barnes et al (2015) highest levels of adherence in a normal population. Whilst this does drop-off over time, the mean average remains above 50%, which some studies define as above adherent level. Interestingly, the results of this evaluation

demonstrate a 78% adherence in 50-59 year olds, which is close to the >79% highest definition of adherence in the Barnes et al (2015) review.

The PROMs for the evaluation had very low completion rates. The most complete set of data was for the EQ-5D-5L (n=29), which was already included in the original suite of measures, whereas the others were implemented later on. For other PROMs sets of pre and post data ranged from 3 to 13 complete sets. Some reasons provided for this included:

- too many questions overall
- too much similarity between groups of questions, with patients confused what was different to the last question they were asked to score
- non-asthma related questions were considered irrelevant by some patients and as a result did not really see the point of answering them

There were some positive perceptions of the project reported, including offering more than usual care and having the potential to benefit patients. A representative of the innovator company stated that engaging with the NHS had enabled iterative development/ design of this combination of complex interventions, and that being able to test them together, out in the real world, with all the challenges that involved, was useful to them as a company. They also experienced valuable learning about how to integrate processes included Information Governance, Data Protection, legal issues, collaboration frameworks, Intellectual Property, and clinical pathways.

However, as would be expected for the first implementation of a complex set of interventions there were a number of reported problems. For instance, the appointments were not coordinated with routine care appointments; they were considered too long and to begin with were only during the day, which was difficult for people that worked. It was considered difficult to engage with people that had well-controlled asthma. Some patients were 'put off' by glitches with the technology, some of which were reportedly not fully resolved. Perhaps most importantly the data recorded on the app was often reported to be inaccurate owing to syncing problems and delays in recording readings.

One of the key findings for this project is the large amount of learning that needs to take place at the early stages of testing out technology (especially combinations of innovations) in real-world settings. The implementation was not designed for streamlined recruitment, or to be particularly user-friendly. The process of collecting data for the evaluation (especially when combined with the data being collected for the intervention) was reported to be particularly off-putting and resulted in very low completion rates. Recruitment was particularly difficult, and the low numbers recruited at each site resulted in having to recruit additional sites and set up working relationships with additional practices, which was increasingly time consuming.

Difficulties were experienced related to coordinating and agreeing the organisational governance, project planning, data sharing etc. These issues slowed the delivery of the project, and were reported to produce a project which lacked the required flexibility and responsiveness.

1.1.2 Digital Care Homes

A pilot care home was used for monitoring from June to September 2017. The recruitment of additional care homes to project happened from September to December 2017 and monitoring began in these six other homes from October 2017 to January 2018. Between January and March 2018 evaluation interviews were undertaken with participating care home staff. There were 67 residents who used the intervention across the seven care homes. The number per home ranged from 5 to 16.

The technology for this project consists of equipment to measure vital signs (e.g. temperature, blood oxygen levels, blood pressure etc.) and a tablet app to transmit these readings via a Digital Health Platform to the Single Point of Access (SPA) team. Care home staff use the technology to submit NEWS observations and share these electronically with the nurses at SPA who are able to view readings and alerts through the Digital Care Home portal. If the readings trigger an alert at SPA, then the care home is called by phone. If a referral is required following an alert and dialogue with the care home team, the

information can be uploaded into patient records via SystemOne or made visible to NHS services via access to the portal.

The issue that this project is trying to address is the high level of emergency attendances and admissions for care home residents. Care home residents have 40-50% more hospital admissions and Accident and Emergency attendances than the general population age 75 and over.

It is possible that translating the National Early Warning Scores (NEWS) from acute care situations to care homes and linking alerts to clinical responses, that emergency attendances and non-elective hospital admissions might be reduced. It was recognised that monitoring has the potential to indirectly improve the care of residents.; it is possible that the act of regular monitoring increases the knowledge and understanding of care home staff, and also increases vigilance and attention. There is evidence in the literature for this and it is corroborated by notes associated with responses to alerts, where care home staff members have taken actions to improve measures. There is also a possibility that involvement in the project has helped to improve the level of service that some homes are receiving from GPs.

This project has been generally well-received; particularly by care home staff and SPA there are also some areas of uncertainty where further investigation, refinements and possible improvements are suggested. The potential impact on other related health and social care services and appropriate support need to be better understood prior to scaling up. It would also be worthwhile understanding whether there are types of care homes that are more receptive to the intervention or stand to benefit more than others. The optimum selection of appropriate residents should also be investigated. There were some concerns raised about the appropriateness of the NEWS to trigger alerts. Therefore, the exploration of modified scores would be beneficial.

Out of 55 alerts 11 (20%) were confirmed as incorrectly entered and there were 3 incidents of SPA staff not being able to make contact with the home. Incorrect entries are followed up and staff supported to provide accurate measures. Of the remaining 41 contacts with care homes, following alerts being triggered, 8 cases were either escalated to GPs or had a GP visit planned for the next day. In two cases, alerts were attributed to contextual factors that could be resolved (i.e. high temperature in room and lying position). In 16 cases care home staff reported having no concerns about the alert. Other alerts resulted in care home staff reporting continuing monitoring.

An exploration of reduced admissions required to achieve cost neutrality found that, across the 67 residents within seven care homes in the DCH study, the intervention would need to avoid 21.5 long stay non-elective inpatient contacts per year at £2,984.71 per contact to achieve cost-neutrality; which is equivalent to a decrease in 0.32 long stay non-elective inpatient contacts per resident/year. If the decision maker wanted to re-coop the technology costs via hospital cost-savings over the first year of implementation, the intervention would have to avoid 0.33 long stay non-elective inpatient contacts, or 1.62 short-stay non-elective inpatient stay contacts, or 6.72 emergency medicine contacts per resident in the first year. The number of admissions avoided per person to achieve cost neutrality can reduce with increased numbers of residents being monitored, but this depends on the tiered costing of the Inhealthcare platform dependent on number of people connected the platform.

A pre-post examination of emergency contacts for care home residents being monitored in the project reveals an interesting effect. When all cases are considered, there appears to be a slight decrease in the number of emergency contacts during the intervention period compared with the baseline period. However, when only considering participants who were resident in the care home at start of the baseline period, there appears to be a slight increase in emergency contacts over the intervention period compared with baseline. The observed number shows a slight increase (a rate of 0.6 per year higher in the intervention period), but the confidence interval here is (-0.4, 1.6), therefore including zero and meaning no firm conclusions could be drawn. Therefore, the conclusion is that there is no difference that can be inferred between the rate of emergency contacts while using the intervention compared with baseline.

The digital care home project provides a good example of the integration of innovative digital technology to create new linkages and pathways of care between existing services, which are already providing similar functions. The technology is creating a direct link between the care home and comprehensive referral services. As the decision whether to provide monitoring for specific residents is made within the care home, with the resident and their relatives, this has the potential to bring decisions about care closer to the service-user. This type of intervention is fairly controllable and predictable at an operational level as the technology is procured and operated by the organisations and there are reasonably straightforward activities required to ensure that the system works as intended. Recruitment has proved to be challenging, but feasible to recruit meaningful numbers with small resources. The SPA team have not reported any problems with capacity as a result of involvement in the project, indicating that the system is not being put under strain as a result of the innovation.

1.1.3 Programme-wide evaluation

The PPP Test Bed programme started in February 2016. It was intended to be completed in March 2018, but was extended until June 2018. The vision of the programme was to integrate data systems across primary, community, secondary and social care and mental health services; to create pooled data resources. Digital monitoring technology would be used alongside self-management tools. An intelligence centre would provide predictive analytics, and strategic decision support. All of these systems would be based around holistic patient needs with the facility for real-time alerts to prompt rapid service decision-making. Whilst this was an ambitious proposal, there was a great deal of activity focused on achieving this vision; infrastructure and processes for implementing and testing digital innovation in the NHS was successfully created and some important learning resulted.

The main characteristics of the programme were interdisciplinary working and cross-organisational partnership between NHS providers, commissioners, University, industry, and patient groups to implement different innovative health technologies in NHS settings as a new way of thinking and working with different stakeholders. However, the different expectations of innovators indicate that there was a lack of alignment about the aims of the programme initially. For instance, innovators felt there was a lack of clarity on procurement frameworks, which contributed to difficulties managing innovators' expectations.

Across stakeholder groups, there was not an agreed perception about the type of evidence that the evaluation was supposed to provide. This was further compounded by delays to the commissioning of the national evaluation team, part way through the programme. However, the various partners developed improved ways of working together effectively over time to coordinate activities and design mutually agreeable interventions. This inevitably involved developing processes and relationships and sharing knowledge to manage conflict and attain compromise.

For the NHS locally, the key value of the Test Bed programme was the testing of methods and infrastructure for the identification, implementation and evaluation of new technologies. This was facilitated by the inclusion of a wide variety of technologies in a wide range of settings. It was a surprise to find that some technology was at the development stage and not completely ready to be deployed as fully functioning products or solutions. However, the programme created an opportunity for these technologies to be tested out with users in real life situations. Some innovators learned an enormous amount from the PPP Test Bed. Their technology was modified along the way and they were able to build market-ready products. The learning from the PPP Test Bed informed technology design and as a result all innovators developed improved iterations of their technology.

The programme had a short timescale, especially considering the implementation and testing of novel technologies and the need to redesign pathways and integrate the requirements of a large number of stakeholders. There were challenges agreeing collaboration agreements with innovators and partners; this took longer than expected. The pace of change required was not always compatible with information governance (IG), organisational governance and ethics approvals. Early pressures from the funders to recruit large numbers of service users had an effect on the design of the programme; promoting the seeking of opportunities for large numbers of recruits. However, owing to the experimental nature of the

programme, recruitment was unexpectedly slow in some projects and more rapid than expected in others. The design of the programme evolved and shifted over time to be more responsive to the emerging context. Also related to the short time-scales and pressure for recruitment, many of the projects were characterised by various issues related to lack of readiness; whether this was in terms of market readiness of the technology, readiness for good quality data collection, or organisational readiness for implementation. More time for planning and achieving a mutually agreed state of readiness amongst all partners prior to roll-out would have been beneficial.

The PPP Test Bed governance structure did not change a great deal. However, the relatively small changes were very important. The focus on project activities to bring partners together was effective and created a core group of evaluators, implementers, and patient representatives who worked closely together, solving problems and creating efficient working practices to design and deliver projects. Project-specific innovators, technical and clinical experts and service delivery teams then worked with this core set of partners as required on a project-by-project basis. However, turnover of staff and a time-limited programme required additional support, flexible approach to resource use and knowledge transfer plans to reduce impact of personnel rotation. Ultimately, high staff turnover, secondments and fixed term contracts were reported to negatively affect the project delivery.

Despite the technologies that were tested being considered market-ready, and had often been used in other settings, there was a considerable amount of feedback that was produced. Recommendations for product improvement and development were produced from engagement with the implementation team, front-line staff, service-users and evaluators. As innovators were not aware that their products could be improved or adapted in the ways that were suggested; this would indicate that everyday technology deployment does not always result in useful feedback to innovators. This provision of feedback for improving and adapting digital technology could be considered a crucial function for NHS infrastructure similar to the PPP Test Bed. The unique combination of organisations and functions integrated into a core team seems to be a critical condition for this mechanism to operate.

The national Test Bed programme was initiated to address the following three issues that were identified as pertaining to the implementation of technological innovations.

1. Innovations are not tested in combination with each other or dependent infrastructure
2. There is little evidence of 'real-world' implementation
3. Innovations are simply added, rather than used to re-think service delivery

There is an incompatibility of short time-scales and real-world evaluation when interventions need to be rapidly and collaboratively designed and implemented, prior to consolidation of evaluation designs. This also creates a natural state in which implementation and evaluation will be out of sync. Pressure from the national Test Bed programme for high recruitment numbers resulted in having to seek new opportunities for implementation when difficulties were encountered; rather than pursuing original plans, which led to rapid change

Combinations of innovations need to be carefully considered, and the risks involved with multiplying complexity understood. In this programme, interventions that were successfully implemented and demonstrated sustainability were single innovations, which could articulate a response to a well-understood problem and were introduced alongside pre-existing services that were stable and well-understood.

2 Introduction

2.1 Purpose and Overview

This report provides content to help inform future decisions on the clinical and economic case for further investment in similar ‘combinatorial innovations’ and/or technology implementation programmes, by reporting on a process evaluation and impact/economic evaluation of the PPP Test Bed. It will contribute to the national Test Bed programme evaluation report (a synthesis of all seven Test Bed evaluations) in September 2018 to share learning, insight and support transparent decision-making. To maximise readability, some technical content is included in appendices. There will also be separate scientific reports published on the ‘White Rose Repository’ (<https://eprints.whiterose.ac.uk/>) to provide access to further details that could not be included due to space constraints.

This report will be used to:

1. Contribute to national findings. Underpin a synthesis report that will be prepared by the National Evaluation Partner. This will highlight key findings, such as:
 - a. the conditions under which combinatorial innovations have been found to be effective in achieving their aims,
 - b. common themes and messages, and
 - c. lessons learned to inform implementers of similar interventions.
2. Inform scale-up & spread. Inform decisions about the potential roll-out or scaling up of existing Test Bed interventions; including the adjustments to patient pathways and staff training that may be needed to implement the interventions.
3. Inform future resources. Inform decisions regarding resources required to support a wider programme of new and innovative Test Bed interventions.
4. Assess use of funding. Demonstrate accountability and value delivered from the public funding invested in the current programme of Test Beds.

The Perfect Patient Pathway (PPP) Test Bed programme began in March 2016 and was initially a 2-year programme to be completed in March 2018. The programme was extended until June 2018 by NHSE to allow further time for testing. The objectives of the programme were:

- Provide an ongoing platform for testing, refining and scaling-up innovations.
- Re-design pathways, bringing combinatorial technologies and system transformations to support holistic and personalised care.
- Embed the culture of transformation and improvement in NHS and other health and care organisations.
- Support co-ordinated decision-making across health and care, informed by real-time data and predictive analytics.

A complex service evaluation (rather than research) approach was chosen for the programme in order to remain flexible and responsive to implementation developments and modifications.

2.2 Research or Evaluation

Owing to the lead-in time for establishing partnerships with innovators and service providers to design the implementations, it was not considered possible to carry out research projects for the PPP Test Bed. Whilst research approaches might have been able to provide better quality evidence, they would not have been feasible within this programme; to remain flexible to real world implementation needs.

Research projects would require long set-up and design periods, NHS/ Health Research Authority (HRA) ethics approval, Governance approvals, Research Passports and Letters of Access to cover each individual element of the programme. Once the interventions were decided with the required level of certainty, this would not be feasible within the tight timeframe of the programme.

As a service evaluation in a real-world setting, evaluation activities were inevitably somewhat constrained. No activities could be undertaken that may have been construed as research. The evaluation used the HRA guidance as the main source to ensure that evaluation activities fell under definitions of evaluation rather than research (<http://www.hra.nhs.uk/documents/2016/06/defining-research.pdf>). Ongoing discussions with the research office at Sheffield Teaching Hospitals (STH) NHS Foundation Trust also ensured that activities were unambiguously defined.

2.3 Ethical approvals

Ethical approval for all elements of this complex service evaluation was provided by the University of Sheffield's School of Health and Related Research (SchARR) research ethics committee (REC). The evaluation protocols, participant information sheets and consent forms were written with the assistance of the PPP Test Bed PMO and the PPP Test Bed Advisory Group (TAG) convened by Healthwatch Sheffield. Amendments to the protocols and participant information materials were submitted and approved by the REC as the evaluation responded to implementation developments.

2.4 Evaluation scope and approach

The evaluation scope was to assess all technology implementation projects and provide a programme-wide evaluation. The Evaluation was formative and summative and used a combination of evaluation methodologies to both take account of and understand the emerging complexity. This approach also allowed the evaluation to isolate activities that could be assessed using quasi-experimental methodologies.

The general evaluation approach was mixed-methods and theory-driven, and included the following three work-streams:

1. Effectiveness (primary and secondary outcomes)
Assessment of key outcomes using trial methodology and comparison data to determine the relative effects of elements of the programme compared to business as usual.
2. Efficiency (economic analysis)
Costs associated with the benefits of the new delivery model & how these compare to the costs and benefits of business as usual.
3. Programme theory (attribution and understanding)
How promising elements of the programme can be scaled-up in such a way as to provide better value and improved patient care. Assessment of the implementation processes. Developing hypotheses regarding assumptions about causality; to explore what works, for whom, in what circumstances and why.

2.4.1 Data collection tools

There were several quantitative and qualitative data collection tools used within the evaluation. The tools used varied between each project. Some tools were 'generic' and used for several projects, such as the patient-reported outcome measures (PROMs) booklet and the participant experience map. However, many projects used project specific tools. A summary of what tools were used in what project is provided within *appendix 1*, alongside a copy of all the tools used. The PROMs booklet and experience map were also piloted within the early stages of the PPP Test Bed programme.

2.4.2 Explanation of the qualitative evaluation approach

The aim the qualitative evaluation was to seek the views and perspectives of key stakeholders in order to identify and explore influences on the programme and individual projects being able to achieve, or not achieve, their stated aims. This approach allows insight into another person's views, perspectives, feelings, beliefs, and thoughts within their own personal, physical, psychological, social, and professional environment.

Each project was delivered by a small team, leading to the often small numbers of potential respondents. The qualitative findings presented in this report are based on the views of the sample interviewed, and may not be transferable to other settings, nor represent the views of stakeholders not interviewed.

However, the key themes identified aim to capture the range of views expressed by those interviewed, that going forward, may be important to consider when launching similar projects.

An experience mapping approach was used to conduct face-to-face and telephone interviews with key stakeholders and service users. This is a practical data collection tool, which facilitates collecting data within a comparative framework. Participants are encouraged to look back and describe their experiences of being involved in the programme. A reflective sense-checking and iterative question development process, during the interview facilitates co-production of interpretation and theory development as part of the interview process. This can help to ensure that interpretation is grounded in participants' perceptions of experiences and begin the theory development process during data collection, which facilitates formative evaluation and rapid feedback of early findings.

2.5 Governance structure

The governance structure over the life of the programme remained relatively unchanged from the planning stage (please see diagrams in appendices). However, there were changes to important operational details and the focus of the day-to-day working that formed naturally from practical difficulties and requirements. The key changes were:

- The increasing centrality of Healthwatch Sheffield and the function of the advisory group (TAG) and project champions, which connected with all governance levels and processes.
- Other advisory groups: Whilst the innovator group did have periods of regular meetings and calls, this proved difficult to coordinate in practice. Apart from the TAG, the use of individual expert advice emerged according to need, rather than formal groups of staff, innovators and technical advisors.
- Project focus of activity: the planned functions of the test bed (platform development, pathway re-design and embedding technology, technology deployment and evaluation) did not operate independently, but coalesced around project activities, and included advice (PPI from the TAG in all cases) and expertise as required. This resulted in developing processes of inter-organisational co-production

The governance and management functions were supported by regular meetings and teleconferences.

- **Project teleconferences.** These took place for every project once every one or two weeks and were forums for responding to emerging issues, operational decision-making, monitoring, risk resolution and escalation of problems. They included PPP Test Bed PMO team members, Healthwatch Sheffield, evaluators, implementation teams, service delivery teams, clinical and technical advisers and innovators.
- **Evaluation Team meetings.** Every two-weeks there was an evaluation meeting (including members of the PPP Test Bed PMO, Healthwatch representative and all available members of the evaluation team). These ensured that decisions were aligned with implementation, evaluation and service-user requirements and perspectives.
- **Programme Planning meetings.** These monthly meetings included representatives from a range of stakeholder groups; evaluators, implementers, innovators, senior sponsors, programme management. These meetings were used for updating stakeholders, signing-off operational decisions and agreeing priorities and next actions.
- **Programme Board meetings.** These quarterly meetings were reporting forums that served to ensure whole-programme oversight and support from senior sponsors.
- **Evaluation Advisory Group meetings.** These meetings included commissioners, evaluation and research experts, representatives from the local authority, NHS Digital, NHSE Test Bed national team, Test Bed national evaluation team etc. They provided oversight, support, advice and strategic direction for the evaluation activities.

Table 1: Meeting frequency for projects

Months	1	2	3	4	5	6	7	8	9	10	11	12
T/C	X4	X4	X4	X4	X4	X4	X4	X4	X4	X4	X4	X4
Eval.	X2	X2	X2	X2	X2	X2	X2	X2	X2	X2	X2	X2
Plann.	X1	X1	X1	X1	X1	X1	X1	X1	X1	X1	X1	X1
Board			1			1			1			1
EAG	1			1			1			1		

2.6 Engagement with the public and service users

A key partner and stakeholder within the PPP Test Bed Sheffield was Healthwatch Sheffield; an independent organisation working in partnership with local people, to ensure that their views are heard by the people making decisions about health and social care. They engaged the public and service users to provide feedback and guidance to the PPP Test Bed about its work and evaluation, working closely with the evaluation team to design participant information sheets, consent forms and evaluation protocols. The team consulted a range of groups and individuals throughout the programme's lifetime, including; The PPP Test Bed Advisory Group (TAG) (which comprised individuals with long term conditions and relatives or carers of people with such conditions), PPP 'Test Bed Champions' (individuals with long term conditions relating to specific projects) and a wide range of community groups. Within this report, they have provided an 'engagement summary' for each of the core projects and the programme wide evaluation.

2.7 Core projects

There were a five core projects included in this PPP Test Bed and formally evaluated by the evaluation team. Some projects underwent transitions, developments and testing in different settings. The core projects were:

1. **Falls Prevention/Strength and Balance:** a falls risk assessment device using sensors and a tablet with bespoke software
2. **Asthma:** an attachment to an asthma medication inhaler which links to software platforms (including a smart phone app) to record medication use
3. **Emergency Care Mobile App (SOS UK):** a healthcare mobile application (app) that records medical information for access in an emergency or routine consultation and can contact carers to alert them in case of emergency or to indicate that the user is safe
4. **Digital Care Home:** using vital signs monitoring linked to National Early Warning Scores (NEWS) and a digital communication platform in care and nursing homes to alert remote clinical teams in case of notable deterioration
5. **Diabetes management:** an attachment to an insulin injection pen that provides information about timing of previous injections

In addition to these technology-specific projects, this report will also cover the PPP Test Bed programme as a whole, with the programme wide evaluation.

2.8 Test Bed 'Plus'

In addition to the core projects, there were three additional PPP Test Bed projects undertaken within the lifetime of the programme that the evaluation team were not commissioned to evaluate.

These were the 3Rings project (evaluated by Sheffield Hallam University), Digital Health Training (Good Things Foundation) project (internally evaluated by the Good Things Foundation) and the predictive analytics project (this was mostly theoretical and had no tangible outputs). Although not formally evaluated by the University of Sheffield team, we received evaluation reports towards the end of the programme, and a summary of these projects is included after the core projects.

3 Falls prevention/Strength and Balance

Table 2: Falls Prevention Project Summary

The project	Falls Prevention/Strength and Balance
Healthcare challenge	
Health care challenge being addressed by the project	The human and economic care costs of falls in the frail elderly population and the need for effective and cost-effective fall-risk assessment and prevention activities
Project rationale	(See <i>appendix 2.1</i> for logic model)
Rational and logic underpinning the project	With minimal training clinicians and health care assistants using the Kinesis QTUG™ device can assess and identify frail people aged 65 and over at risk of falls before they fall, which could enable appropriate clinical decisions and referrals based on the information obtained. This could include early referrals to community falls specialists for intervention, assisting patients to keep well and maintain a healthy independent life.
Intended outcomes	<ul style="list-style-type: none"> • Reduced falls risk • Reduced injurious falls (resulting in reduced mortality & morbidity) • Improved quality of life and wellbeing of people at risk of falling
Project timing (see timeline <i>appendix 3.1</i> for details)	
Dates of project [numbers of participants assessed]	<ul style="list-style-type: none"> • Project set-up/planning – Jun-Dec 2016 • Dec 2016-Project started and first GP practice recruited • Feb 2017 – SchARR ethical approval received Feb 2017 Evaluation data collection started <p>March 2017 – Second GP practice recruited</p> <ul style="list-style-type: none"> • Sept 2017 – Third GP practice recruited • Oct 2017 – participant recruitment finishes • Feb 2018 – Evaluation primary care data collection stops • Feb-Jul 2018 - Data analysis and reporting [20 participants] <p>Two further projects were instigated (with limited evaluation):</p> <ul style="list-style-type: none"> • Dec 2017 – project opens in secondary care balance clinics. This project is continuing beyond the PPP Test Bed [45 participants as of June 2018] • Feb 2018 – project opens in community group settings, closed Apr 2018 [281 participants]
Summary of how the project changed over time	<p>Details of changes</p> <ul style="list-style-type: none"> • The project was initially split into 2 sections, QTUG 1 to evaluate the acceptability of the intervention and the evaluation tools, QTUG 2 to evaluate the intervention. However, recruitment was slow and these projects merged with minor ongoing iteration. • Mar 2017 - Nurses and Health Care Assistants (HCA) from the telehealth team were seconded for assessment and collection of evaluation data. • Mar/Apr 2017 - Vouchers were offered in an attempt to improve recruitment <p>Two additional projects were introduced (Community Strength and Balance Project & Secondary Care Balance Clinic Assessment). Owing to time and resource constraints evaluation was limited to brief descriptions and concise process assessments.</p>
Technology and service delivery model	
The technology deployed	<p>Kinesis QTUG™ (Quantitative Timed Up and Go).</p> <ul style="list-style-type: none"> • The technology uses body worn sensors and a mobile software app including a falls risk questionnaire to assess frailty, mobility, and falls risk. It can be used by non-specialists with minimal training, and is wireless and portable. QTUG uses proprietary algorithms to give an objective assessment of falls risk based upon the ‘Timed up and go’ test. There is a NICE Medtech Innovation Briefing on QTUG https://www.nice.org.uk/advice/mib73

Model of service delivery (including pre-Test Bed pathway)	<ul style="list-style-type: none"> • “The NICE guideline on falls in older people recommends that older people with a history of falls, or who are considered to be at risk of falling, should be observed for balance and gait deficits”.¹ • NICE have listed QTUG in their Falls Pathway (Assessment Section) <p>https://pathways.nice.org.uk/pathways/preventing-falls-in-older-people#content=view-node%3Anodes-assessment</p> <ul style="list-style-type: none"> • Usual care: Patients are generally referred for specialist falls risk assessment and intervention after a fall has occurred. • PPP Test Bed intervention: The aim was to use a ‘moderate’ Electronic Frailty Index (eFI) Score (calculated through GP records) to identify people that might be at risk of falling, to give primary health care professionals the ability to assess falls risk using QTUG at an earlier stage (before a reported fall). These assessments were carried out in GP practices. They could then make appropriate clinical decisions and referrals based on the information obtained. Referrals were made to the Integrated Community Therapy (ICT) team.
Those involved	
Target population	Patients registered with one of three GP practices in Sheffield. Patients were to be aged 65+, identify as ‘moderately frail’ on the eFI and have no clinically reported falls. Patients that are known to have fallen should be automatically referred to (ICT) falls prevention services.
Those involved (stakeholders)	Patients at potential risk of falls, relatives, GPs, Kinesis Health Technologies Ltd., Primary Care Sheffield, Integrated Community Therapy team at Sheffield Teaching Hospitals NHS Foundation Trust, PPP Test Bed PMO, Healthwatch Sheffield, University of Sheffield (SchARR), NHSE.
The evaluation	
Key evaluation themes	<ul style="list-style-type: none"> • Describe the intervention, participant characteristics and falls risk scores • Assess the acceptability and accessibility of QTUG and associated activities to practitioners and individuals at risk of falls • Assess the appropriateness of the eFI and associated process to screen for people at risk of falling • Explore changes to patient pathways and care • Establish if falls risk assessments and referral to a falls prevention intervention reduces individuals’ falls risk • Explore the extent to which technology can identify people at risk of falls and measure their risk after an intervention • Assess whether there are economic benefits for the use of the QTUG technology as part of a falls-prevention care pathway • Evaluate the replicability of the project to support future scale up and spread outside of the scope of the project
Evaluation questions considered but not examined (including reason why not)	<ul style="list-style-type: none"> • Actual reduced falls rates in the sample (limited time to observe events, limited recruitment, ethical and practical difficulties in obtaining comparison group data)

3.1 Engagement report provided by Healthwatch Sheffield

The PPP Test Bed Advisory Group and wider public viewed the project as worthwhile and valuable, with many people expressing their desire to take part. Several community group organisers also thought the project would be worthwhile for their members. The clear focus on prevention was cited as one reason why this project was valued, and a possible benefit was thought to be that people who were assessed for

risk of falls and accepted the intervention on offer would potentially be more confident in going outside the home and being socially active. It was questioned whether all GP practices would have capacity to offer the assessments, and whether some participants may lose confidence and become less active if they had a high falls risk score. However, it was thought that people with low or moderate scores may become more confident and active as a result of being assessed.

The PPP Test Bed Advisory Group and Champion for the project provided suggestions for why recruitment rates might be low and offered suggestions for developments. They felt that use of the word 'falls' was viewed as being likely to provoke fear and anxiety in some potential participants. It was also suggested that some people may not want to take part due to 'being proud' or associating the project with a potential loss of independence. It was suggested that some people may think the assessment could trigger unwanted interference from support services or cause professionals to question their ability to cope in their own home. Furthermore, letters sent to patients, inviting them to take part were thought to have the potential to scare people and deter them from participating. The PPP Test Bed Advisory Group and Champion for the project suggested adopting a more community-based approach, working with the voluntary sector to offer assessments. These recommendations were taken on-board by the implementation team and along with consultations with clinical staff and the evaluation team, led to a change in the wording of the invitation letter for the GP practice-based project and also led to the instigation of the community group-based project.

Although participants had to be aged 65 and over for an accurate score to be produced by the QTUG device (the assessment has only been validated for use in this age group), it was felt that many younger people are also at risk of falls and would benefit from such an assessment.

However, it should be noted that the evidence provided by the evaluation team does not currently support this. The model developed by the evaluation team provides qualification for the assessment of younger people, according to our economic model the assessment of younger people, unless otherwise assumed to have an associated risk factor, may not be cost effective. The problem of validity of assessment in younger people would also need to be investigated through further research, prior to recommending this approach.

3.2 Evaluation Focus

Whilst the device was used in three settings (GP practices, community groups and balance clinics), the evaluation is mostly focused on the GP practice-based intervention. This is because it was the original intervention design, and the other interventions (community groups and balance clinic) were introduced later; when there were limited remaining evaluation resources and time. The later iterations of the project were therefore briefly explored to understand the implementation processes, acceptability, utility and potential effects on the patient pathways and standards of care.

3.3 Methods

3.3.1 Process evaluation methods

See prior section (2.4 Explanation of the qualitative evaluation approach) for an explanation of the experience mapping approach used to conduct face-to-face and telephone interviews with key stakeholders and service users.

Convenience sampling was carried out, whilst all efforts were taken to gain a range of views from the small group of people involved in design and delivery of the project. Ten participants were interviewed; two practice managers, five participants with clinical roles, one innovator and two service users. We used thematic analysis^{2,3} to generate themes and discuss the themes supported by participants' accounts.

3.3.2 Descriptive statistics of patient reported outcome measures

Patients were asked to complete a variety of patient-reported outcome measures (PROMs) to quantify their health status, health-related quality of life, and ability to self-manage their condition/health, at baseline, and then three-month and six-month follow-up post-baseline. These PROMs were the:

- EuroQol Five Dimension with Five Levels questionnaire (EQ-5D-5L);⁴
- EQ-5D-5L Visual Analogue Scale (EQ-5D-5L VAS)⁴
- ICEpop CAPability Measure for Adults (ICECAP-A);⁵
- Recovering Quality of life 10-item (ReQOL-10);⁶
- Patient Activation Measure 13-item (PAM-13);⁷
- Modified Falls Efficacy Scale (MFES);⁸
- Self-reported, Dementia-related Quality Of Life with Utility-index (DEMQOL-U).⁹

If a single score was missing for one domain/item of any of the aforementioned PROMs, the overall summary or index score could not be calculated; therefore, the index or summary score was treated as missing. Demographic information (e.g. age and gender) is provided as descriptive statistics of the patient sample as well as PROM completion rates (e.g. if 'completed', 'not completed', or 'lost to follow-up') and PROM descriptive statistics (means, standard deviations, 95% confidence intervals) for each of the aforementioned measure at baseline, follow-up, and difference between baseline and follow-up. All analysis was conducted using Stata software version 15.¹⁰

3.3.3 Economic evaluation methods

Due to lower than expected recruitment to the PPP Test Bed QTUG evaluation, insufficient data was collected to perform a within-trial analysis. Instead, an economic model has been populated utilising parameter estimates from the empirical literature.

A cohort-based economic modelling approach which incorporated an initial decision-tree followed by a Markov model was chosen to perform an exploratory cost-effectiveness analysis in the form of a cost per quality adjusted life year (QALY) analysis. This is an exploratory economic model which utilises existing evidence from the empirical literature rather than evidence from the study itself (this decision was made post-hoc due to lower than expected recruitment to the study which would limit the generalisability and statistical power of the results obtained from the study, thus providing unreliable results).

The exploratory model is focussed on a falls-prevention care pathway which incorporates three key aspects:

- (Aspect 1) falls-risk assessment using either the QTUG or Timed up and Go (TUG) test when fall-risk screening is implemented within primary care in an 'average' size primary care practice;
- (Aspect 2) falls-prevention intervention in relation to one of four types:
 1. Sheffield Integrated Community Therapy (ICT) team (note, this is classified as a "multifactorial intervention, which includes individual risk assessment" as described within Gillespie, Robertson, Gillespie, Sherrington, Gates, Clemson *et al.* ¹¹)
 2. Falls Management Exercise (FaME) as a form of "multiple-component group exercise"
 3. Tai Chi
 4. Home safety Assessment and Modification (HAM)
- (Aspect 3) downstream effects on number of fallers, rate of falls, and subsequent health and social care resource-use and cost implications

A more detailed description of the economic model is available in *appendix 5*. Summary key results from the deterministic and probabilistic results only are provided later on in this report. The decision-tree and Markov model are represented in *Figure 1* and *Figure 2*, respectively. A summary of key parameters (i.e. transition probabilities, cost, and health-state utility values) used in the model related to the decision-tree and five Markov model states are presented in *Table 3 to Table 6*; footnotes for all tables in this section are available in *appendix 4*. Full details of this model, parameters, unit costs, utility values, methods for analysis, and full results (including one-way sensitivity and expected value of perfect information [EVPI] analysis results) can be found in the separately published scientific report

<https://eprints.whiterose.ac.uk/>.

Table 3: Falls prevention cohort size and probability of fall, injurious fall, or transition to long-term care by age group

Cohort age	Cohort size ^a	Prob. Fall ^b	Prob. Minor + Major fall ¹²	Prob. Minor fall ¹²	Prob. Major fall ¹²	Prob. Long-term care post major fall ¹²
65 to 69	425.11	14.44%	2.87%	2.35%	0.52%	0.00%
70 to 74	333.89	18.40%	3.68%	2.76%	0.92%	8.60%
75 to 89	557.35	47.27%	9.46%	5.77%	3.69%	27.40%

Table 4: Falls risk assessment sensitivity, specificity, and costs

Falls risk assessment	Sensitivity (95% CI)	Specificity (95% CI)	Cost per test/person ^c	Set-up cost ^d
QTUG	0.67 (0.53, 0.79) ^{13a}	0.81 (0.63, 0.94) ^{13a}	£10.50	£2806
TUG	0.32 (0.14, 0.57) ^{14b}	0.73 (0.51, 0.88) ^{14b}	£7.50	£24

Table 5: Falls prevention intervention efficacy and costs¹¹

Falls prevention intervention	Risk Ratio (RR) of falling (95% CI) ^{11a}	Rate Ratio (RaR) of falls (95% CI) ^{11a}	Cost per intervention/person
ICT Team ^{11b}	0.93 (0.86, 1.02)	0.76 (0.67, 0.86)	£170.19
FaME ^{11, 15c}	0.85 (0.76, 0.96)	0.71 (0.63, 0.82)	£220.96
Tai Chi ^{11, 15d}	0.71 (0.57, 0.87)	0.72 (0.52, 1.00)	£374.99
HAM ^{11, 15e}	0.88 (0.80, 0.96)	0.81 (0.68, 0.97)	£247.41

Table 6: Falls prevention Utility values and costs associated with Markov model states

Markov model state	Utility (65 year old base) ^{16a}	Cost ^{17, 18b}
Well ^{19c}	0.780 (base)	£0 (per cycle)
Minor fall ^d	0.755	£422 (per fall) ^{12f}
Major fall ^d	0.682	£4,048 (per fall) ^{12f}
Long-term care ^e	0.586	£24,960 (per cycle)
Dead	0.000	£236 (one-off)

Figure 1: Decision tree example – Care pathway versus no intervention

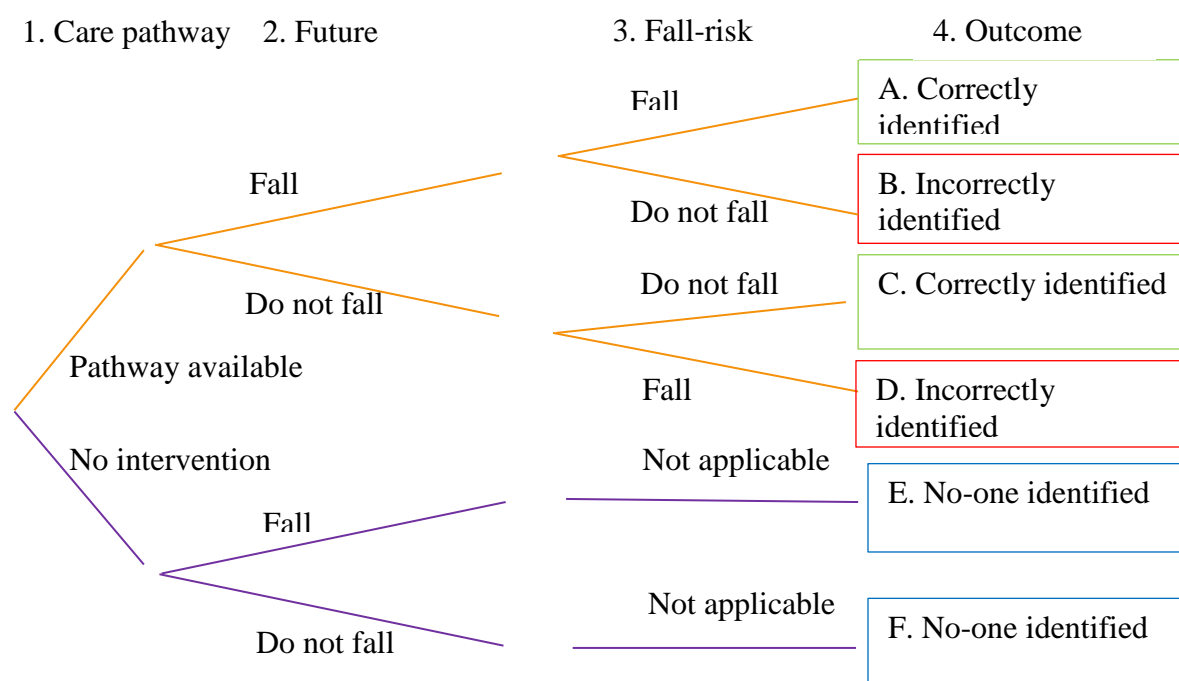
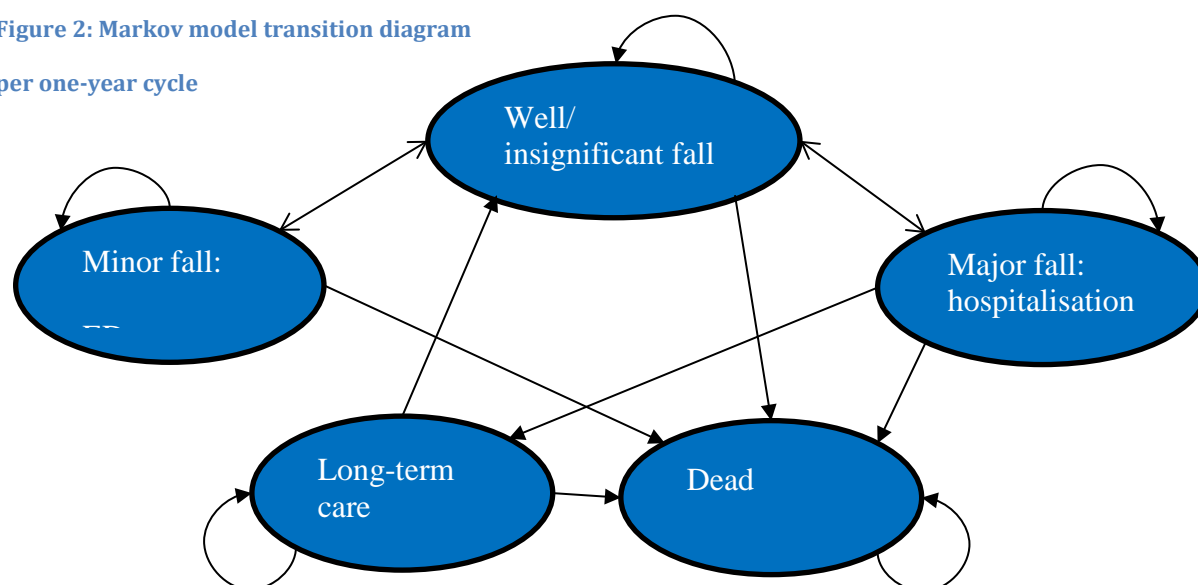


Figure 2: Markov model transition diagram
per one-year cycle



3.4 Key Findings

3.4.1 Process Evaluation Key Findings

3.4.1.1 Design and Set Up

3.4.1.1.1 What was the process to design the project?

Falls and falls prevention were identified as a clinical priority within workshop/engagement events held with clinicians in March 2016. Four months later (July 2016) the PPP PMO had appointed a project manager to oversee a 'falls prevention project'. This project would use the QTUG™ technology to undertake a falls risk assessment using sensors and a tablet with patients from three Sheffield GP practices aged 65 years and over, whose records state they have not fallen but are shown to be 'moderately frail' on the Electronic Frailty Index (eFI), a recognised scale used by GPs. Project planning started in September 2016, it went live in December 2016 (three months later).

Between September and November 2016, the project entered a rapid 3-month design phase. The PPP Test Bed PMO's aim was to get the project up and running in December 2016, immediately after the SchARR REC meeting. The project was up and running at this point, but evaluation data collection did not start until two months later (February 2017) due to the issues noted below.

During the 3-month design phase the project manager worked with GP practices, Healthwatch Sheffield, the TAG, clinicians, the technology innovator (Kinesis) and evaluation team to identify clinical outcomes, evaluation methodology, patient recruitment processes, identify potential participants and agree standard operating procedures. This culminated in the first draft of a project initiation document (PID) being circulated to all involved in October 2016 and a second in November 2016. This was complex and challenging. It took these three months to collaboratively identify project aims, outcomes, practical implementation designs and methods, securing staff to undertake the recruitment, assessments and data collection.

This was done via project conference calls, project meetings and email correspondence. Using these methods, the full team collaborated in an attempt to produce project plans, evaluation protocols and documentation (including project information sheets and consent forms).

The evaluation lag was a result of not being in a position to start drafting the evaluation protocol and associated documentation until the project's aims, outcomes and implementation plans were clear. This clarity did not appear until November/December 2016 because all involved struggled to identify an agreed evaluation methodology and outcome measures. The pace and demands to set up and implement a

project and evaluation in three months was challenging. The project was designed from scratch; documentation, collaborative working relationships, communication and management processes. The speed, workload demands, governance structures within different organisations and novelty of the working relationships impacted on communication and the stakeholders understanding of each other and their needs. Thus misunderstandings about methodologies arose, which were dealt with pragmatically.

Once drafted, the evaluation protocol and associated documents were submitted to ScHARR REC in December 2016. Ethical approval was gained two months later (February 2017) (following requests for amendments which were referred back to stakeholders and addressed collaboratively). A pilot methodology was agreed and subsequently reviewed in an attempt to speed up ethical approval and thus data collection.

3.4.1.1.2 What changes had to be made during implementation to ensure effective delivery of the intervention, and why?

The assessment was fairly simple but the process was lengthy including the completion of a number of questionnaires. It was considered that this might be contributing to low recruitment rates. The implementation process benefited from good communication; the implementation team, evaluation team and Healthwatch Sheffield received regular feedback and updates from the delivery team, allowing rapid, minor improvements to be made. The project had support from GP practices and was responsive and changed to be more adaptive to the needs of the users; including changing the description of the project in the invitation letter and realising that the described length of the assessment sessions could be reduced. Members of staff from the STH Telehealth team became involved in the assessments to offer more assessment clinics at more flexible times, rather than relying on GP practice personnel. It was also decided to improve recruitment rates by offering shopping vouchers as an incentive for participants and to reimburse travel costs.

3.4.1.1.3 Were the governance arrangements for the intervention effective and why?

Governance arrangements for the intervention were effective in managing relationships and specifics of delivery in the GP practices. This was managed through regular pre-set up meetings at the GP practices with members from the implementation team, evaluation team, Healthwatch Sheffield and senior sponsors from the regional healthcare economy and the innovator company. There were weekly teleconferences with members of the implementation team, evaluation team, Healthwatch Sheffield, and Innovator Company. These were used for reporting progress and finding mutually agreeable solutions to emergent problems. Issues requiring further input were taken to other decision-making forums for resolution.

Once the project was running, it was clear that changes were needed. A short pilot phase was useful for these purposes. At an operational level, this required close working between the PPP Test Bed programme personnel; rapidly responding to emergent difficulties and co-producing solutions. For instance, the evaluation team had an organisational mechanism for purchasing shopping vouchers, to promote recruitment, whilst the distribution of these was managed by the delivery team. Changes to project delivery also required consideration of any necessary amendments to the ethical review process, negotiation with GP practice personnel and consultation with the TAG.

3.4.1.2 Partnership

3.4.1.2.1 Did the partnership of the NHS with innovator firms work as intended and why?

The innovators expectations of the project were to have the technology evaluated using a clinical trial to get large scale economic evaluation and then to be deployed on large scale within the NHS. Innovators did not expect that their technology would be tested for practical elements of integration within a care pathway.

3.4.1.2.2 Has the engagement by each party to the partnership been sufficient and why?

The project demonstrated excellent levels of engagement with a range of stakeholders including the PPP Test Bed Programme Management Office (PMO), innovators, community research nurses, health care

assistants, GPs, patients' groups, front-line staff, clinicians and evaluation team. This was largely due to having a dedicated project implementation team to coordinate activities and communication between partners.

3.4.1.3 Implementation and uptake

3.4.1.3.1 What were the barriers to effective delivery (and uptake of technology/ services) and how were they overcome?

The front-line clinical staff members were required to administer questionnaires which placed additional burden on them. Clinicians considered some items on the questionnaires were around sensitive issues (such as suicidal thoughts) with no measures in place to refer patients to if needed; this mental health referral route then needed to be accommodated within the pathway. The way in which the project was initially described to patients was not considered suitable, and was changed. They appreciated terms such as balance, mobility and strength more than falls or falls prevention as some have not experienced falls.

The evaluation team was supported by the team through communication, receiving regular feedback and updates. The project also changed to be more adaptive to the needs of the users by offering vouchers to participants as an incentive to increase recruitment rates. Participants were informed that the session would last 60 minutes. However, following experience of delivering the sessions, to understand that they could be delivered more quickly than this and developing more streamlined processes, the length of the session was reduced to 30 minutes.

3.4.1.3.2 Were there any unintended consequences that needed to be managed and how was this done?

It became clear that the project was identifying patients that had fallen recently, but this had not been serious enough to require treatment and therefore was not in their GP records. They were considered to be at risk of falling again and were directly referred to the ICT falls prevention service. When the additional project incorporating community group assessments began to assess large numbers of people, it was noted that the ICT falls prevention service had developed a long waiting list, as a response to this new demand.

3.4.1.3.3 To what extent is the intervention likely to be scalable and why?

There are elements of the GP practice-based intervention that are likely to be scalable, beyond the PPP Test Bed phase. There are also limitations of the approach that need to be considered. Without entering a sustainability phase it is difficult to fully assess the effects of changes that would be required. However, the main sustainability issues have been identified and are discussed below.

GP Practices: The key barrier for the sustainability of the GP practice-based intervention would be identifying how a long-term assessment service would be organised. This includes identifying financial flows and organisational incentives and finding a solution that would incorporate an appropriate funding model.

- The use of the eFI worked well to identify people at risk of falls; resulting in one-third of those identified, and subsequently assessed using the QTUG, being judged to be at risk and referred to the ICT fall prevention service. However, there is currently insufficient evidence to demonstrate that this is more efficient than using age alone (e.g. 70+ or 75+).
- Reasons for non-response to invitation could be explored further to understand whether non-responders are likely to be at risk, and what the barriers are to attendance.
- The shopping voucher incentive is unlikely to be sustainable. Whilst the offer of shopping vouchers seemed to help with recruitment, this might not be required if the length of the assessment sessions were reduced (e.g. less questionnaires).
- Flexibility and more choice of clinics for delivering the assessment improved access for respondents. However, if this increased the cost of delivery, it could affect sustainability.
- The combination of in-house admin support to identify and invite participants and an external team to deliver the assessment worked well, and would limit the training required to use the

technology. This would also allow the intervention to be implemented efficiently in small practice groups that would otherwise only identify small numbers of people to invite for assessment.

The community group approach showed potential for sustainability, particularly in terms of the numbers of people identified at risk of falls. However, the consequence of this was that a long waiting list of eight to ten weeks soon developed for the ICT falls prevention service. This indicates that a sustainable assessment service would require improved coordination with prevention and support services and possibly increased capacity or alternative support to manage referrals.

The hospital-based balance clinic is continuing to use the QTUG assessment technology. This approach provided an efficient solution for an existing problem within this service, and also demonstrated unanticipated additional benefits.

3.4.2 Impact evaluation key findings

3.4.2.1 Stakeholder benefits

3.4.2.1.1 Did the NHS get better products or processes as a result of collaboration/testing/learning?

The project provided an opportunity for the NHS to evaluate the use of an innovative technology into falls prevention pathway as an integrated therapy service. The QTUG device is more accurate than a traditional TUG test in detecting falls risk, and is relatively easy to use. It is being integrated into hospital balance clinics.

3.4.2.1.2 What have the benefits to innovation partners been of engaging with the NHS as part of the Test Bed programme?

The innovators expected large economic evaluation leading to their technology to be procured by the NHS. However, the technology has now been trialled effectively in a number of different settings, and we have worked with them to develop an economic model that commissioners could potentially use to define the benefit assumptions around using the technology. The device was also improved slightly, through the addition of Velcro straps to attach the sensors, which were much preferred by users.

3.4.2.1.3 Patient experience: What were the impacts of the intervention on patients' experience?

The key impact in patient experience has been the preventative approach to falls, rather than waiting for someone to have an injurious fall prior to an intervention. There has also been the benefit of awareness raising in the community, and more systematic referrals from the balance clinic.

3.4.2.2 Participant recruitment and assessments

180 patients were invited to attend assessment clinics, and 60 attended. Seven were assessed as high risk, but declined referral. Of the 20 people successfully referred to the ICT falls prevention service, 19 patients were assessed as 'high risk' of falls with the QTUG device, one was assessed as 'moderate risk' of falls but reported a history of falls, so was also referred onto the ICT. Mean age of participants was 77 (range 67-87). Of the 20 who were assessed at baseline, 17 continued on to receive ICT falls prevention intervention.

The first follow-up assessment took place approximately three months after referral. Final follow-up assessments were done on discharge from the ICT falls prevention service.

Figure 3: Recruitment rates by month for QTUG assessment

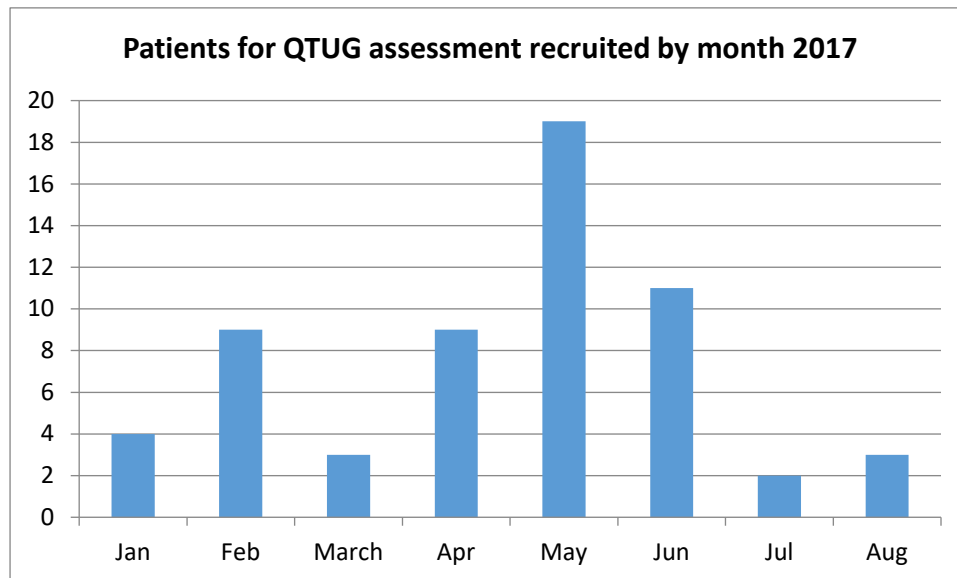
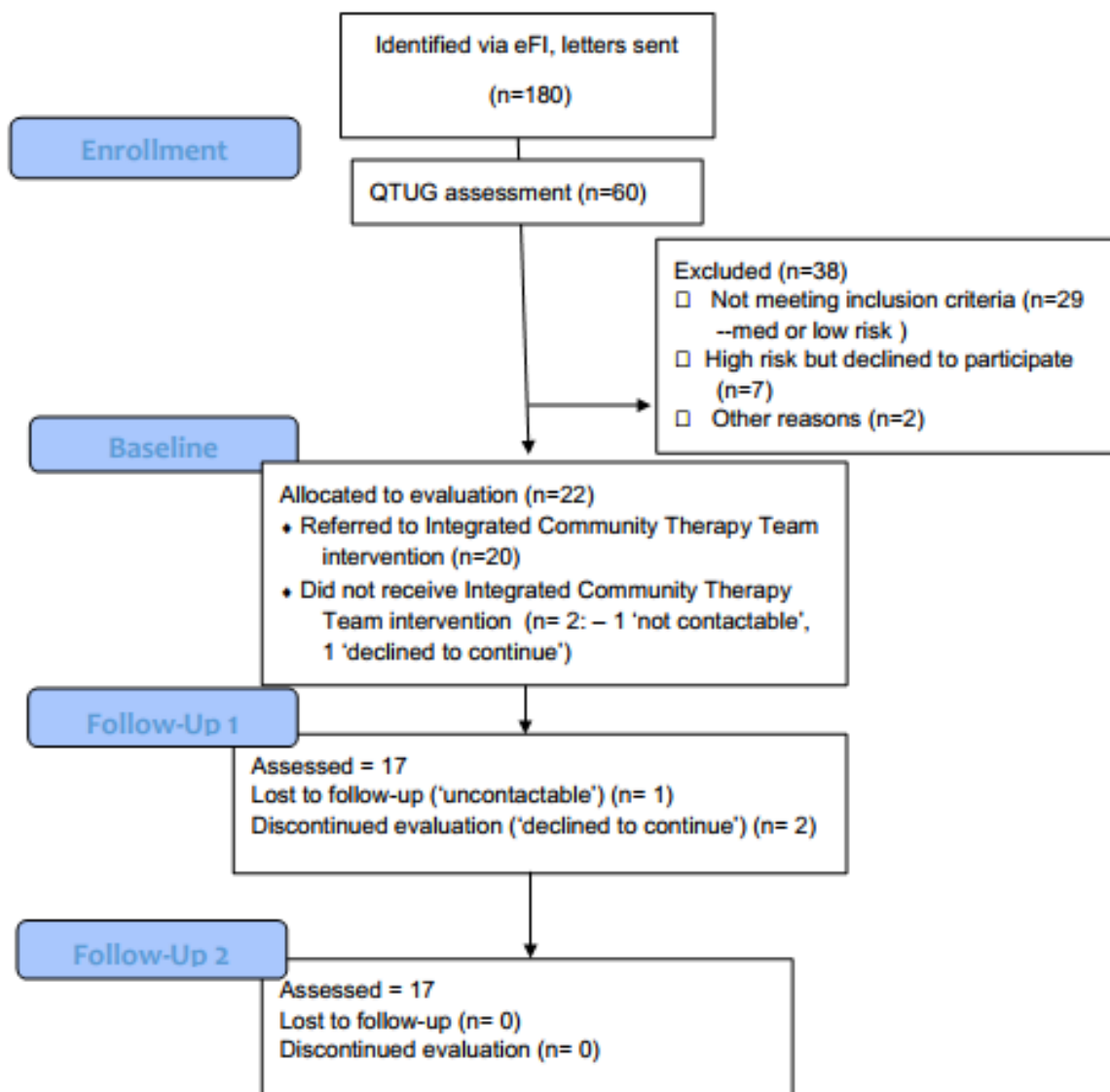


Figure 4: Recruitment Flow Diagram



3.4.2.3 QTUG device questionnaire data

The QTUG device allowed the collection of data for each assessed individual. QTUG falls questionnaire data was available for 12 of the participants at baseline assessment. It also collected falls risk scores and an estimate of frailty.

3.4.2.4 Combined falls risk from the QTUG device

Of the 17 participants who continued past the baseline evaluation, baseline and final assessment QTUG technical data are available for ten. There are some discrepancies in the biometric data for some of the subjects e.g. height & weight, gender, which raises questions regarding the accuracy of the data. If height is being measured at each appointment, this might fluctuate owing to measurement error, weight would also be expected to fluctuate within reasonable expectations.

Of the ten subjects for whom we have baseline and 6-month follow-up data:

- Five showed a decrease in falls risk estimate (range -30.76 to -1.13 percentage points).
- Five showed an increase in falls risk estimate (range 0.3 to 14.7 percentage points)

3.4.2.4.1 Frailty Estimate from QTUG assessment

3 participants showed a decrease in their frailty estimate (i.e. less frail), 7 showed an increase (i.e. more frail).

3.4.2.5 PROMs results

The PROMs results are reported on available data. However, these should be treated as descriptive only, as challenges in recruitment led to very small numbers of complete data sets.

PROM (EQ-5D-5L, EQ-5D-5L VAS, ICECAP-A, ReQoL-10, PAM-13, DEMQOL-U, and MFES) response and completion rates for the PPP QTUG study are shown in *Table 7*.

A total of 20 people were enrolled onto the PPP Test Bed falls prevention study. Baseline demographics suggest that there were 2 (11%) males and 8 (44%) females in the sample (data on gender was obtained for only ten people). There were 7 (39%) patients aged 65-79, and 6 (33%) patients aged 80-89 years (age data was obtained for 13 people). The majority of patients were not living on their own (10 patients) and only three people were living in a residential/care home setting (living status data was obtained for 13 people).

The PROM with the highest completion rate was the EQ-5D-5L VAS, PAM-13 and MFES. The EQ-5D-5L VAS was completed by 18 (90%) people at baseline, 16 (80%) people at 3-month follow-up, 17 (85%) people at 6-month follow-up, and 14 (70%) people at all time-points. The PAM-13 was completed by 16 (80%) people at baseline, 16 (80%) people at 3-month follow-up, 17 (85%) people at 6-month follow-up, and 12 (60%) people at all time-points. The MFES was completed by 17 (85%) people at baseline, 16 (80%) people at 3-month follow-up, 17 (85%) people at 6-month follow-up, and 12 (60%) people at all time-points. These completion rates were generally lower for all other PROMs; however, it should be noted that there was some confusion with the data collection team if this DEMQOL-U measure needed to be collected as part of the data collection schedule (i.e. if the measure only needed to be asked of those with a diagnosis of dementia or not), which resulted in some of 'not completed' records which explains part of the DEMQOL-U low completion rate. Only two (10%) people completed all PROMs for cross-comparison across all PROMs and time-points (when excluding the DEMQOL-U).

Due to the very small sample size, the PROM scores and score changes overtime could be perceived as a biased and unrepresentative sample of people who did or could use this technology; therefore, no key findings related to health-related quality of life can be obtained from these PROMs. Further descriptive statistics related to the PROMs are provided in the separate scientific report (<https://eprints.whiterose.ac.uk/>); however, these results should be interpreted with caution due to the small sample size.

Table 7: Falls prevention PROMs completed at baseline, three and 6-month follow-up

Measure	N	Baseline		Follow-up (3-months)			Follow-up (6-months)			All ^c
		C, n (%N)	NC ^a , n (%N)	C, n (%N)	NC ^a , n (%N)	LFU ^b , n (%N)	C, n (%N)	NC ^a , n (%N)	LFU ^b , n (%N)	
EQ-5D-5L	2	15	5	15	2	3	16	1	3	9
	0	(75%)	(25%)	(75%)	(10%)	(15%)	(80%)	(5%)	(15%)	(45%)
EQ-5D-5L	2	18	2	16	1	3	17	0	3	14
VAS	0	(90%)	(10%)	(80%)	(5%)	(15%)	(85%)	(0%)	(15%)	(70%)
ICECAP-A	2	11	9	13	4	3	16	1	3	6
	0	(55%)	(45%)	(65%)	(20%)	(15%)	(80%)	(5%)	(15%)	(30%)
ReQoL-10	2	11	9	13	4	3	12	5	3	5
	0	(55%)	(45%)	(65%)	(20%)	(15%)	(60%)	(25%)	(15%)	(25%)
PAM-13	2	16	4	16	1	3	17	0	3	12
	0	(80%)	(20%)	(80%)	(5%)	(15%)	(85%)	(0%)	(15%)	(60%)
DEMQOL-U	2	6	14	2	18	0	0	0	0	0
	0	(30%)	(70%)	(10%)	(90%)	(0%)	(0%)	(0%)	(0%)	(0%)
MFES	2	17	3	16	1	3	17	0	3	12
	0	(85%)	(15%)	(80%)	(5%)	(15%)	(85%)	(0%)	(15%)	(60%)
All PROMs ^d	2	8	12	11	6	3	15	2	3	2
	0	(40%)	(60%)	(55%)	(30%)	(15%)	(75%)	(10%)	(15%)	(10%)

Acronyms. C = Completed; NC = Not Completed^a; LFU = Lost to follow-up^b; N = number enrolled to the study

3.4.2.5.1 Service user outcomes

No quantitative evaluation on the impact of this intervention was possible due to the nature of the intervention and the constraints of the evaluation. The desired impact of the intervention is to prevent falls. However, one inclusion criterion for involvement in the study was that no falls had been recorded previously. Therefore, the only way of evaluating this would be to compare the number of falls between the intervention group and a matched comparison group, and in order to observe the number of falls likely to show a difference this would have required both a large number of participants and a follow-up period of several years. Even if a comparison group could have been identified and consented (which would be difficult given the designation as service evaluation rather than research), the follow-up period and numbers required would be well beyond the scope of the study.

Two service users were interviewed. They had different backgrounds and motivations for taking part in the project. One user was a 67 years old woman with multiple health problems living on her own but receiving support and help from her family with food shopping and cleaning. She has pain in her limbs which prevents her from going out. She agreed to take part in the study because she does have falls sometimes and the falls assessment was a new experience for her. Although she found her participation in the study interesting, the assessment did not have an impact on her general confidence. She has a fall alarm for safety and copes with the situation by not thinking too much about falling.

The other user was an 80 years old man living with seven members of his family in a stressful economic situation. He is dependent on his wife for activities of daily living except personal hygiene. He has multiple health problems, pain and dizziness which make him prone to falls.

There was some inconsistency in his perception about how QTUG affected his confidence. He stated that after the first assessment, he was less confident about going out because he was scared of walking. However, he did not attribute this increased fear to the test itself. Overall, he believed the test had “no effect” on his overall walking confidence and was “not bothered” by having the assessment. The suggested exercises that followed up the initial QTUG assessment were, in his view “not very helpful”. No rails or other equipment was recommended for him. He enjoyed visits to his house every month by the ICT team members during the 6-month intervention period that followed the initial assessment. He was disappointed when they stopped coming. His risk score remained at 84 at both the first and second risk assessment.

3.4.3 Potential cost-effectiveness

All cost-effectiveness results are presented in the separate scientific report and only a summary of key results are now described. However, there are some further details in *appendix 5*. These cost-effectiveness results are based on a cohort model of people in certain age groups when screening for falls-risk is conducted in primary care and based on either:

1. QTUG versus Timed Up and Go (TUG) as alternative forms of screening, both of which are followed by a falls-prevention intervention;
2. QTUG with falls prevention intervention versus no care pathway (i.e. no intervention), although downstream health and social care to aid people post-fall are accounted for in both circumstances.

Therefore, cost-effectiveness in these cases are QTUG versus TUG or no intervention and the QALY and cost-savings are for the whole cohort of people which represents a specified cohort of older people in an average primary care practice over a 2-year period (i.e. time horizon of the model).

Based on the results of the modelling analysis, there is a general suggestion that in all cohorts of people (aged: 65 to 69; 70 to 74; 75 to 89; 65 to 89; 70 to 89) that screening with QTUG *dominates* (produces more QALYs and cost-savings) over screening with TUG irrespective of which falls prevention intervention follows this assessment for those at risk of falling. This result makes logical sense, because although screening with QTUG costs more per person (£10.50 per person with a sunk cost of £2806) compared to screening with the TUG (£7.50 per person with a sunk cost of £24), this additional cost for the improved sensitivity and specificity enables those at risk of falls to be better identified and therefore referred onto a falls prevention intervention to avoid future falls whereby the cost-savings are produced by avoiding downstream health and social care costs. Therefore, if falls-risk screening is perceived as desirable by decision makers or clinicians, then based on this analysis, QTUG should be considered over TUG for this assessment in older people aged 65 to 89 (i.e. the remit of this modelling analysis) assuming QTUG can maintain the higher sensitivity and specificity for fall-risk assessment relative to TUG.

The modelling results focussed on screening with QTUG followed by a falls prevention intervention compared to no intervention when the choice to screen is based on age are more complicated than the results for “QTUG Vs TUG”, as the results and general recommendations based on these results differ in the following five age-based cohorts: 65 to 69; 70 to 74; 75 to 89; 65 to 89; 70 to 89. The exploratory economic model suggests that the falls prevention care pathway has a higher probability of being cost-effective at a specified willingness to pay (WTP; e.g. £20,000 to £30,000 per QALY) threshold when screening and falls prevention interventions are utilised in a population aged 75 to 89 (89 is the upper age limit due to the perceived inappropriateness of using specific fall prevention interventions in people aged 90+; i.e. group exercise and Tai Chi) than in those aged between 65 to 74. In fact, the probability of the care pathway being cost-effective when screening is implemented in those aged 65 to 69 compared to no intervention is almost zero, around about 50% in those aged 70 to 74 (i.e. about the same probability as a coin-flip), and around the high 90% to 100% in those aged 75 to 89.

In those aged 65 to 69, the largest QALY gain was observed when screening with the QTUG followed by Tai Chi (0.14), although the lowest additional cost was observed when screening with the QTUG followed by ICT (£15,557); the lowest ICER was £219,375 per QALY when using QTUG with Tai Chi.

In those aged 70 to 74, the largest QALY gain was observed when screening with the QTUG followed by Tai Chi (0.23), although the only cost-saving result was observed when screening with the QTUG followed by FaME (£1,281); only QTUG followed by FaME produced a result which *dominated* no intervention when accounting for health and social care costs (although only at 56% probability of being cost-effective at a WTP threshold of £30,000 per QALY).

In those aged 75 to 89, the largest QALY gain was observed when screening with the QTUG followed by Tai Chi (1.18), although the highest cost-saving result was observed when screening with the QTUG followed by FaME (£86,211).

It should be noted that although QTUG followed by Tai Chi produced the highest mean QALY gain across analyses, QTUG followed by FaME produced the highest probability of being cost-effective at the given WTP thresholds due to the relative cost-savings and larger uncertainty around Tai Chi's efficacy relative to FaME which is accounted for in the probabilistic analysis but not the deterministic analysis.

Table 8: Cost effectiveness summary

Age group	Largest QALY gain	Lowest additional cost (+) or largest cost-saving (-)	Probability of cost effectiveness at a WTP threshold per QALY of £30,000
65 to 69	QTUG+Tai Chi (0.14)	QTUG+ICT (+£15,557)	Almost zero
70 to 74	QTUG+Tai Chi (0.23)	QTUG+FaME (-£1,281)	Around 50%
75 to 89	QTUG+Tai Chi (1.18)	QTUG+Tai Chi (-£86,211)	High 90% to 100%

It should also be noted that in the deterministic analysis, a care pathway in those aged 70 to 89 produced higher QALY gains and cost-savings than for those in aged 75 to 89 cohort due to the extra people within this cohort for whom a QALY gain and cost-saving could be obtained (e.g. highest QALY gain from QTUG with Tai Chi: 1.41; highest cost-saving from QTUG with FaME: £90,263), but at a lower probability of cost-effectiveness due to the uncertainty of producing a cost-effective result in those aged 70 to 74.

The cost-effectiveness results suggest that the relative cost-effectiveness of the care pathway is dependent on five key factors in tandem:

1. sensitivity and specificity of the falls risk assessment;
2. efficacy of the falls prevention intervention to reduce the risk of falling and rate of falls;
3. the cost of the care pathway (i.e. cost of the falls prevention intervention and falls risk assessment);
4. the rate of falls which require medical attention and has a perceived effect on health-related quality of life which could be avoided (i.e. those falls which require a visit to A&E and, for some, an inpatient admission and perhaps subsequent long-term care), the rate of which increases with age based on the observational figures included in the model;
5. the cost of hospital care and social care.

Therefore, for the care pathway to be cost-effective, there is a need to identify those most likely to have an injurious fall using a falls-risk assessment which has high sensitivity and specificity, in order to correctly refer people to a falls prevention intervention which has a high rate of efficacy, at a reasonable cost which is offset against downstream cost-savings associated with when an older person has an injurious fall.

The above paragraph simplifies the decision problem, as what is a 'high' rate of sensitivity and specificity and what is a 'reasonable cost' are traded off against each other to determine what is cost-effective. That is, a lower rate of efficacy for a falls prevention intervention could be traded-off against a higher rate of efficacy if the prior intervention is much cheaper to implement than the latter intervention (assuming the decision making process is based purely on cost-effectiveness rather than a certain weight being placed on reducing falls themselves as an objective rather than cost-effective outcome).

As this is an exploratory economic model and analysis, all results are indicative and not definitive (i.e. suggests the care pathways with the highest probability of producing a cost-effective outcomes, not the care pathways which will produce a cost-effective outcome) and therefore any decisions made based on these results should be made alongside the rest of the evidence around the use of screening and falls prevention interventions in different cohorts of older people (not just based on age, but also frailty as an example) as well as policy objectives (such as the requirement to make cost-savings in the healthcare systems and avoiding falls or improving quality of life in older people).

3.5 Limitations

The exploratory economic model has a number of limitations despite key parameters being driven by randomised control trial (RCT) data (often considered the 'gold-standard' for evidence-based decision making). A key limitation of any model is that there is a need to utilise data from multiple sources, with varying degrees of quality (e.g. from systematic reviews, to single RCTs, to observational data for what could be described as large or smaller cohorts of people who are heterogeneous). An advantage of the economic modelling analysis is that it can quantify this range of information and the uncertainty around the point-estimates (e.g. taking into account the confidence intervals around the mean point estimate, not just the mean estimate) to provide useful information to decision makers around not only if a care pathway could be cost-effective, but the probability of the care pathway being cost-effective based on the information and parameters used in the model in a timely and relatively cost-effective manner; for example, the cost of running multiple RCTs to assess the use of different falls-risk assessment tools alongside various falls-prevention interventions against each other and against offering no pathway would cost millions of pounds (£) compared to running an exploratory economic model.

The model is limited to parameters included in its design and key assumptions associated with the model design and data available to conduct the analysis. In this regard, a range of sensitivity analyses was run alongside the key results presented in this report which are presented in the separate scientific summary (<https://eprints.whiterose.ac.uk/>). It should also be noted that a key difference between age cohorts is the assumed number of injurious falls per year, as this is an important parameter within the model; although it should be noted that given modern healthcare, number of injurious falls per year might be assumed to decrease rather than increase as older people become healthier than historically, which would make the care pathways less cost-effective rather than more.

As described in the more detailed account of the economic model in *appendix 5*, further research is required to understand the dynamic between the first fall and subsequent falls while controlling for other risk factors associated with falling and injurious falls such as age, frailty, and other related co-morbidities; however, actually identifying such a causal effect would require quite sophisticated statistical analysis in a large cohort of well characterised older people.

3.5.1 Community strength and balance project

Between February 2018 and April 2018 a community strength and balance project, using the Kinesis QTUG, was established, led by Healthwatch Sheffield and the telehealth team.

Why: The PPP Test Bed programme planning group discussed various opportunities to assess recruitment in different settings which aligned with the PPP Test Bed Advisory Group (TAG) suggesting using the QTUG in community groups.

What happened: Healthwatch Sheffield contacted voluntary sector groups (such as lunch clubs and assisted living groups) with members over 65 years, based on their local knowledge and experience. Once the project had started, other groups also made contact with Healthwatch Sheffield after hearing about it. The Telehealth team also made contact with community groups and sheltered housing schemes.

Groups were sent information in advance and were visited by a member of Healthwatch Sheffield staff on a small number of occasions, together with community nurses and a Health Care Assistant (working with the telehealth service). People were supplied with information about the project prior to the visit, a co-ordinator spoke about the project to the group and information flyers were handed out so they had time to consider/discuss this with friends/family members prior to attending the session. Other groups also contacted Healthwatch Sheffield to become involved, once they had heard about the project.

Members of the telehealth team then visited the group to offer the assessment. The QTUG assessment was conducted on those people who gave signed, informed, consent. The QTUG results (high/medium and low risk of falls) were discussed with the person. A referral for a falls prevention service with the Integrated Care Team (ICT) –Therapy was offered via Single Point of Access (SPA) if the person was either:

- identified by the QTUG as at high risk of falling,
- appeared to have needs in the clinical judgement of the community nurse although not identified as high risk by the QTUG score
- reported a fall in the past

Assessments and referrals: 281 people had been seen by the time the project closed in April 2018.

Nine of these were not fully assessed (one under 65 years of age, one equipment failure, one unable to walk three meters, six not carried out due to concerns by nursing staff (these were immediately referred)). The following table provides a breakdown of the completed assessments.

Table 9: Results of Community strength and Balance Assessments

Risk Level	N (%)	Referred to ICT	Refused Referral	Successful Referrals
High	137 (50.4)	137	13	124
Moderate	72 (26.5)	16	2	14
Low	63 (23.2)	1	1	0
Total	272	154		138

Of the 272 assessments conducted 138 (50.7%) were successfully referred to the ICT for the falls prevention intervention.

The following issues were reported by the implementation team:

- Healthcare assistants (HCA) cannot work alone in the community due to the need to undertake risk assessments and make clinical decisions when people assessed are not attending as NHS patients / NHS service where an HCA is usually based, with access to clinical supervision from senior staff
- Availability of assessors prevented work being done 5 days a week as some groups only function on specific days. This prevented some groups from being able to take part. Visits to assisted living schemes (extra care housing) were introduced as the day of the week was less of an issue at these venues.
- WiFi connection at some locations was patchy, preventing contemporaneous data from being uploaded from the QTUG tablet to the database (this was conducted later when Wifi were available, although there were still some difficulties with connections at the nurse team's base). However, the technology is available with a 4G option that was not used on this occasion. This would however require a SIM card and data contract

There were a number of practical difficulties that were encountered, including:

- organising interpretation for people without English as their first language
- large groups expecting to be assessed all at once. However, the team did manage 30 assessments in one hour. In other sessions time slots were introduced to reduce waiting times
- some disagreement with assisted living scheme managers regarding who should be assessed. This was addressed by visits prior to the session with information for participants

The staff involved in project implementation considered the high numbers of recruits and interest in the project may be because:

- The population and groups approached saw a need for this approach
- A group setting means people are affected by their peers' behaviour

It is also likely that taking the assessment to people in the community to sessions that they are already attending removes barriers experienced with travel and setting up additional appointments as is required

for the GP practice approach. However, it should be noted that high referral rates resulted in a waiting time of eight to ten weeks for the ICT intervention.

3.5.2 QTUG secondary care balance clinic description

Duration: (3 months when the information was gathered)

Summary: The QTUG assessment device was introduced into a weekly, hospital-based tertiary balance clinic. Previously, whilst referrals were made to falls prevention services, these were not systematic; there were no formal assessments of falls risk or systematic referral criteria to falls prevention services. Initial feedback indicates that the assessment process is suitable for a busy clinical setting, aids interactions with patients, provides systematic assessment and referral routes and is well received by patients and clinicians.

Usual pathway:

- Referrals into neurotology out patients from GPs and speciality secondary care (e.g. neurology, rheumatology, Neurological Ear, Nose, Throat (ENT), out of area consultants)
- Attend clinic for assessment by specialist otology team
- Full neurotology assessment (no formal questioning/investigation of falls risk/gait)
 - Given advice/self-help and discharge
 - Referred for investigations/follow up (including falls service, but ad hoc) and then review in 6 months back in clinic
 - Referred for customised ENT, physiotherapy and follow up and discharge

PPP Test Bed pathway (same as above but with bold parts added):

- Referrals into neurotology out patients from GPs and speciality secondary care (e.g. neurology, rheumatology, Neurological Ear Nose and Throat (ENT), out of area consultants)
- Attend clinic for assessment by specialist otology team
- Full neurotology assessment plus QTUG assessment on those over 65 (this provided an enriched objective assessment of falls risk and a platform to talk about issues with patient / relative / carer- including reassuring those who's risk is low but who feel they are struggling)
 - Given advice/self-help and discharge
 - Referred for investigations/follow up (**including formalised route to fall service based on QTUG result**) and then review in 6 months back in clinic
 - Referred for customised ENT physiotherapy and follow up and discharge

Results:

The following results were reported by clinicians using the technology within the revised patient pathway:

- The QTUG was tried within weekly balance clinics for three months (about 50 clinics in a calendar year)
- This clinic has seven patients attend approx. – their age range is 29-98 years.
- The QTUG was tried in the balance clinic for three months (a total of 84 people attended the clinic approx.)
- The QTUG was tried on all those over 65 years who attended the clinic (n=45)
- Seven of these (16%) were referred to the falls service (those were people with higher falls risk)
 - More targeted referrals than previous random ones
 - Takes around 5-7 minutes to do the test; 'hence sits well in a busy clinic'.
 - Trialled by two users (Consultant and Clinical Fellow)
 - 'Easy to put on patient' (with the new velcro bands - not the original bandage)
 - 'Easy to set up and use with the software'
 - Reported by clinicians as 'well received by patients - user friendly'
 - 'Data is very well presented for explaining to the patient'

3.6 Conclusions and implications

The exploratory economic modelling analysis suggests that falls risk screening with the QTUG device with referral to a fall prevention intervention has a high probability of being cost-effective compared to offering no falls-prevention care pathway at willingness to pay thresholds commonly used by decision makers (e.g. NICE) when primary-care-based screening is utilised in those aged 75 to 89, and at lower probabilities in those aged 70 to 74. However, factors such as willingness to be screened and referred onto a falls prevention intervention should also be considered in relation to the practical issues within implementing such a care pathway. The exploratory economic model in the base case assumes all older people in the primary care practice are willing to be screened and be referred on if identified as at risk of a fall; in reality, this may not be the case. Indeed, in the primary care project only one-third of those invited were actually assessed (60/180), and two of the twenty-two referred on did not receive the (ICT team) falls prevention service.

The exact number of people who need to be screened for the care pathway to be cost-effective is dependent on a number of contextual factors such as how many people are eligible for screening whereby 'eligibility' should be to target those most likely to have a fall/injurious fall. It should also be noted that the cost-effectiveness of the pathway is highly dependent on the efficacy of the pathway to prevent falls in a target population who have a high rate of injurious falls which could be avoided (e.g. those aged 75 to 89). This requires for the care pathway to be implemented with high fidelity to how it was designed to work and that if the care pathway can't be implemented in such a way to obtain the efficacy of results as was determined in a trial-based setting as described by Gillespie, Robertson ¹¹, this will have implications for the cost-effectiveness of the care pathway. Aspects such as long term planning to maintain fidelity to the original care pathway design and improving efficacy using cost-effective solutions should be considered as part of the project scale up and implementation plans. For an additional discussion of conclusions and implications of the economic modelling please see *appendix 5*.

The spread to other settings appeared to be successful in terms of implementation and numbers of people assessed. Implementation in community groups returned a high number of assessments, perhaps in part because of the informal setting and removal of barriers experienced by the primary care cohort of frail people in scheduling a visit to the GP practice. It is also worth noting that 50% of respondents to the questionnaire in the GP setting reported a fall in the previous 12 months. Feedback from the community project indicates that similar numbers were found. However, the project team reported that the majority of people were not aware that this was an issue of particular concern or that there were services that could assist with reducing their falls risk. Therefore, this community awareness raising role could be valuable.

Implementation in the balance clinic indicated that the technology can be introduced in a straightforward fashion in order to modify and formalise assessment and pathways, where service-users might be at risk of falls. The QTUG assessment technology was very well received by the clinical staff, being both quick and easy to use, systematising assessment and referral procedures and facilitating improved interactions with patients.

The formation of the two follow-on projects resulted from discussions between the implementation team and key stakeholders (e.g. Healthwatch, TAG, community groups and acute service consultants) to scope out opportunities for spread to other settings. This is a good example of how the existence of a team of implementers can raise the profile of technologies that might have potential, identify settings and people that are receptive and facilitate implementation. This flexible support for promoting and supporting emerging demand was a key mechanism of the PPP Test Bed programme.

3.7 Recommendations

A key recommendation is that the use of the QTUG device has the potential to improve assessment rates and target prevention services at people at risk of falls. However, the intervention works to form new pathways across services and therefore (in the current organisational structure) requires a dedicated team that can work across organisations and services to seek opportunities, promote, coordinate and facilitate this screening and assessment.

Although the economic model is exploratory, the findings suggest that the care pathway should be targeted at those aged 70-89 as there is a high probability of the pathway being cost-effective in those aged 75 to 89 and cost-effectiveness is borderline in those aged 70 to 74. The very low probability of the care pathway being cost-effective compared to no intervention in those aged 65 to 69 warrants further consideration as it would suggest that unless the rate of injurious falls increases in this age group this care pathway would continue to have a higher probability of not being cost-effective. However, with improving healthcare it seems reasonable to assume this age group are remaining healthier and less frail than historically, and therefore the rate of injurious falls would be expected to fall rather than rise, thereby further decreasing the cost effectiveness of the care pathway for those aged 65-69. Although, the benefit of avoiding the first fall (compared to subsequent falls) in this younger age group could not be quantified in the modelling analysis and so should form an area for future research and consideration.

The model is stratified by age where rate of falls (not just injurious falls) is perceived to increase with age; however, there has been a move to assess adverse events, such as falls, as a deficit associated with frailty which is in itself associated with age. What this means for screening is that if a person is perceived to be particularly frail at a younger age (e.g. if the eFI suggested a 65-year-old was severely frail) they may be eligible for screening and a fall prevention intervention as there is a higher probability of a severely frail person having an injurious fall than someone defined as 'mild frailty' of the same age (e.g. aged 65). The eFI was used to identify the target population in the primary care setting; however, there was not enough evidence from a large enough sample size to suggest if screening for frailty using the eFI before fall risk assessment was a better method for signposting to the care pathway than screening based purely on age. As frailty screening using the eFI can be done without the patient present, based on GP system coding in the patient records, this may limit the number of people who need to be contacted and brought in for screening which could reduce the overall screening cost (e.g. screening without the patient present and limiting the number of people who need to be assessed with QTUG which itself involves time and associated cost). As more is learnt about the relationship between frailty, age, and adverse events such as falls, this should be an aspect for further research when deciding on a care pathway associated with falls risk screening and falls prevention intervention.

It is interesting to note that 50% of the 12 people completing questionnaires in primary care reported having a fall in the past 12 months, and yet this had not been reported and did not appear on their GP records. As history of past falls is a good predictor of future falls, an alternative or complement to screening maybe to first identify those with a history of falls. This could also be combined with awareness raising and altering perceptions in older people about reporting when they have a fall so that an intervention can be implemented before an injurious and/or more severe fall occurs in the future. However, whilst a previous reported fall is a good predictor of future falls, the previous falls being detected in this population were not serious or injurious to the extent that they were considered worthwhile reporting or receiving care for.

The community group and balance clinic-based projects have shown potential to be effective approaches to identifying people at risk of falls. Further investigations would be required to estimate the cost-effectiveness of these approaches. However, there is already some evidence that the benefits are not necessarily directly related to cost-savings. For instance, the balance clinic staff noted that the QTUG assessment provided a useful interface and opportunity to discuss issues related to falling (e.g. reassuring people that considered themselves at high risk of falling), and awareness raising about the risk of injurious falls was reported as an outcome for the community groups. As already suggested, there is a need to improve reporting of falls history in an older population group (e.g. aged 65+).

4 Asthma

Table 10: Asthma project summary

The project	Asthma
Healthcare challenge	
Health care challenge being addressed by the project	Management of asthma, adherence with treatment regimes
Overview of project	
Project rationale (See <i>appendix 2.2</i> for logic model)	
Rational and logic underpinning the project	<p>The provision of the CareTRx Programme will enhance the self-management of asthma by providing patients and healthcare professionals with quantitative data relating to inhaler usage and access to a patient support programme</p> <p>The CareTRx Programme also included a methodology to cluster patients based on their adherence behaviour allowing for tailored interventions (SMS and emails) as part of the programme journey. The journey also included access to behavioural changes techniques that could be accessed via the CareTRx website.</p>
Intended outcomes, according to programme theory assumptions	<ul style="list-style-type: none"> • Adherence data allowing more effective and informed professional-led management of asthma • Digital coaching encouraging activation and motivation of people with asthma to self-manage their condition • Access to adherence data facilitating shared decision-making between healthcare professionals and patients • Improved patient adherence with treatment regime
Project timing (see timeline <i>appendix 3.2</i> for details)	
Dates of project (including Evaluation dates if different) [numbers of participants]	<ul style="list-style-type: none"> • June – Dec 2016 Project set-up/planning • Dec 2016 Project started with first GP practice recruited to 6-month programme • Mar 2017 – Plans discussed to open in secondary care • Apr 2017 – SchARR Ethical approval gained for evaluation • May - December 2017 – three SchARR REC ethical amendments required to manage project iteration [110 participants: 6-month pathway] • Nov 2017 – shift to 3-month primary care programme [22 participants: 3-month pathway] • Jan 2018 – Secondary care pathway live [4] • Feb 2018 – recruitment ceases all projects • Jun 2018 – all patients unenrolled, data collection completed • Jun/Jul 2018 – analysis and report writing <p>Recruitment rates are important to note:</p> <ul style="list-style-type: none"> • 132 people were recruited from 6,043 people with asthma on GP lists (2.2%) • The market research report states that approximately 4,500 invitations were sent out, which equates to a 2.9% response rate
Summary of how the project changed over time	Details of changes made
	<ul style="list-style-type: none"> • Jan 2017 - Decision not to include Chronic Obstructive Pulmonary Disease (COPD) services and mental health problems for resource reasons • Jan-Mar 2017 – decision to recruit more GP practices • Mar 2017 – decision to include patients from Secondary Care • Nov 2017 - 3-month primary care programme implemented (previously 6-month programme)

	Data were made available to GPs via printed paper copy reports (rather than electronically as intended, owing to practical considerations of managing live data and Information Governance (IG) responsibilities)
Technology and service delivery model	
The technology deployed	<p>The CareTRx programme developed by Teva Pharmaceutical Industries Ltd.:</p> <ul style="list-style-type: none"> • CareTRx Sensor (to record dosing events and lights up to remind user to take medication) and linked Smartphone app (to display inhaler usage history) • CareTRx Clinic • Smartphone mobile application (App): CareTR App • Data analytics dashboard and a data hosting cloud (for users and clinicians to view medication history) • Patient Support Programme (PSP digital coaching) (including email, text messaging and website, linked to behaviour change approaches).
Model of service delivery (including pre-Test Bed pathway)	<ul style="list-style-type: none"> • Usual care: First point of diagnosis is usually self-referral via GP (or could be flagged by other services that patients attend). Patients assessed by GP, treatment started, then ongoing management, monitoring and review consultations. • PPP Test Bed intervention: the use of the CareTRx Programme. The CareTRx intervention aims to supplement the usual care pathway by providing data on inhaler usage and treatment adherence to the patient and the healthcare professional reviewing the patient's medication.
Those involved	
Target population	<ul style="list-style-type: none"> • Patients with a diagnosis of asthma, aged 18 and over and using a preventative inhaler that is compatible with the CareTRx Sensor. • Registered with a Sheffield GP Practice and/or receiving secondary care through Sheffield Teaching Hospitals. • 16 GP practices were involved in patient recruitment
Those involved (stakeholders)	Patients with asthma, carers, GPs, Teva Pharmaceutical Industries Ltd., specialist respiratory services at Sheffield Teaching Hospitals NHS Foundation Trust, PPP Test Bed PMO, Healthwatch Sheffield, University of Sheffield (SchARR), NHSE, Asthma UK
The evaluation	
Evaluation themes	<ul style="list-style-type: none"> • Acceptability and utility (engagement) of the intervention • Implementation process • Description of service-users characteristics and quantitative data
Evaluation questions considered but not examined (including reason why not)	<ul style="list-style-type: none"> • Effectiveness of the intervention compared to business as usual (The intervention was not considered to be designed for sustained delivery in the NHS. There were numerous practical difficulties in obtaining comparative data) • Economic benefits. There were a number of reasons why we could not estimate the economic benefits, for instance: <ul style="list-style-type: none"> ○ The sensor was not available on the open market, and the company had not arrived at a sale price for the technology. Therefore, reliable costs for the technology were not available ○ There were practical problems in establishing counterfactual data (i.e. what would have happened without the intervention). For instance, the recruitment rate was very low (see 'participants' above). This indicates that there were a large number of potential users

	<p>systematically excluded and the eventual recruits were also highly self-selecting (i.e. probably a biased sample)</p> <ul style="list-style-type: none"> ○ Another practical problem in establishing counterfactual data relates to difficulties gaining comparable data about adherence rates. Adherence levels are a key proximal outcome in the chain of programme theory (i.e. adherence is assumed to lead to better health outcomes). However, the device records adherence, and other measures of adherence (not using the device e.g. Medication Possession Rate (MPR)) would not give comparable measures ○ The on-boarding process was experimental, resource intensive and incorporated consenting processes for the PPP Test Bed and the evaluation, as well as trialling questionnaires for the evaluation and the intervention. Therefore, this took longer than would have been expected if the process was refined, and was therefore far costlier in this testing stage than it would be if it was implemented as business as usual. However, the testing (and therefore costing) of a refined process was not feasible within the constraints of the programme.
--	--

4.1 Engagement report provided by Healthwatch Sheffield

The following section was compiled from a report received from Healthwatch Sheffield and includes feedback from the advisory group, project champions and other patient and public engagement activities.

Overall, the project was viewed as a good idea because it promoted better self-management and had the potential to lead to better health and fewer asthma attacks for patients. In particular, the education element of the CareTRx programme was valued, as it was widely acknowledged that some patients have issues with knowing how to administer their inhalers effectively. Giving patients reminders to take their inhaler was thought to be particularly useful, as was the provision of ongoing support for patients who may lack the ability or motivation to adhere to a prescribed routine. It was thought that the project would be of benefit to children and young people, and it would be favourable to enrol newly diagnosed patients.

It is worth noting that these types of devices are more commonly used for children and young people. However, this device was not licenced for children.

It was questioned whether the project would attract patients who are non-compliant with their prescribed inhaler routine, even though they might benefit the most from taking part. Linking possible social benefits with an improvement in asthma was thought to be helpful in encouraging non-compliant patients to consider taking part.

Concerns about data security and 'big-brother checking up on you' were thought to be issues that might deter some patients from enrolling onto the programme, whilst shortening the time of the initial consultation clinic and offering flexibility in clinic times to accommodate working patients, was thought to be important in boosting recruitment.

Some people believed the project had the potential to bring wider benefits, such as alleviating pressure on GP resources due to participating patients having better control of their condition. There were concerns that some patients would not be able to participate in the project because they could not afford a smart phone and internet data, or didn't possess the required digital skills to use the app. However, it was

acknowledged that these patients could still benefit indirectly if services were used less by participating patients as their asthma improved.

Views differed on how engaging with the CareTRx programme might influence patients' interactions with healthcare professionals at their GP surgery. It was suggested that it may encourage patients to discuss their condition with their GP, but there was thought to be a risk that patients might not attend their regular asthma reviews if they felt their asthma was more controlled. There were doubts that GPs and nurses would have the time to utilise data produced by the sensor to inform their consultations with patients, and it was felt that there was a need to promote the benefits to GPs, who would need a strong incentive to be involved in the project.

When people handled the sensor they reported that it was quite difficult to attach and remove the device from the inhaler, and some thought that people with limited dexterity in their hands would struggle to carry out this process when swapping the sensor from an empty inhaler to a new one. Ideas to improve the technology were suggested, such as making the sensor beep as well as flash to remind users to take their inhaler, adding a dose counter to overcome the issue of not knowing how many doses are left in inhalers, and provision within the app to record who administered the inhaler, as some patients might rely on carers for assistance.

4.2 Evaluation Focus

Qualitative: The innovators engaged a market research company to carry out interviews with clinicians and service-users, which mainly focused on acceptability and usability. To minimise respondent burden, it was mostly decided to mostly rely on the market research reports for qualitative elements of the evaluation. The evaluation team had the opportunity to contribute to designing interview schedules.

Effectiveness: It was not feasible to collect comparative data to assess the effectiveness or cost effectiveness of the intervention. There were several reasons for this:

- Owing to time pressures related to the short time-scales of the PPP Test Bed programme and time required to collaboratively design an evaluation approach and gain required ethics and governance permissions, it was agreed that the intervention would begin, prior to finalising the evaluation protocol. During this time the feasibility of conducting cost effectiveness analysis was considered. Implementation was therefore not designed at the outset with consideration of an effectiveness evaluation (e.g. collection of baseline measures).
- The key interest of Teva in the PPP Test Bed was to understand the value of the technology and PSP though understanding the acceptance and engagement by both patients and healthcare professionals, which might then inform value/costs.
- The technology was not available on the open market; due to the relatively early stage of market testing. Therefore, no costs for the intervention were available.
- The intervention package was not considered to be designed in such a way that it would be delivered in the same way if it were taken up by NHS services. There were a number of challenges to designing an iteration of the project suitable for scaling and adoption.
- The patient burden (completing baseline and follow-up questionnaires, attending an hour-long initial clinic and follow up clinics etc.), was considered to result in high levels of self-selection, which was confirmed by the low response and retention rates.
- Response rates for recruiting comparison cohort, which would not have the incentive of being offered the technology, would be expected to be even lower than the intervention group.
- A key proximal outcome (adherence to medication regimen) was actually measured by the technology. A comparison group representing business as usual would need to have adherence measured using an alternative method, the comparability of which would not be possible to predict. This would be likely to introduce bias or possibly increase adherence as a result of highlighting adherence through the recording process.

The quantitative evaluation is therefore focused on describing service-users' characteristics, including adherence and patient-reported outcome measures (PROMs). Qualitative elements are focused on the

implementation and collaboration process, and also incorporate elements of the market research report that was commissioned by the innovators with input from the evaluation team.

4.3 Methods

4.3.1 Process evaluation methods

Semi-structured face-to-face and telephone interviews were conducted to collect qualitative data from key stakeholders. Interviews were guided by interview schedules. The interview schedule was a practical data collection tool called an experience map, which facilitated collecting data within a comparative framework. Participants were encouraged to look back and describe their experiences of being involved in the programme.

Of the stakeholders involved in the implementation three participants agreed to be interviewed:

- A Teva staff member
- A specialist respiratory nurse
- A GP practice manager. Two other practice managers were contacted but did not respond

Two programme managers also provided data regarding this project within their programme wide interviews. Thematic analysis of evaluator notes of interviews was used to identify key themes, barriers and facilitators to implementation. Evaluator notes were analysed to identify codes, codes were adapted iteratively as each interview was completed and overall themes identified. Codes and themes were then reviewed by two other members of the evaluation team.

4.3.2 Impact methods

Data were gathered by Teva from CareTRx participants and shared securely and anonymously with the evaluation team. For each participant, data included demographic information (age, sex, height), medication schedule (number of planned doses for each medicine), a list of all scheduled doses by day and time, and a list of all actual medication events, including when a scheduled dose was taken, and any unscheduled (extra) doses.

Descriptive statistics were produced for each variable. The main outcome was adherence, calculated as the proportion of scheduled doses within a given period that were recorded as taken. As there was no comparison data, adherence was calculated on a monthly basis (defined as periods of 30 days from the date each participant joined the study), and summarised overall, by demographic groups, and by groups defined by medication schedule. Latent growth curve analysis was used to determine whether there were trends over time in terms of change in adherence.

4.3.3 Economic evaluation methods

4.3.3.1 Economic analysis methods: descriptive statistics of PROMs

Patients were asked to complete a variety of patient-reported outcome measures (PROMs) at baseline, and then either three-month or six-month follow-up, dependent on the programme to which they were enrolled. The PROMs included the EuroQol Five Dimension with five levels measure (EQ-5D-5L), ICEpop CAPability Measure for Adults (ICECAP-A), Recovering Quality of life ten-item version (ReQOL-10), and Patient Activation Measure with 13 items measure (PAM-13).

If a single score was missing for one domain of any of the aforementioned outcome measures, the overall summary or index score could not be calculated; therefore, the index or summary score was treated as missing. Three-sub samples of people were identified for whom specific PROM descriptive statistics are produced:

- those who completed a PROM at *either* baseline or follow-up (i.e. largest overall sample)
- those who completed a PROM at *both* baseline and follow-up
- those who completed *all* PROMs at *both* baseline and follow-up (i.e. smallest overall sample)

Demographic information (e.g. age and gender) is provided for each of the three aforementioned patient samples, as well as PROM completion rates (e.g. if 'completed', 'not completed', or 'lost to follow-up') and PROM descriptive statistics (means, standard deviations, 95% confidence intervals) for each of the aforementioned measure at baseline, follow-up, and difference between baseline and follow-up. All analysis was conducted in STATA version 15.

4.4 Key findings

4.4.1 Process evaluation key findings

4.4.1.1 Design and Set Up

4.4.1.1.1 What was the process to design the project?

Stakeholders involved in the initial design were Teva, the PPP Test Bed PMO, and the GP practices where the intervention was deployed. Healthwatch Sheffield were involved in terms of the proposed project protocol being taken to the PPP Test Bed Advisory Group (TAG) for feedback, and the evaluation team became more involved as time went on.

The CareTRx project was designed to try a combination of five services with people who experience asthma and test out these services' acceptability.

1. CareTRx Sensor (to record dosing events and lights up to remind user to take medication) and linked Smartphone app (to display medication history)
2. CareTRx Clinic
3. Smartphone mobile application (App): CareTRx App
4. Data analytics dashboard and a data hosting cloud (for users and clinicians to view medication history)
5. Patient Support Programme (PSP digital coaching) (including email, text messaging and website, linked to behaviour change approaches).

Teva had previously implemented all but one of the services (the Patient Support Programme) in other geographical locations. It therefore had established routinely used data collection tools and methods for use with them. These same tools and methods were used within the CareTRx project. This was a complex, multi-faceted intervention, which arrived already pre-designed to a large extent, to collect a great deal of information from participants, and included a time-consuming on-boarding process.

Interview accounts from Teva, programme and evaluation team staff suggest that initially, designing the project was time consuming. Discussion between the partners about what could be done was, in part initially due to a lack of alignment about what was wanted by different partners from the PPP Test Bed, and waiting for plans surrounding the concept of a 'command/coordination centre' to become clearer, in order to inform project design. There were also challenges experienced in alignment of information governance (IG) and data sharing across all organisations, particularly regarding the expectation to feed data into the coordination centre and providing data for evaluation purposes.

Whilst the intervention design, implementation plans were agreed and recruitment began in December 2016 and the evaluation protocol was agreed about two months later (February 2017), there were ongoing discussions about what could be achieved in terms of establishing evidence of effectiveness and cost-effectiveness. The intervention/implementation blueprint was designed by Teva in collaboration with the PPP Test Bed PMO which provided details of the project's methods, data collection tools, recruitment and consenting process and project plans. Initially the evaluation proposal only involved secondary data analysis by the evaluation team.

Teva is a large multi-national pharmaceutical company, which has established routes to market for its products, including stages of testing related to producing specific types of evidence. The innovators approached the PPP Test Bed programme with an expectation that they would be testing out a novel

combination of interventions to explore feasibility, acceptability and usability; to inform refinement of the interventions. In interview, the innovators talked about current project aims, such as:

- testing out the CareTRx programme in the real world to understand how patients and health care staff engaged with it,
- testing and adapting the technology in response to problems,
- considering if they were targeting the right people,
- whether the right data was being collected,
- methods of data collection and storage.

However, there was an increasing influence on the development of the project caused by the understanding that clinical and cost effectiveness outcomes were the kind of evidence that NHSE, the funder, wanted and expected. At a later stage, it was agreed to try the implementation of a range of patient-reported outcome measures (PROMs), so that a pre-post analysis of changes to health related quality of life could be assessed, which proved to be challenging.

Introducing additional data collection during the CareTRx clinic was difficult, as these clinics were already very time-consuming. Follow-up telephone calls were used as an alternative approach to collect additional PROMs. The nurses (collecting data) were employed by the innovator company, which meant that there were indirect lines of communication and incompleteness of data could only be identified when data were collated and compared to recruitment figures. The success of changes to data collection or encouragement of data collectors could also only be monitored with time-delays.

The market research commissioned by Teva was completed by a company called 'Customer Faithful' and their report was shared with the evaluation team in June 2018. Elements of the report are used below and identified as market research to distinguish these findings from those of the evaluation team.

4.4.1.1.2 Were the interventions delivered in line with the proposed plans?

Recruitment was slower than anticipated and GP practices soon became exhausted of potential participants; requiring further recruitment of GP practices along with further identification and invitation of potential participants. Some interviewees (market research report) reported that there were technical glitches, especially initially; with the asthma sensors not updating or syncing easily via Bluetooth, sensor battery failures, sensors falling off the inhalers, and false data readings. It was reported that some of these initial problems were resolved. However, the market research reported enduring problems with accuracy of recorded inhaler use, 'often requiring manual input' and reported to result in duplication when the device synced with the app later on. A decision was also made to roll out this programme to secondary care.

4.4.1.1.3 Were the governance arrangements for the intervention effective and why?

Interview accounts suggest that some staff felt governance arrangements lacked flexibility and responsiveness. However, this can be understood as relating to the common difficulties of working collaboratively across large organisations. The challenges of a global company such as Teva and a large NHS organisation such as STH trying to ensure they met their obligations for data protection whilst negotiating data sharing between organisations were reported as very challenging and given as one reason why the project took so long to implement. However, once the project was running, and NHS patient adherence and usage data and patient reported outcome data was shared between the NHS, Teva and the evaluation team, governance arrangements appeared effective. Data was anonymised and shared electronically through sharing access rights to a secure server.

4.4.1.2 Partnership

4.4.1.2.1 Has the engagement by each party to the partnership been sufficient and why?

It is well accepted that collaboration on inter-organisational programmes can be challenging, in terms of alignment of organisational objectives, incentives and processes. Whilst all staff talked about the challenges of partnership between the PPP Test Bed PMO, Teva, and the evaluation team, none appeared

to feel engagement was an issue. All parties were committed to the project and worked together to be flexible and to overcome challenges.

As would be expected with the delivery of a novel, complex intervention, some variability of engagement of GP practices with the project was reported. Some areas where this could be addressed are; enhanced promotion to GP personnel of benefits for patients and ensuring administrative staff members are fully informed about the project and have a protocol for engaging with patient enquiries.

4.4.1.3 Implementation

4.4.1.3.1 What were the barriers to effective delivery (and uptake of technology/ services) and how were they overcome?

Recruitment: All interviewees talked about the challenges of meeting the agreed recruitment target. eventually 132 were recruited from 6,043 patients with asthma on GP lists. It is worth noting that 46 patients attending the CareTRx Clinic could not be enrolled in the programme, due to several reasons (e.g. did not bring their inhalers, did not have a suitable phone). As noted above, GP practice engagement was reported as variable.

The amount of set up and lead time to set up the CareTRx Clinics in each practice, is a reflection of the demand on the front line to accommodate innovation.

Recruitment rates are important to note:

- 132 people were recruited from 6,043 people with asthma on GP lists (2.2%)
- The market research report states that approximately 4,500 invitations were sent out, which equates to a 2.9% response rate

Appointment times and length: the practice manager interviewed thought timing and length of appointments were off-putting. They reported that their practice had learnt that evening appointments led to increased uptake as many patients worked, and the length of appointments had led to some patients declining.

Engaging patients who are well: the practice manager interviewed talked about known difficulties engaging those with well controlled asthma, in their usual clinical practice, explaining that these people often do not attend appointments as they may perceive they do not need to.

The technology and adaptations required: The nurse reported that some patients had been put off by the glitches in the technology (reported above).

Reports for GPs: The digital data link with GP systems was not operationalised. Whilst this ambition was discussed at the start of the PPP Test Bed, it was never set as an objective. Providing access to live data and the responsibility that came with this was not fully understood and this was the main determining factor in deciding not to seek achieving this objective. Therefore, a nurse visited monthly to download and print-out paper reports for GPs to access patient data. The visiting nurse required access to a Personal Computer (PC), which was regarded as somewhat inconvenient and not providing data in an easily accessible way.

Turnover of respiratory nurses: the practice manager reported a turnover of the Teva -employed nurses that visited their practice, and felt this did not help smooth running of the project.

The market research report findings stated that 'overall, the multi-stage induction process was a manual and lengthy process. This contrasts to the 'works-straight-out-of-the-box' experience that people have come to expect from consumer technology devices such as smartphones, tablets, Fitbits, etc. Both patients and nurses felt that this encouraged a dynamic where patients had to put effort into making the system work, often with technical glitches along the way.' For example, feedback from patients identified:

- Low levels of participation and engagement with the app
- Lengthy set-up & close down clinics
- Technology glitches early on
- Lack of personalisation and updated content
- Concerns about accuracy of data

Market research feedback sessions with four of the GP Practices that took part in the programme and an additional group session with nurses who ran the induction clinics identified:

- Insufficient data recorded by patients, especially regarding triggers and exacerbations.
- Reports were not easy to interpret. It was felt that the various graphs and graphics provided in the report were not very clear in leading to an actionable output. As a result, HCPs felt that their ability to interpret clear messages from the data reports in a timely and efficient way was not well supported by the programme. They viewed this as a missed opportunity, as they felt that an asthma review following receipt of the patient data could have been a real chance to engage patients further in their awareness and compliance of their asthma treatment.

Market research discussions with nurses identified that:

- Patients were put off by or disliked the content and length of patient reported outcome measures. They were regarded by some as irrelevant or intrusive and that only numbers, not comments, were recorded. Nurses felt this affected engagement and retention and also felt this negatively affected the supportive nature of their role.
- No non-English language materials were available, despite recruiting from practice populations with significant numbers of people who may not read English

Non English materials were considered, yet as this was a project for the purpose of testing and for a limited period a decision was made by Teva to print English materials only.

4.4.1.3.2 What were the facilitators of effective delivery (and uptake of technology/ services) and how were they ensured?

- Positive perceptions of the programme:
 - **Offering more than usual care:** Offering an additional service to patients, and it potentially meeting a need for their patient population, had appealed to the GP practice.
 - **Regarding the programme as having potential to benefit to patients and cost / resource savings:** all interviewees spoke about *the potential* the programme offered to patients and the NHS by improving self-management.
- **Data shared to inform treatment and reviews:** the GP practice manager reported the GP had read the reports and used these to inform the asthma review process in some cases.
- **Iteration of design:** The Teva staff member and nurse talked about how the project had facilitated iteration of design of the technology. The GP practice manager also reported how there had been issues with batteries running out, mobile phones and inhalers not connecting, and the app being upgraded but stated the issues had been resolved.
- **Pragmatic collaborations:** Despite the challenges, a pragmatic approach and regular communication was reported to have enabled project delivery.
- **Costs covered by Teva:** The cost for the nurses (Teva employed nurses used within the project) for patient-facing activities was covered by Teva. No costs for delivering the project were reported to have been experienced by GP practices, which had helped to encourage other local practices to participate.
- **Costs covered by PPP Test Bed:** The GP practice manager reported how helpful it was that the project covered costs such as GP/nurse time and using practice rooms. She reported that without these costs being covered, the practice would not have felt able to participate.

The market research identified that some patients found the programme useful as an adherence reminder.

HCPs regarded the programme concept as:

- supportive of encouraging self-management
- supportive of patient education of their individual condition

4.4.1.3.3 Were there any unintended consequences that needed to be managed and how was this done?

Market research interviews with the nurses suggest that some patients may have had negative experiences of participating in the programme. For example, nurses reported that some patients found the questionnaires during the clinic (which were used to define the type of intervention that would be appropriate) were not perceived as relevant support.

Following the initial clinic, nurses were required to set up a follow-up call the next day, taking a further 30+ minutes, to check that patients were comfortable with the technology, as well as completing another patient-reported outcome measures (PROMs) questionnaire. Nurses felt that patients quickly became disengaged, finding the questions boring and repetitive, also that:

- *Some patients were frustrated* that nurses had nowhere to record their comments as to why they offered particular scores. Nurses themselves felt it didn't reinforce the supportive role they felt they were supposed to offer to patients.
- *Questions that asked about depression and suicidal thoughts* (e.g. "I feel life is worth living"), nurses received some antipathy from patients ("patients really, really did not like those questions"). This point was reinforced by patients during the focus groups.
- *Some of the question language was considered Americanised* e.g. for the question "how much love do you feel you have", some patients found this intrusive about their family life, some misconstruing this as a question about frequency of sexual intimacy, making them feel embarrassed.
- *Perceived 'non-asthma' related questions* embedded the feeling that the project was really not about them as a patient and/or benefiting them. Some refused to respond to all the questions, which caused the system to record "incomplete"

4.4.2 Qualitative findings regarding impact

4.4.2.1 Stakeholder benefits

4.4.2.1.1 Did the NHS get better products or processes as a result of collaboration / testing/ learning?

One practice manager reported that a GP had used the data provided by the CareTRx programme to inform some asthma reviews. Practices also reported the value of a respiratory nurse specialist conducting an asthma review. Findings from the market research commissioned by Teva indicated that healthcare professionals saw potential for the concept.

As one of the earlier projects in the PPP Test Bed, the data flows and Data Protection Impact Assessment (DPIA) generated in collaboration between NHS Information Governance experts and Teva were shared widely across the Test Bed Programme in general.

4.4.2.1.2 What have the benefits to innovation partners been of engaging with the NHS as part of the Test Bed programme?

A Teva staff member reported that engaging with the NHS had enabled iterative development/ design of this combination of complex interventions, and that being able to test them together, out in the real world, with all the challenges that involved, was useful to them as a company. They also experienced valuable learning about how to integrate processes included Information Governance (IG), Data Protection (DP), legal issues, collaboration frameworks, Intellectual Property (IP), and clinical pathways.

4.4.2.1.3 Patient experience: What were the impacts of the intervention on patients' experience?

Market research with patients identified that some patients found the programme useful as an adherence reminder but that overall patients had to put effort into making the system work, often with technical glitches along the way. The market research reported the following patient feedback:

- Low levels of participation and engagement with the app and website
- Lengthy set-up & close down clinics
- Technology glitches early on
- Lack of personalisation and updated content
- Concerns about accuracy of data

4.4.2.2 To what extent is the intervention likely to be scalable and why?

The combination of interventions and the associated delivery model is not likely to be scalable or sustainable in its current format. The data collection burden on participants appeared to be deterring participation, and the extent of involvement of nursing staff indicates that the resource use requires further investigation and refinement. Evidence of benefits is also currently not clear. However, there has been considerable learning from the project and if any of the technologies are required to address a service delivery issue, then this should not require as much development of the details of the implementation.

Scalability in the short term (i.e. without clear evidence of effectiveness) would be dependent on adoption by a service that recognises an intrinsic benefit in elements of the intervention. The theory of the intervention is rational and has evidence-based internal logic, which demonstrates that the intervention is capable of addressing (or at least monitoring) low adherence rates. However, considering the monitoring device on its own, these types of monitoring devices have traditionally been used with children and young people. This provides a different prospect in terms of monitoring and enforcement of adherence than in an adult population, hence the inclusion of behavioural change and motivational elements of the programme.

The relatively high levels of adherence, in this intervention when compared to available literature indicate that there is potentially a positive effect. However, the participants were highly self-selecting and therefore probably represent a biased sample that might be more likely to have an interest in improving their adherence levels.

4.4.3 Evidence of impact

4.4.3.1 Cost effectiveness impact

Without the appropriate study design and data collection methods put in place to inform a formal economic evaluation (i.e. intervention and counterfactual information in regards to efficacy and effectiveness both in terms of resources used and change in health status/health-related quality of life), a study-based cost-effectiveness analysis cannot be performed. In order to inform an economic model, efficacy data on the technology's effect on asthma exacerbations (or other asthma-related outcomes) and the cost of the intervention (which was not available) would be the minimum desirable information in order to inform such a model, assuming all other parameters could be obtained from the empirical literature or other accessible sources (e.g. routine healthcare datasets). A more formal economic evaluation should be undertaken once the complex intervention has reached a more finalised stage of development.

4.4.3.2 PROMs findings

There are a number of challenges with the PROMs data. As discussed earlier, these were mostly introduced after the initial evaluation proposal had been agreed, in order to attempt to improve available evidence of effectiveness. Whilst the EQ-5D-5L was collected from the start as part of Teva's original plan, and for the 6-month intervention 95% were collected at baseline, only 29% were collected at baseline and follow-up. Details of implementation and data flows were time consuming to design and agree between all stakeholders, and involved a significant amount of work for implementation personnel. Whilst data collection improved as the project progressed; baseline data had poor completion rates initially.

- The 6-month data has very few genuine baselines
- There is a large percentage drop-out from initial baseline to final assessment

PROM (EQ-5D-5L, ICECAP-A, ReQoL-10, and PAM-13) response and completion rates for the six-month (first planned study) and three-month (second planned study) programmes are shown in *Table 11*.

A total of 83 people were enrolled onto the six-month programme; however, seven people discontinued at the early stages of the programme and the ICECAP-A, ReQoL-10, and PAM-13 were only included in the data collections schedule at a later date (i.e. 1st June 2017) than the EQ-5D-5L such that 18 people were never asked these measures; therefore, the data for those 76 people who continued with the study are the basis for these descriptive statistics for the EQ-5D-5L, and for those 58 people who joined the study after the 1st June 2017 for the ICECAP-A, ReQoL-10, and PAM-13. The PROM with the highest completion rate was the EQ-5D-5L as this was the PROM that Teva was already planning to collect as part of their own suite of measures. After removing duplicates (i.e. cases where the patient had more than one recorded PROM score entry) and the records where the time point was uncertain (e.g. we could ascertain if the PROM was completed at baseline, six-month follow-up, or an 'intermediate' time-point between baseline and follow-up), the EQ-5D-5L was completed by 72 (95%) people at baseline, 30 (39%) people at six-month follow-up, and 29 (38%) at both time-points. These completion rates were generally lower for all other PROMs. Only three (4%) people completed all PROMs for cross-comparison in the six-month study.

The PROM completion rates for the three-month pathway (the second planned study) were much higher, albeit for much lower overall enrolled patient sample to this study (see

Table 11). Only 5 (36%) people completed all PROMs for cross-comparison in the three-month study.

Some reasons for low completion rates are provided by the market research report, which states:

The final clinic invitation followed a similar process to the initial clinic, with an invitation letter and a follow up call. However, only c. 10 per cent of patients actually attended, with a further 20 per cent completing a postal version of the final clinic questionnaire. Nurses felt that patients were reluctant to go through the same issues with the survey, especially:

- too many questions overall (survey fatigue)
- too much similarity between groups of questions, with patients confused what was different to the last question they were asked to score
- non-asthma related questions were considered irrelevant by some patients and as a result did not really see the point of answering them

Due to the very small completion rate of PROMs between baseline and follow-up within and between PROMs, the PROM scores and score changes overtime represent a potentially biased and unrepresentative sample of people who did or could use this technology; therefore, no key findings related to health-related quality of life can be obtained from these PROMs. Further descriptive statistics related to the PROMs are provided in the separate scientific report; however, these results should be interpreted with caution not only due to the small sample size but complications with how the data was collected and scored which brings into question the reliability of these PROM scores (these factors are described and discussed within the separate scientific report).

Table 11: Asthma project PROM completion rates for six and 3-month programmes

Measure	Measure asked, N	Baseline		Follow-up (6 or 3 months)			Both
		Complete d, n (%N)	Not completed, n (%n) ^a	Complete d, n (%n)	Not completed, n (%n) ^a	Lost to follow-up, n (%n) ^b	Complete, n (%n)
Six-month programme							
(1) EQ-5D-5L	76	72 (95%)	4 (5%)	30 (39%)	0 (0%)	46 (61%)	29 (38%)
(2) ICECAP-A	58 ^c	11 (19%) ^d	47 (81%)	34 (58%)	0 (0%)	24 (42%)	5 (8%)
(3) ReQoL-10	58 ^c	16 (28%) ^d	42 (72%)	37 (64%)	0 (0%)	21(36%)	10 (17%)
(4) PAM-13	58 ^c	18 (31%)	40 (69%)	33 (57%)	0 (0%)	25 (43%)	13 (22%)
PROMs (1-4)	58 ^c	7 (12%)	51 (88%)	25 (43%)	0 (0%)	33 (57%)	3 (5%)
PROMs (2-4)	58 ^c	8 (14%)	50 (86%)	29 (50%)	0 (0%)	29 (50%)	3(5%)
Three-month programme							
EQ-5D-5L	14	10 (71%)	4 (29%)	10 (71%)	2 (14%)	2 (14%)	8 (57%)
ICECAP-A	14	9 (64%)	5 (36%)	10 (71%)	4 (29%)	0 (0%)	9 (64%)
ReQoL-10	14	5 (36%)	9 (64%)	7 (50%)	0 (0%)	7 (50%)	5 (36%)
PAM-13	14	10 (71%)	4 (29%)	10 (71%)	0 (0%)	4 (29%)	9 (64%)
All PROMs	14	5 (36%)	9 (64%)	7 (50%)	0 (0%)	7 (50%)	5 (36%)

Footnotes for this table and the following tables are available in *appendix 4.6*

4.4.4 Descriptive analysis of medication doses

Without the appropriate study design and data collection methods put in place to inform a formal economic evaluation a study-based cost-effectiveness analysis cannot be performed for the intended health technology. Therefore, a descriptive analysis of outcome measures was undertaken.

Data were received from 83 patients (46 women and 37 men). Patients' age ranged from 18 to 77, with a mean of 41.9 (standard deviation 18.1). There were a further ten patients for whom scheduled doses were included in the data, but no actual events (either planned doses that were taken, or unplanned doses) were recorded; another three had no planned doses taken, but did have unplanned doses recorded. These 13 patients are excluded from the analysis. In addition, one patient had recorded Ventolin as the medication being used. As Ventolin is a rescue medication, and would not normally be used for planned doses, we excluded these planned doses from the analysis also.

There were 21,043 scheduled doses across these patients. These scheduled doses were times for reminders to take medication. Of these, 5683 (27.0%) were recorded as taken. If the inhaler sensor had not registered a dose of medication, or the patient had not manually entered a dose of medication in the app prior to the reminder for the next dose, this was recorded as a missed dose. Additionally, there were a further 8441 recorded doses not linked to scheduled doses.

However, for many patients there were long periods at the end of the study where no more doses were taken. It is thought that this probably represents patients dropping out of the study/not using the devices any more, and therefore these scheduled events were removed from further analysis. This makes a substantial difference to the number of scheduled doses, removing 10,898 (51.8%) and leaving just 10,145. This means that the 5683 scheduled doses that were recorded as taken represents an overall adherence level of **56.0%**.

4.4.4.1 Adherence to medication regimen by groups of patients

4.4.4.1.1 Adherence to medication regimen by gender

Adherence was slightly higher among women (61%) than among men (50%).

4.4.4.1.2 Adherence to medication regimen by age group

Adherence was generally higher amongst older patients, although the peak (78%) is for patients in their 50s. It is unclear how much, if any, of this difference may be due to use of the device.

Table 12: Adherence to medication regimen by age group

Age group	Number of patients	Adherence
Under 30	29	48%
30-39	10	48%
40-49	14	48%
50-59	11	78%
60-69	12	66%
70+	7	56%

4.4.4.1.3 Adherence to medication regimen by medication type

There are no appreciable differences in adherence between different medication types. Those medications with only one patient using them cannot be compared meaningfully with the others; all other medications have adherence rates of between 49% and 60%. Even if adjusted for length of time after start of the intervention, these differences are small.

Table 13: Asthma project adherence to medication regimen by medication type

Medication	Number of patients using	Adherence
Serevent	1	82%
Qvar	3	57%
Clenil	48	58%
Fostair	19	60%
Seretide	10	49%
Sirdupla	6	52%

4.4.4.1.4 Adherence to medication regimen by number of doses

There is no meaningful difference between adherence where the scheduled dose is a single dose (53%) and adherence to two doses (56%). There were three patients who had a quadruple dose, and the adherence level to these was 70%.

4.4.4.2 Adherence to medication regimen over time

Whilst we do not know about adherence levels before the project (as patients did not have a device to measure adherence prior to the project), one objective was to see whether the level of adherence changes over time during the project, and whether this differs by groups of patients.

Because each patient started at a different time, “months” are defined here as successive periods of 30 days after that patient joined. Overall adherence by month is as follows:

Table 14: Asthma project adherence to medication regimen over time

Month	Total number of scheduled doses	Adherence
1	3524	65%
2	2169	54%
3	1706	51%
4	1470	49%
5	866	48%
6	322	54%
> 6	88	50%

Clearly there is a greater level of adherence in month 1, but after that it remains relatively constant. Within month 1, there is a similar effect: on day 1, adherence is as high as 84% (this may include some doses that were set up and taken in the initial clinic), is down to 72% by day 8, then never reaches 70% again.

Examination of adherence over time by group reveals that the significant drop-off in adherence after month 1 is predominantly due to men, with women showing relatively little change over time. There is a slightly sharper drop-off in adherence amongst younger patients than amongst older patients.

4.5 Limitations

An important limitation to this evaluation is the lack of availability of equivalent comparison data, which means that there is no reliable evidence against which to assess efficacy or effectiveness of the combination of interventions. Therefore, the findings are descriptive only, and no substantive findings about impact or cost-effectiveness can be concluded.

4.6 Conclusions and implications

There were many positive perceptions of the project. These included meeting a need for some patients, enhancing self-management, supporting patient education, providing detailed information to GPs, iteration of technology design, and developing relationships and processes of collaborative working.

Difficulties were experienced related to coordinating and agreeing the organisational governance, project planning, data sharing etc. Whilst, the PPP Test Bed provided resources to work through these problems and reach mutually agreeable solutions, IG support was not immediately available from NHSE and Teva provided legal and data protection support. These issues slowed the delivery of the project, and were reported to produce a project which lacked the required flexibility and responsiveness.

Recruitment and retention were lower and also took longer than expected, meaning that the initial targets were not met. This was perceived to be related to the burden placed on participants to provide information, both for the innovator's and evaluator's purposes, and the contact time required to 'onboard' participants. It was also considered difficult to engage patients that didn't see themselves as being unwell to the extent that they required additional support. It is also worth considering that whilst testing to understand and seek the value for patients, it is not possible to provide patients with an assurance of potential benefits to them.

Within the constraints of the evaluation, it was not possible to estimate cost effectiveness or the effectiveness of the intervention compared to business as usual.

4.7 Recommendations

Implementation:

- **Agree aims and expectations with commercial organisations** - the innovators reported that to work on a PPP Test Bed again they would need clearly agreed objectives, and a clear timetable
- **Support NHS partners to negotiate with commercial organisations** for example regarding data protection, data sharing and intellectual property. Some interviewees suggested that a central, legal team, perhaps led by NHSE could have been beneficial, rather than local NHS personnel negotiating, as they were not skilled in this area, and there were different cultures in the NHS and commercial companies
- Consider:
 - to aid recruitment, if aiming to recruit participants of working age, more and flexible appointment times, including evening appointments; group meetings for all staff in a practice was suggested by the respiratory nurse as a way of possibly aiding recruitment so that all staff know how to respond to queries from patients
 - working with linked local GP practices to share available rooms and co-ordinate clinic slots offered, e.g. patients from one practice could be seen at another
 - seek advice from patient and public representatives about the acceptability of length of appointments and outcome measures, pilot and adapt for future projects
- A more formal economic evaluation could be undertaken once the intervention has reached a more finalised stage of development. If this is required, a research-based study design with appropriate data collection methods (e.g. focused on health outcome and downstream care resource-use) and costings (e.g. costing the intervention) would be desirable. However, comparable adherence data would be problematic to gather, without first conducting a study to establish a rigorous method, so the theoretical link between the intervention and outcomes would be difficult to evidence. There are also individual contextual factors that would need to be accounted for, and an understanding of which elements of the intervention might be causing what effects for whom.

Intervention:

The market research identified areas that indicate recommendations to improve the intervention. Some of the main topics are listed below.

- Improve introduction to website content. "...few patients had experienced any of the website element of the programme – some of the lack of engagement with it was simply down to access not being embedded from the start and email introduction being missed."
- Review information. "The information sheet materials were interpreted by patients as being more about supporting asthma device/technology research rather than potential to benefit them personally."
- Translate materials into other common languages
- Consider linking the induction with routine asthma care appointments
- Reduce length of appointments and provide at flexible times. "75 min. initial appointments dissuaded some from taking part – considered by some prospective participants as too long, especially during a work day,"
- Improve reliability of sensor and connectivity with app
- "careful consideration should be given to when the patient receives the programme i.e. after diagnosis, after exacerbation" "clarify the reason for the length of programme and how this is personally aligned to the objectives and needs of the patient."
- Consider offering the ability to opt out of text messages. Some patients saw these as a nuisance.

4.7.1 Relationship between adherence to medication regimen and health outcomes

There are numerous difficulties interpreting the current evidence on the relationship between medication adherence and exacerbations; most importantly there are variations in the methods of

measuring adherence (for instance medication possession rate (MPR)) and measuring/ defining exacerbations or related health service or additional medication use. Additionally, studies include various types of medication.

Adherence levels vary considerably (e.g. between 22% and 63%,²⁰ and higher adherence seems to be associated with higher severity of symptoms. Therefore, those with higher adherence rates might have a higher symptom burden and therefore experience more frequent exacerbations, despite better adherence.

Studies using electronic monitoring devices have tended to focus on children, rather than adult populations. However, a literature review conducted in 2015 concluded that “good adherence was associated with fewer severe asthma exacerbations in high quality studies”.²¹

“The four adult studies of good quality (q=8), reporting MPRs [28, 29, 32, 33] concluded that 25% increased adherence was associated with approximately 10% reduction in severe exacerbations (adjusted hazard ratio 0.89 [28], relative rate 0.75 [29], adjusted odds ratios 0.90 [32] and 0.86 [33])”²¹

Another systematic literature review in 2015 concluded that “[o]verall, 24% of exacerbations and 60% of asthma-related hospitalizations could be attributed to poor adherence”.²⁰ However, establishing a level of ‘adherence’ also varies between studies (e.g. >79%, >49%, >74% (ibid)).

Although there is no way to estimate change or attribute adherence levels to the intervention, it is worth noting that adherence in month-one was 65%, which is higher than Barnes et al (2015) highest levels of adherence in a normal population. Whilst this does drop-off over time, the mean average remains above 50%, which some studies define as above adherent level. Interestingly, there is 78% adherence in 50-59 year olds, which is close to the >79% highest definition of adherence in the Barnes et al (2015) review.²⁰

4.7.2 Suggestions from the market research report:

- Simple ‘how-to’ reminders may have resulted in greater early engagement.
- Tailoring clinic times and materials to working age and ethnic communities and UK audience.
- Any messaging / emails/ website in the future needs to be personalised to the individual use of the patient.
- Some HCPs felt the programme would appeal more to younger patients/students, but they had no direct feedback on this from patients to evidence this.
- Some HCPs suggested using the programme for education for those more recently diagnosed for a specific period of time, in order to achieve a particular purpose, such as appropriate use of maintenance inhaler usage.
- HCPs from all of the practices interviewed made the point that the value of the programme should not be seen as simply creating more data. They felt that creating a lot of data can even make it harder to find useful insight and takes longer to review the information. Instead, data that provides actionable insight, perhaps as exception reporting, would be better – e.g. frequent use of a rescue inhaler, or some kind of compliance summary view.
- HCPs felt that the various graphs and graphics provided in the report were not very clear in leading to an actionable output. As a result, HCPs felt that their ability to interpret clear messages from the data reports in a timely and efficient way was not well supported by the programme.

4.8 Asthma project: Secondary care asthma clinic roll out

Secondary Care adopted the CareTRx Programme as per the previous design and ‘Service Evaluation’ framework; the main difference, apart from the setting was that the CareTRx Clinic was facilitated by an NHS secondary care asthma nurse. It was reported by the innovator contact that an NHS nurse was required in the secondary care context to retain ownership of the process.

The timelines for this element of the project were set by the timelines for delivery of the PPP Test Bed programme, making this artificially constrained since the request came towards the end of the

programme. This element of the project was treated as an opportunity to try implementation in a new setting.

A PPP Test Bed programme manager worked to identify secondary care contact. The Teva representative project manager and Consultant met face to face mostly to set the project up (5-10 meetings approx.). A nurse was then trained (one full day face-to-face).

The discussions focused on contracting; agreeing the aims and objectives of the CareTRx Programme, and roles and responsibilities which were reflected in the contract. Recruitment was limited to a period of 2 weeks due to working within existing timelines (including Christmas, data cut-off date for collation cleaning and analysis for the evaluation, and time needed for training). We are not aware of any targets set for recruitment. However, 4 patients were recruited in total.

Key differences between primary care and secondary care projects:

- Background of nurse
- Contractual issues – new contract needed. Primary care contract was with the Clinical Commissioning Group (CCG) and did not cover secondary care
- IG governance – review of IG arrangements, policies and procedures.

The main outcomes for this element of the project are related to organisational learning about implementation in an acute setting, rather than General Practice. It is worth noting that the innovator contact assumed that the hospital Trust would not have worked with the pharmaceutical company in the way that it did if it was not a part of the PPP Test Bed programme. Therefore, implementation experience would not have been acquired.

5 Emergency Care Mobile app (SOS UK)

Table 15: Emergency Care Mobile app (SOS UK) project summary

The project	Emergency Care Mobile app
Healthcare challenge	
Health care challenge being addressed by the project	It can be difficult to ascertain personal details or appropriate medical history information in the case of an emergency admission to hospital where a patient cannot recall critical health history or is unconscious. Where a person is in possible danger of needing emergency assistance, an emergency contact can easily be messaged with that person's location or informed that they do not need assistance.
Project rationale (See <i>appendix 2.3</i> for logic model)	
Rational and logic underpinning the project	There are five key elements to the programme theory for this intervention. 1) This information can be used by the individual for self-management or 2) as an aide memoire during consultations with clinical staff. 3) A quick read code on the lockscreen of the phone (or separately printed Quick-Read (QR) code) can be read in a medical emergency to provide emergency care services (or bystanders) with health information. 4) The 'Red Button' functionality alerts contacts in an emergency. 5) The 'Green Button' functionality informs contacts that assistance is not required. See <i>appendix 2.3</i> for logic model.
Intended outcomes	<ul style="list-style-type: none"> Improved outcomes from medical emergencies Improved health status (self-management, improved accuracy of information, enhanced clinical consultations) Peace of mind and time released for carers (improved wellbeing, quality of life) Improved independence of users (improved wellbeing, quality of life)
Project timing (see timeline <i>appendix 3.3</i> for details)	
Dates of project (including Evaluation dates if different) [numbers of recruits]	<p>Jun 2016 – May 2017 – design and start-up phase</p> <p>May 2017– App goes live</p> <p>May- November 2017 – ongoing marketing of app via various forums/media</p> <p>Dec 2017 – free download of app closes</p> <p>Jan 2018 – quantitative evaluation data collection complete</p> <p>May 2018 – qualitative evaluation data collection complete</p> <p>Feb-Jun 2018 – evaluation data analysis and reporting</p> <p>[131 recruits]</p>
How the project changed over time	Details of change
	Once the app had been developed and offered to download, the key changes to the project were focused on the marketing and promotion of the app and further development of the online analytics definitions
Technology and service delivery model	
The technology deployed	This project involved the implementation of a phone app (SOS UK), which aims to provide individuals with an easy way to keep self-recorded health information on their smartphone. A quick-read (QR) code allows remote access to these details in an emergency. There is also functionality to alert contacts in an emergency (Red button), or to indicate that the user does not require assistance (Green Button). The app was offered free of charge for a limited period; from May 2017. There was a marketing campaign to raise the profile of the app with specific potential users
Model of service delivery (including pre-Test Bed pathway)	The app was offered on the open market, at first as a free 'trial'. Individuals could download it.

Those involved	
Target population	Anyone in the UK but with marketing focused solely on Sheffield/South Yorkshire region
Those involved (stakeholders)	UK population, Humetrix.com, NHS Trust A&E departments, Sheffield Teaching Hospitals NHS Foundation Trust, PPP Test Bed PMO, Healthwatch Sheffield, University of Sheffield (SchARR), NHSE.
The evaluation	
Evaluation questions (aims/objectives) <ul style="list-style-type: none"> Process evaluation Impact evaluation 	<ul style="list-style-type: none"> Identify early stage usage of the app Determine the utility, engagement with and acceptability of the SOS UK app Explore the effectiveness of app promotion Assess the co-production of the project
Evaluation questions considered but not examined (including reason why not)	<ul style="list-style-type: none"> Qualitative investigation of users' experiences and opinions (there were no respondents that agreed to being contacted by the evaluation team)

5.1 Engagement report provided by Healthwatch Sheffield

The following section was compiled from a report received from Healthwatch Sheffield and includes feedback from the advisory group, project champions and other patient and public engagement activities.

The app was thought to be able to provide peace of mind for people with health conditions and their relatives and carers. Some people who were confident in using smartphone apps thought the app was a good idea, and particularly liked that relatives could be informed in an emergency situation. The app was thought to have the capacity to save hospital staff time as they would have useful information to hand.

It was suggested that reasons for not downloading the app might include fear of using technology and concerns over data security, particularly regarding who might be able to access personal information held in the app with a quick-read (QR) scanner.

Some people questioned the simplicity of using the app and thought it seemed quite complicated, stating that they didn't want to download the app because they were happy with their current arrangements for informing health care professionals of relevant information in an emergency, and some felt that the app lacked the visibility of their tried and trusted methods. Putting the QR code on a wrist bracelet was thought to be a good way of improving visibility of the app to emergency service staff, as it was believed that staff would check an unconscious patient's wrists to establish whether they wear an SOS Talisman.

A Personal Assistant pointed to the value of the app in situations where a carer had been in an accident and couldn't give information about the person they cared for. They also thought it would be useful because they felt that the information they give to medical staff about their client is often disregarded, whereas the information in the app might be more readily accepted.

Some people thought the app could be improved in some ways. It was suggested that if the app displayed the date the information was last entered or confirmed by the user, then this might overcome the possible issue of medical staff not trusting that the information held in the app was up to date. Providing practical support to download the app, publish a profile and helping them initially navigate the app was thought to be a good way of encouraging people to download and use the app.

5.2 Evaluation Focus and methods

All data collection tools can be found in the separate scientific report (<https://eprints.whiterose.ac.uk/>).

The app was offered on the open market, which meant that it was not possible to systematically identify users or to establish a comparison group. We worked with the app developers to insert a message that could be dismissed, ignored or responded to, inviting app-users to take part in evaluation activities. If users responded to the link, they were taken to a website, where they were asked some simple questions about their situation and experience of using the app. They were also given the opportunity to have further involvement in the evaluation by agreeing to a telephone interview.

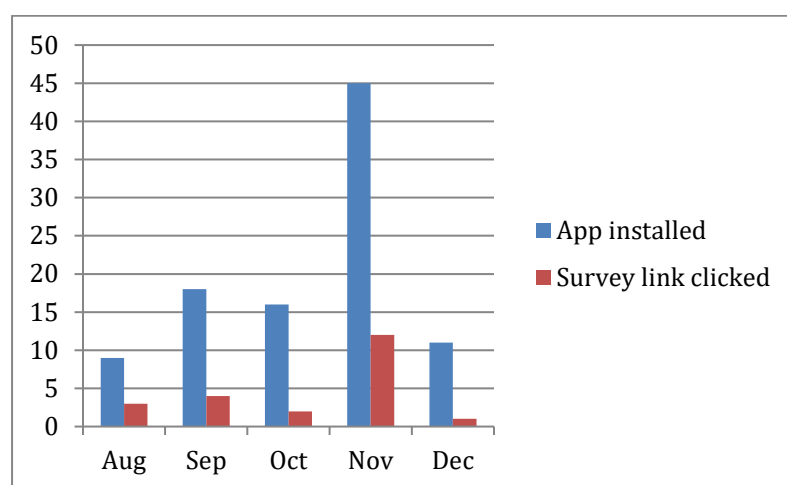
We also had access to the app developer's 'Flurry' analytics login, so that we could explore usage statistics.

5.3 Key findings

5.3.1 User Survey

From May to December (inclusive), 134 people installed the app and 25 clicked the link to the evaluation survey site (19%). Between August and December out of 104 downloads 22 clicked the link to the evaluation survey website (21%). The only completed surveys were filled in during November between the 5th and 23rd November 2017 (5 completed surveys). Two of the respondents found out about the app through their workplace, one from a GP surgery, one from a GP practice website and one from the STH Royal Hallamshire Hospital site (RHH); demonstrating that the marketing activities had a wide reach within Sheffield.

Figure 5: Emergency care mobile app installation and evaluation survey link use



Two respondents were white, two were Asian/Asian British and one did not say. There were two female and three male respondents, and there was a wide range of ages represented; from 18-25 to 65-79 and two respondents were in the 41-64 age group. According to the EQ-5D-5L the respondents were reasonably healthy, two respondents reported no health-related quality of life issues; one had 'slight' pain or discomfort; one other had slight problems with mobility, usual activities and pain/discomfort. However, one respondent reported that they were unable to perform usual activities, were extremely depressed/anxious, had moderate pain/discomfort and slight mobility problems.

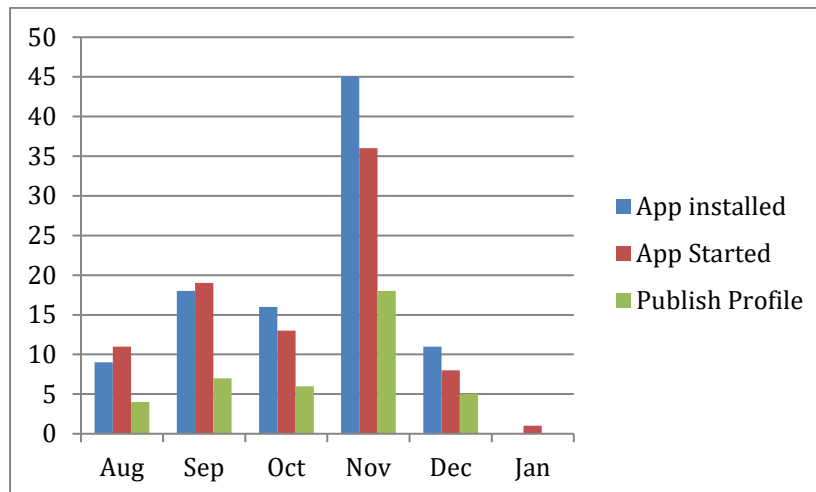
5.3.2 Reasons for downloading the app

The most attractive feature of the app for these respondents was the 'Red Button' for making an emergency call. Four respondents wanted people to be able to access their health details in an emergency (QR code on lockscreen), three wanted to record health details to show a clinician and two wanted to track their own health details (self-recorded details). Only two of the five respondents wanted to use the app for all four of the main functions. One of the respondents only wanted to use the app for the emergency call function.

5.3.3 App usage and retention

Between May 2017 and January 2018 there were a total of 134 installations of the app, 100 instances of the app being started (75%) and 50 published profiles (37%).

Figure 6: Emergency care mobile app usage and retention



The red (emergency contact) buttons on devices were activated a total number of 12 times. The green (no assistance required) buttons were activated a total of 20 times. The QR codes were used to access health records a total number of 37 times. However, there is no way of recognising whether or not these were test activations.

The rolling retention rates for both Android and Apple were around the 22% mark at 3 months. This had dropped to around 9% for them both at 6 months. At 12 months, Apple had dropped to 0% and Android around 4%. This demonstrates that fewer Apple users return to use the app after 6 and 12 months than Android users. The return rates for this app would be expected to be very low. Once details have been entered and profiles published, it is expected that these will need minimal editing and upkeep. Use of the app would otherwise be limited to activation of the red or green buttons or accessing health information by the user or a health professional.

5.3.4 Communications activities

Around 10,000 flyers were sent to individuals on the Carer's Centre register and all Sheffield GP practices in Jun and July, which caused an initial spike in installations of the app. Healthwatch Sheffield also took flyers to various hospital waiting areas, community groups and Patient Participation Groups (PPG's) and mentioned the app during engagement activities in the community. There were emails sent to 71 GP practices in October. There was a further press release (23 newspapers, 14 TV and radio, 15 clinical journals, 18 med tech journals and websites) and email and Facebook activity in October and November, there was a correlating further increase in the rate of app installations during November.

5.3.5 Qualitative process evaluation findings

Semi structured interviews were conducted with NHS implementation staff and one member of Humetrix staff to explore key issues, experiences and perspectives. No service-user interviews were conducted, as no one provided their details and permission to be contacted. However, throughout the project the TAG members tried the app and provided feedback to the innovators through Healthwatch Sheffield.

The lack of funding for innovators through the national Test Bed programme was a significant issue for innovators to manage. This was particularly a barrier to further adaptations that were suggested by user-group feedback; as innovators did not have budget set aside for these development activities.

There were difficulties aligning the innovator and implementation team approaches to promotion, which was a fundamental requisite for the project. Some difficulties were due to limitations and regulations surrounding the promotion of products by the NHS. In particular, it was felt that the content of promotional materials was compromised through the associated governance processes. This project demonstrates a tension between testing and simultaneously promoting the use of technology in the NHS, whilst not being able to guarantee benefits. This is acutely noticeable for interventions that are critically dependent on attracting high numbers of users.

There was also a lack of alignment in the expectations of the innovators and the readiness of the NHS to work with app developers. There was an understandable assumption on the part of the innovators that scale-up and spread would take place as they believed that the Test Bed was promoted nationally as an opportunity to test technology at a meaningful scale. However, as this was ultimately a forum for testing rather than scaling-up, there were no firm plans made for this in the absence of evidence of impact. The evaluators could not recommend a minimum number of downloads to detect a meaningful impact on service use as there was no available evidence upon which to build such an assumption and it was considered that timescales were not long enough to detect any changes. The frustrations of the innovators and misalignment of expectations of partners are clearly demonstrated in the following quote:

“...we committed \$250,000 ...we went in on premise that it would be scaled up, commercialised. There was no plan for this...always talked about never defined”

However, the PPP Test bed PMO were keen to emphasise that the national Test Bed programme was always about testing and evaluating in real world settings. There were never any guarantees to procure. From the perspective of the PPP Test Bed PMO, there was important learning and experience gained from the project.

“...for example, the methods we’ve used, how we’ve worked with tech companies, how we’ve been evaluated, more flexible ways to evaluate” (PPP Test Bed PMO team member, STH Testbed team member from evaluator notes of interview)

5.4 Conclusions and implications (including limitations)

There were some important limitations to the evaluation. Notably, the date set to launch the app and ethics permissions to launch the evaluation website were not synchronised. Therefore, a number of users that clicked through to the evaluation website in the first few days were not able to access the survey. The ‘Flurry’ analytics took some time to finalise metrics and align apple and android telemetry, which makes data before August 2017 unreliable and difficult to interpret.

Problems associated with coordinated timing between evaluation and implementation and readiness of data collection tools are key risks of rapid evaluation. Although these risks can be limited, it is prudent to expect a minimum ‘bedding-in’ period, during which time the quantity and quality of data should be expected to be problematic.

It is not clear why 17 users clicked through to the evaluation website, but did not complete the survey. However, they were required to read lengthy information about the evaluation and how their data would be used and stored, prior to completion, which might have deterred them. We also had one user who reported difficulties connecting with the website using an Apple iPhone, which we were unable to replicate on android devices. There could also have been connectivity and bandwidth problems, preventing access to the evaluation website and associated Google form for the survey.

Estimating usage

It is not possible to determine whether or not recognised activities are simply a result of users testing the functionality of the app or whether they are incidences of actual usage. It would be useful in future for key functionalities of the technology to have a ‘test’ option within the app, which could distinguish it from, for instance actual emergencies.

Linking communication activity with uptake

It is not clear to what extent promotional marketing activities directly relate to additional downloads, or what residual effects are being observed. Indeed, one of the respondents to the evaluation survey in November reported finding out about the app from their GP practice, it is not clear whether this was from flyers distributed towards the end of July, or directly related to the GP emails in November. However, there are noticeable peaks in download activity, which coincide with intensive periods of promotional activities.

Co-production

The project demonstrated some successful approaches to collaborative working between an international innovation organisation, University evaluators and an NHS implementation team. The experience of promoting the SOS UK app and engagement between innovators and the NHS resulted in innovation partners making other application developments that considered intrinsic incentives for front-line clinicians.

Linking the app to an evaluation website developed and owned by a University evaluation team was an innovation in evaluation methods, which showed great promise; over the main study period, 22% of people that installed the app tried to click through to the survey. However, owing to scheduling difficulties and potentially technological issues, only a small number of surveys were completed. This method of evaluation could be worthwhile pursuing in further studies.

Uptake and alternative technology

Ultimately, the project did not achieve sustainability in terms of the intervention being continued in the longer-term, largely owing to the relatively low take-up. In order for the 'emergency access to health records' function to make a difference to the local health services, it would need to have widespread use so that emergency response staff would know to look for the phone and open the data with a QR reader. Although it could be argued that this solution is just one of a number of similar applications and smartphone functions, and emergency staff might increasingly look for accessible information on a person's phone in case of an emergency.

Whilst discussions with emergency hospital staff indicated that they would often look for a person's phone, it is worth noting that android phones have a function to access emergency health details from the lock screen without requiring a QR reader. However, emergency health data available from an android smartphone or iPhone are created from free text entry and are error prone. The SOS UK app includes the entire NHS Dictionary of Medicines and Devices (dm+d) database enabling more accurate selection of actual medicinal products. It also includes a curated list of medical conditions.

It is worth noting that recent iPhone and Samsung Galaxies have SOS call functions, and there are a number of free alternative apps such as 'Shake2Safety', 'SOS – Stay Safe!', 'SOS Emergency App' and 'Shake2Alert'. The unique feature of SOS UK is the combination of the emergency call feature and systematic health record storage and emergency access in one app, which some people might find preferable to using different features or apps for each of these functions. The app also offers the possibility of printing the QR code so that health records could be accessed in an emergency without even having access to the user's phone.

5.5 Recommendations

Importance of clarity of expectations: Agreeing expectations about possible expenditure and development funds: if the technology needs adapting further to meet the identified needs of a particular user group or context, how are these costs going to be met? Innovators need to be aware that real world testing may result in recommendations for technology adaptations to meet the identified needs of a particular user group or context. In such cases, consideration will need to be given to the costs of future development plans.

Scale up and spread: In terms of the scale and spread of apps generally, clarifying evidence required and the process by which decisions about roll out, scale up, spread will be made is important. This could enable technology companies to understand the circumstances under which this may happen, and so to plan and cost for this. E.g. what type of evidence would have been needed to convince the NHS locally to have promoted the app more widely or purchase it? Service users, frontline staff and services need to want the app and see the value for scale up and spread to be possible.

Promotion and communication: Consider clarifying responsibilities between the NHS and innovators for promotion of testing (without endorsement) and promotion of the product respectively. The extent to which an NHS organisation is in a position to support project communication and promotion needs careful consideration if there are future similar projects.

Project/product improvements: It would be useful to agree at early stages how improvements, changes or refinements will be identified and actioned. It would also be useful to gauge the receptiveness of the innovators to invest in change in response to feedback from service providers and service users.

6 Digital Care Home

Table 16: Digital Care Home project summary

The project	Digital Care Home (DCH)
Healthcare challenge	
Health care challenge being addressed by the project	Care home residents have 40-50% more hospital admissions and Accident and Emergency attendances than the general population age 75 and over. It is possible that translating the National Early Warning Scores (NEWS) from acute care situations to care homes and linking alerts to clinical responses, that emergency attendances and non-elective hospital admissions might be reduced.
Project rationale	(See <i>appendix 2.4</i> for logic model)
Rational and logic underpinning the project	This project aimed to enable care home staff to electronically record observations for residents such as temperature and blood pressure and securely communicate that data through the hospital Single Point of Access (SPA). This enabled direct dialogue between SPA and the care home to identify the best course of action (e.g. referral to community nursing or GP) if there were early signs of patient deterioration and so avoid or reduce unplanned emergency hospital admissions or A&E attendances for residents.
Intended outcomes	<ul style="list-style-type: none"> • Reduced number of A&E attendances from the care homes. • Reduced number of emergency (non-elective) hospital admissions from the seven identified care homes. • Reduced number of re-admissions to and from the Sheffield Teaching Hospital's Frailty Unit within 30 days of discharge. • Reduced costs to the emergency care system across the region.
Timing of project (see timeline <i>appendix 3.4</i> for details)	
Dates of project [numbers of recruits]	<p>Jan – June 2017 Planning and discussion</p> <p>June-Sept 2017 -Pilot home monitoring starts</p> <p>Sept-Dec 2017- Recruitment of six additional care homes to project</p> <p>Oct 2017- Jan 2018 Monitoring of care homes ongoing</p> <p>Jan 2018 approved amended evaluation protocol and participant materials (to take into account changes in evaluation methods required)</p> <p>Jan -March 2018 interviews with participating care home staff</p> <p>June-July 2018 Evaluation data analysis and reporting</p> <p>[67 recruits]</p>
Technology and service delivery model	
Technology	Equipment to measure vital signs (e.g. temperature, blood oxygen levels, blood pressure etc) and a tablet app to transmit these readings via a Digital Health Platform to the Single Point of Access (SPA) team. If the readings trigger an alert at SPA, the SPA nurse calls the care home.
Innovative Model of service delivery	Care home staff use the technology to submit NEWS observations and share these electronically with the nurses at SPA based at STH who are able to view readings and alerts through the Digital Care Home portal. If a referral is required following an alert and dialogue with the care home team, the information can be uploaded into patient records via SystemOne or made visible to NHS services via access to the portal.
Those involved	
Target population	Frail elderly living in care homes in Sheffield and their family/carers/friends
Those involved (stakeholders)	Care home residents, carers, care home managers and staff, GPs, Single Point of Access (SPA), Inhealthcare, services at Sheffield Teaching Hospitals NHS Foundation Trust, PPP Test Bed PMO, Healthwatch Sheffield, St Luke's Hospice, University of Sheffield (SchARR), NHSE.

The evaluation	
Evaluation questions (aims/objectives) <ul style="list-style-type: none"> • Process evaluation • Impact evaluation 	<ul style="list-style-type: none"> • Evaluation of the acceptability to care home residents, family members and staff of adding technology to the usual care provided by care homes. • Review of the Single Point of Access (SPA) Digital Care Home clinical support framework. • Assessment of types of alerts and responses
Evaluation questions considered but not examined (including reason why not)	<ul style="list-style-type: none"> • Effectiveness of the project in reducing non-elective and emergency health care service-use (practical difficulties in obtaining comparator data)

6.1 Background

The difficulty of the health care system to respond appropriately to the growing demands of the ageing population is widely recognised.

“A growing frail, elderly population are living with one or multiple long-term conditions. Between 2001 and 2011, the number of people aged 85 or over in England increased at three and a half times the rate of the rest of the population. Older people are far more likely to have immediate or chronic health problems, more likely to need to go to an A&E department and more likely to be admitted into hospital once in A&E.”

(National Audit Office, 2013 p.34 “Emergency admissions to hospital: managing the demand”)

A large proportion of elderly people currently live in care homes. There are approximately 416,000 people living in care homes (Laing and Buisson [survey], 2016). 28,471 emergency admissions from care homes were made to hospitals in England in 2016. This is compared to around 17,539 in 2010, which represents an increase of 62%. Whilst admissions from NHS and local authority run homes have decreased over this time, admissions have more than doubled from other homes (10,510-22,089 (110%)).

(Hospitals: Admissions: Written question – 117811. December 2017

<https://www.parliament.uk/business/publications/written-questions-answers-statements/written-question/Commons/2017-12-06/117811/>)

These increases in the demand for A&E services from care home residents (particularly the private sector), suggest that innovative solutions are required to provide more appropriate and pre-emptive services to this population. Technological solutions facilitate new approaches to care and provide part of this solution to this growing problem. Current evidence suggests that, whilst there are some barriers to overcome, digital innovations could provide appropriate benefits.

A systematic review and meta-ethnographic synthesis²² to examine the experiences of nursing staff using computer-based records in Germany showed that the improvement in quality of residents’ records has a positive impact on quality of care; quicker and clinical and care decision making by providing the staff with automatic alerts. There were mixed views about attitudes toward the technology. The authors suggested a model that works between benefit and burden. Therefore, the application of computer-based records does not automatically lead to positive outcomes. The outcome depends on the implementation process and the staff experience which will change over time based on their roles and responsibilities.

The use of telehealth by patients was shown in a study by McCall and colleagues (2008)²³ where a self-administered electronic questionnaire on a mobile phone collected data on patients’ symptoms with advanced cancer. The data was sent to the hospital where it could be used by clinicians for symptoms management and decision making. The review revealed the role of telehealth in supporting patients, carers and professionals remotely.

A systematic review to examine the barriers and facilitators to the adoption of EHR in long-term facilities²⁴ recognised several barriers including increasing cost, negative user perceptions, implementation challenges, cultural change, to lack of proper training and implementation planning, security, staff retention and technology/system issues. The facilitators included long-term cost savings, error reduction, clinical and administrative efficiency, health outcomes, time saving and access and transfer of information. The initial cost of implementing the EHR and the cost of upgrading the EHR was the most common identified barrier. Many studies stressed the need for better coordination between acute care hospitals and long-term facilities in order to transfer information more effectively and to make improved clinical decisions.

A systematic review exploring the impact of EHR on long-term facilities²⁵ concluded that quality of care was improved either as the direct use of EHR or the EHR improved the inter-professional relationships and integration, hence improving the quality of care indirectly. Health outcomes were improved by reducing the occurrence of infections, high-risk pressure sores, neuroleptics and improving activities of daily living, range of motion and timely medication. Remote access and real-time availability of patient data improved access to patient records.

The Digital Care Home project is a preventative intervention, based on the National Early Warning Score (previously used in acute services, such as frailty units). Regular monitoring of important physiological measures is relayed to an NHS single point of access (SPA). There are set parameters, which if exceeded, trigger alerts that SPA staff respond to; initially by contacting the care home. Advice is given to care home staff and referrals made to other appropriate services if required.

Reflecting the huge variety within the care home sector; a variety of homes were involved in the project, with varying levels of admissions to A&E, different numbers of residents, and different roles supporting the intervention. Importantly, as these homes are early adopters there is largely a culture of wanting to invest in preventative approaches.

6.2 Engagement report provided by Healthwatch Sheffield

The following is a summary from Healthwatch Sheffield as a result of their public and patient consultations:

The potential negative impact of care home residents attending A & E and staying in hospital appears to be well understood and this is one of the main reasons the project has been regarded as worthwhile and beneficial by both the PPP Test Bed Advisory Group (TAG) and wider public. The opportunity to benefit from more timely access to care was greatly valued, along with the potential for increasing peace of mind of residents and their relatives.

Ensuring that residents fully understand the implications of participating in the project has been considered as a challenging aspect of implementation. The TAG has queried the rationale used in selecting residents to take part and whether this should be led by care home managers. It is not clear what alternative methods could be used, perhaps this could be medically led by a GP or geriatrician. However, engagement could be difficult.

Conversations between the PPP Test Bed PMO and care home teams during the recruitment phase indicated that it may have caused more anxiety to residents to be approached by strangers to participate in the project. Equally the PMO had no ability to contact residents within a home other than through working with the manager and team.

The TAG have also highlighted the need to be aware of the potential for measurements being taken to provoke anxiety in some residents, but also acknowledge that many residents will be accustomed to having measurements taken.

The TAG believe that vigilance is needed to ensure that the 'human element' carers bring to residents' care is valued and considered when making decisions about residents' care rather than taking a wholly data driven approach. However, the implementation team are keen to emphasise that decisions about

care are made through dialogue with care home staff, residents and SPA staff. Additionally, they feel it is important that staff enter residents' readings into the digital platform at the point they are taken rather than retrospectively.

Carers have suggested that using a digital platform as a way of sharing health information with health care professionals could lead to more responsive care than is usually possible, particularly as their concerns about a resident's health might not be disputed if confirmed by the data collected. Monitoring weight has been suggested as a measure that should be better monitored within homes. There is a common held view that vital signs monitoring would benefit people living in their own home with care packages.

6.3 Methods

6.3.1 Process evaluation methods

Semi-structured telephone and face-to-face interviews were used to seek perspectives and experiences of the project, of staff, residents and their relatives'. Interviews were guided by interview guides. The interview guide was a practical data collection tool called an experience map, which facilitated collecting data within a comparative framework. Participants were encouraged to look back and describe their experiences of being involved in the programme.

A total of thirty-one participants were interviewed; twenty-one staff and eight residents and two relatives of residents. The University evaluation team interviewed a convenience sample of Twenty-one care home staff and other stakeholders. This consisted of five care home managers, three deputy managers, three nurses, two GPs, two senior palliative care staff (consultant and community nurse), two SPA staff, two innovators and two programme managers. The interviews were recorded using handwritten notes and audio recordings (with consent) to enable accuracy and fact checking.

Residents who were having observations taken, as part of the project, whom care home staff felt would be able to participate in an interview, were invited to participate in a face to face semi-structured interview carried out by Healthwatch Sheffield staff. The topic guide was written collaboratively by the SchARR evaluation team, Healthwatch/PPP Test Bed Advisory Group (TAG) group and PPP Test Bed PMO. Two members of Healthwatch attended each interview, one interviewing whilst the other made notes and observed. Notes were typed up afterwards, discussed and shared with the SchARR evaluation team.

Thematic analysis of evaluator notes of interviews, based on Braun and Clarke (2006)² were used to identify codes, key themes, barriers and facilitators to implementation. Codes and themes were reviewed by two other evaluators.

6.3.2 Impact and economic analysis methods

Intervention resource-use, costs and achieving cost-neutrality

Discussions were held with Inhealthcare, the PPP Test Bed PMO, clinical staff and care homes associated with implementing the intervention to understand the resources and costs associated with implementing the Digital Care Home (DCH) interventions. This included the assessment of the cost of the DCH interventions themselves and any staff, training or additional resources associated with the practical use of the device. The resource-use implications were based on these discussions and assumptions as needed (e.g. it was necessary to make assumptions about the time taken for taking the readings or otherwise in real-world settings where it was not possible to conduct a time-and-motion study). Unit costs for the Digital Care Home interventions themselves were based on those costs suggested by the developers (Inhealthcare).

Two types of costs are considered in this cost estimation of the DCH intervention:

1. **Implementation costs:** this includes the cost of the technology for vital sign monitoring (i.e. pulse oximeter, thermometer, and blood pressure monitor) and data entry (i.e. a tablet and case), the Inhealthcare platform itself (i.e. where data is stored), and monitoring modules for care home residents as part of the Inhealthcare platform (e.g. for monitoring long term conditions or malnutrition)
2. **Operational costs:** this includes staff time at the care home (for monitoring and recording the relevant residents' data) and Single Point of Access (SPA) staff (for responding to alerts due to the monitoring of care home residents).

Although the DCH study ran for 11 months (June 2017 - May 2018), all intervention costs are estimated for one year to provide a standardised estimate of the potential intervention costs for comparison against the statistical analysis conducted for this DCH study (i.e. change in hospital contacts per year). For the purpose of providing cost estimations as generalisable examples, we assume care home sizes can be dichotomized into three broad groups based on the number of residents as suggested by the Care Quality Commission ²⁶; these care home sizes are: 'large' homes (50+ beds); 'medium' homes (11 to 49 beds); and small homes (1 to 10 beds; small homes are not included in these costing examples). For the purpose of these cost estimations, we will base the costing assumptions on those care homes and residents included in the DCH study; that is, seven 'large' homes (seven homes in total), from which 67 residents supplied data which were sent to the Inhealthcare platform for the purpose of one monitoring module for Long Term Conditions (LTCs). Two types of costing are performed for the technology involved in the DCH intervention costing estimations:

1. total sunk costs (i.e. the initial purchase cost of the technology);
2. equivalent annual cost (EAC; i.e. the cost of the technology per year assuming a 3-year capital life).

For the EAC, it is assumed that the capital life of the technology is three years and that an interest rate of 3.5%, which is based on the future discounting rate suggested by NICE ^{27,28} is paid each year as a depreciation rate equal to the cost of maintaining the technology; the annuitization and calculation procedure for estimating the EAC are described by Drummond, Sculpher, Claxton, Stoddart and Torrance ^{29,30}. All additional methods and details about applying unit costs to the DCH intervention are provided in the separate scientific report (<https://eprints.whiterose.ac.uk/>).

The DCH intervention cost estimates were compared to those unit costs associated with a long or short-stay non-elective inpatient admission, or an A&E visit (i.e. hospital contacts which could be avoided using the DCH intervention) in order to suggest how many contacts may need to be avoided to achieve cost-neutrality when investing in the DCH intervention (i.e. the intervention cost is equal to the cost-savings of avoiding down hospital contacts).

Unit costs associated with healthcare staff and resources were obtained from appropriate reference cost sources, such as NHS Reference costs¹⁸ and Personal Social Services Research Unit (PSSRU) Unit Costs of Health and Social Care.¹⁷

Quantitative analysis

For each resident involved in the intervention, data were collected on their age, sex, date of joining the intervention, date of entry to the care home, and numbers of emergency contacts both between the start of the intervention and 29nd May 2018, and in the equivalent time period 12 months earlier, before the start of the intervention. Emergency contacts included both Accident & Emergency (A&E) attendance, and Inpatient contacts that were listed as emergency cases (and therefore excluded elective and day cases). Where there were multiple inpatient episodes with the same admission date, these were treated as a single emergency contact.

Descriptive analysis was conducted for all variables. The primary hypothesis was that the rate of emergency contacts would be lower for residents using the intervention than it was for these residents before they used the intervention. This was tested using a multilevel Poisson regression analysis, details of which can be found in the scientific report. Primary analysis was restricted to those participants who were already resident in their care home at the start of the baseline period, but a secondary analysis included all data, regardless of when residents entered the care home.

6.4 Key Findings

6.4.1 Process evaluation key findings

6.4.1.1 Project design

6.4.1.1.1 What was the process to design the project?

The PPP Test Bed PMO identified care homes as being one area that may benefit from technology. Inhealthcare were already working closely with the PMO, and reported prior experience of using this technology to support other forms of monitoring in care homes in other UK locations.

Initial aspirations of the PPP Test Bed were to develop a 'command centre' and whilst commissioning challenges meant that was difficulty in achieving in its entirety, this project provided a demonstrator of the principles in a care home setting linked to acute service responders.

The National Early Warning Score (NEWS) was chosen as the vital signs assessment tool to digitise, because it is an existing assessment, based on Royal College of Physicians (RCP) endorsement, observations required could be obtained using easily available equipment. Whilst we could not find evidence of prior use of the NEWS in care homes, there have been uses of the Modified Early Warning Score (MEWS) in care homes, which might be more suitable. However, there is very little evidence about its influence on outcomes.³¹

Inhealthcare and PPP Test Bed PMO reported positive collaboration with each other to design the project. There was a pilot phase to this project, involving one care home, and feedback from the manager of this care home was taken into account.

6.4.1.1.2 Were the interventions delivered in line with the proposed plans?

The only change made to the proposed plan was that it was initially planned that a community matron service would also be available, to support the care home to care for the resident in the home environment, if needed. However, when the community matron service was no longer available to the project this was replaced with the availability of a specialist palliative care nurse from the local hospice providing community nursing support. However, this element of the service has not been used as to date all alerts have been resolved via dialogue with SPA and GPs.

6.4.1.1.3 What changes had to be made during implementation to ensure effective delivery of the intervention, and why?

Initially daily vital signs observations were planned as part of the pilot phase at one care home, but interview accounts suggest that this was likely to be too demanding for the care home to manage so it was adjusted to be weekly observations (although in one home this is twice weekly).

6.4.1.2 Recruitment:

6.4.1.2.1 Were the governance arrangements for the intervention effective and why?

The governance arrangements have been effective for the set-up, implementation and daily running of the project. The intervention fitted into an existing responsive service, which had used a text messaging system to prompt people at risk of deterioration to send observations to SPA. These would be responded to if required. The digital care home project augmented this and opened the service up to a wider group of service-users. There have been problems with some homes not returning observations at times. These

missing reports are monitored and if a systematic problem is suspected, then the PPP Test Bed PMO will contact the home to try to resolve the issue.

The PPP Test Bed DCH implementation team remain crucial for liaison with care homes, recruiting new homes and keeping them on board, when problems are encountered.

6.4.1.3 Partnership

6.4.1.3.1 Did the partnership of the NHS with innovator firms work as intended and why?

In this project partnership occurred between the NHS (the PPP Test Bed PMO and SPA) and Inhealthcare, and also between these partners and care homes. All partners reported positive collaboration. NHS SPA staff did also report that, initially, it had been difficult not having a written 'how to' step by step manual detailing the tasks they needed to do but this was resolved. An Inhealthcare staff member reported it could be difficult to speak to care home staff to resolve technology queries staff had raised, but when they did speak to the right person these queries were resolved. An Inhealthcare staff member also reported that visiting different care homes on different days, for set-up meetings was a significant cost for their company, suggesting, ideally, seeing several homes in one day would have been more effective for them, but also recognised this may not have been possible for the care homes.

6.4.1.3.2 Has this partnership/ different engagement with the NHS resulted in improved technology pull-through?

It is not clear that the technological opportunities for the project have been fully realised. No interviewees reported GPs or NHS professionals, other than SPA using the data. One member of care home staff noted that information sharing between care homes and the NHS offered an alternative to fax or telephone communication. The two GPs interviewed reported they had not accessed the data available, although one reported they had problems accessing SystmOne on their practice PC. Some care home staff reported they were not aware that GPs could access the data. This element of the project has been promoted and supported. However, it has proved to require more engagement efforts than originally anticipated.

This project remains ongoing to develop closer links with commissioners in terms of wider plans for care home monitoring.

6.4.1.3.3 Has the engagement by each party to the partnership been sufficient and why?

SPA, the care homes, PPP Test Bed PMO and Inhealthcare were engaged to ensure delivery of the project. Care home staff reported the project was well run and felt able to contact the programme manager or Inhealthcare easily if needed. Some interview accounts suggest that it could be difficult for Inhealthcare or SPA staff to contact members of care home staff, to resolve queries, particularly within the timeframe needed (i.e. as data needed to be entered by midday). However, it was acknowledged that this was due to the demands that care home staff experienced, and partners were mindful of this, doing their best to build relationships with care home staff to facilitate collaborative working.

Some care home staff suggest they were not convinced that the project was preventing admissions or benefitting the residents or staff of the care home they worked in, as they felt it did not enhance usual care. This is possibly due to the perceived lack of any additional service, and the advice received from SPA confirming what they would consider usual care. Yet, this did not appear to prevent them engaging in the project by doing the tasks required. Several hoped that participation in the project would be of benefit to other homes, and the NHS. Some staff did express doubt whether the care home they worked in should participate again, if another similar project were proposed.

Care home staff also felt the project (i.e. the technology + doing weekly observations using the NEWS score + the link with SPA) had the potential to support less experienced or non-nursing staff not used to doing vital signs observations regularly or 'struggling' care homes. Other members of care home staff recognised that the project had provided them with access to support from health services, which might have been previously lacking.

6.4.1.4 Implementation and uptake

6.4.1.4.1 What were the facilitators of effective delivery (and uptake of technology/ services) and how were they ensured?

The technology: (i.e. the app, uploading the data via the inhealthcare platform) was regarded as usable and acceptable, staff learnt how to use it after an initial period of training and use.

The link with SPA: was regarded as beneficial by those staff who reported the care homes they worked in sometimes had difficulty accessing GP input.

Communication and support: most staff reported they had been well informed and kept updated about the project, that it had been well run and they could phone for support with any problems.

Manageable workload: most care home staff reported the tasks were manageable for the number of residents they were asked to do observations for. Agreeing named members of staff to do the observations, setting aside time and uploading data before the medication rounds in the morning appeared to help staff complete tasks. Management support was also important, some managers reported uploading data on behalf of staff, or doing observations when staff were unable because of other demands.

Staff skills and knowledge of residents: some staff reported assessing non-verbal communication with residents with severe dementia to judge if they were happy to have observations taken, or knowing residents complex medical histories, and explaining to SPA why a resident may repeatedly be getting alerts raised.

Recruitment of care homes (September- December 2017): the programme manager's efforts to engage care homes resulted in seven care homes participating and coming on board with the project within a short time period. Regular communication and flexibility to work around the demands faced by care homes, from all partners appeared key.

Recruitment and selection of residents: Some staff reported they were able to identify only a limited numbers of participants who had the capacity to consent to participate or expressed concern that most of those selected (from the care home they worked in) to participate were well and may be unlikely to benefit from this project. However, this observation contradicts the data. The reality is that there is a wide range of level of care need represented. Some residents have low level needs and can be supported to remain well, whereas other residents are more frail and require early interventions to prevent emergency admissions.

Some staff reported that some residents with limited verbal communication and behavioural symptoms of dementia objected to having the observations done as they did not understand what was being asked of them. Care home staff reported using their knowledge and skills to assess residents' ability to consent to observations and selecting residents that would benefit.

Time required to recruit care homes: An Inhealthcare staff member talked about how challenging it had been to recruit care homes, and required numbers of residents. With hindsight, they considered the recruitment process could have started whilst the pilot was running.

Advice from SPA: Some staff felt the advice from SPA, following a raised alert, had not added to the usual care residents received. This appeared to be because the advice had been to contact the GP or continue monitoring, and staff reported this was already occurring, some reported cases where medication had already been prescribed and administered before SPA contacted the home. This is an interesting area for further investigation. It is not clear whether this perception by some staff is representative, or whether the monitoring process and knowledge that when alerts are triggered this will prompt a call from SPA is in itself contributing to improved care and extra vigilance. It could be that an additional focus on these residents' care might result in subsequent advice being seen as superfluous.

Limited understanding of hospice nurse's role: staff interviewed appeared unclear about what the role was for the community nurse support in this project (which was to fulfil the community nursing support role for the purposes of the project). There were no reports of this service being used at the time of interviews.

Timing for uploading observations: uploading observational data by midday was reported as difficult, by some, to balance with other workload demands (e.g. when a care home had only one qualified nurse working, who was responsible for doing the observations, uploading data and doing the home's medication round). However, this is considered necessary in order to arrange a response during the working day.

Communication - between care home staff, SPA and Inhealthcare: Inhealthcare and SPA staff reported it could take several attempts to speak to the care home staff they needed. SPA staff reported that sometimes, particularly initially, some care home staff appeared to feel somewhat affronted to have the care they were providing questioned by the SPA nurse advisors, and SPA staff reported they could feel awkward doing this

Overcame by: SPA staff tried to build friendly working relationships with the care home staff that they contacted, and said this improved over time. Care home staff explained residents' medical histories or actions taken to SPA staff; SPA staff and Inhealthcare kept trying to contact the right people, understanding how busy care home staff could be. SPA clinical operations manager, Inhealthcare staff and the PPP Test Bed programme manager all attended set up meetings with care home managers. Inhealthcare provided written information for care homes about how to upload the data and a short 'how to' video was produced during the course of the project.

Limited data sharing: some care home staff questioned whether SPA staff could see the residents' history or history of previous alerts; a SPA staff member questioned whether care home staff knew that if they completed and entered in comments this would raise an alert and SPA staff would see this, even if the NEWS score was not high enough to raise an alert. A SPA staff member reported they did not telephone the home when this happened. One GP reported they could not access System One (where the data was recorded by SPA) because their practice used EMIS and neither GP nor any staff member interviewed reported they were aware of GPs accessing or using the data from these observations to inform their clinical care. As mentioned earlier, there have been activities aimed at encouraging wider data use.

Lack of feedback for residents: Some staff had expected residents to be given feedback, perhaps in the form of a print out, about their observations

6.4.1.4.2 Were there any unintended consequences that needed to be managed and how was this done?

The GPs interviewed expressed concern about possible negative unintended consequences. They described the potential for this project to increase workload, should numbers of participants increase. This could be due to staff or themselves thinking more actively about interventions or treatments, or increased admissions due to over-reliance on a mechanised/protocol for assessment rather than individualised, holistic medical assessment taking account of multiple factors (including for example pharmacology, medical history, long term conditions, usual or current behaviour and physical presentations).

A positive unintended consequence that was reported, relates to the potential for monitoring to indirectly improve the care of residents. It is possible that the act of regular monitoring increases the knowledge and understanding of care home staff, this could also increase vigilance and attention. There is evidence in the literature for this and it is corroborated by notes associated with responses to alerts, where care home staff members have taken actions to improve measures. There is also a possibility that the link to SPA and referrals to GPs has helped to improve the level of service that some homes are receiving from GPs.

6.4.1.4.3 To what extent is the intervention likely to be scalable and why?

The intervention could potentially be scaled-up with the support of a small implementation team. However, there are a number of elements of the intervention that require improved understanding prior to attempting larger scale implementation.

Further consultation with local stakeholders about the next steps for such a project is recommended, as staff views about the actual and potential benefits of such a project were mixed. Residents interviewed did not object to having weekly observations and appeared reassured that their health was being monitored. Also some would be happy to have observations done more often, especially if tailored to their individual health condition. However, some residents did also report they would rather go into hospital to be cared for if unwell, rather than remain in their care home setting.

The potential impact on other related health and social care services and appropriate support need to be better understood prior to scaling up. It would also be worthwhile understanding whether there are types of care homes that are more receptive to the intervention or stand to benefit more than others. The optimum selection of appropriate residents should also be investigated. There were some concerns raised about the appropriateness of the NEWS to trigger alerts. Therefore, the exploration of modified scores would be beneficial.

6.4.1.5 Recommendations

Qualitative interview accounts suggest that if considering continuing with, scaling up or rolling out this project the following issues should be considered:

Context: the usual care context in relation to hospital admission, the relationship with GP and advanced care planning processes:

- recruiting care homes which recognise need to reduce their admissions or who want support with this may influence the success of similar projects
- consulting and gaining information about how an individual care home works with a GP and consulting the relevant GPs, care home staff and managers about whether or not they think this relationship would benefit from the additional support of SPA and/or community nursing (or what kind of support they would find most helpful)

Recruitment and selection of residents: consider the inclusion/exclusion criteria for residents and the potential impact of this. Some staff reported people who were very unwell or had severe cognitive impairment had often been excluded, as they could not give consent, whereas they may have benefitted from increased observations, and/or those who were generally well were included, so this project may see limited impact. However, some residents with dementia are participating. One staff member suggested participation could become part of usual care, so that residents could be consented to take part as part of the care home admission process, but could opt-out, if wished. This was suggested as one way of enabling those without capacity to consent, to participate. However, the consenting process is part of the Test Bed programme and would not necessarily form part of any wider roll out.

The NEWS score and types and frequency of observations: consider if NEWS score is the most appropriate assessment tool for this client group or identify characteristics of residents who might benefit from NEWS score i.e. is it appropriate for nursing / residential care residents / particular health conditions/ end of life? Also consider the frequency of observations; some care home staff felt observations should be more frequent if they were to identify early signs of ill health, yet the pilot study identified that daily observations were too much of a demand for care homes.

Consider and consult stakeholders about the roles for SPA & other support: clarify roles including considering whether SPA and community nurses are the most appropriate services to support the care homes or perhaps other resources, such as the addition of mental health nursing. Consulting with front line staff as well as managers and providing examples of what both services can offer to care homes to support them care for residents at the home may be helpful.

Timing of observations: consider if and how the deadline for uploading observations needs to be adjusted to minimise the burden on care home staff whilst leaving enough time for SPA staff to respond.

Dissemination & Feedback loop: Several staff wanted to know whether the project had led a decrease in admissions. Consider how to update staff and stakeholders throughout the course of the project and about when findings will be available and how they will be shared:

Consider how to share data with residents and care homes: There was an appetite from some residents for knowing their 'results' from the observations, even if no problems were identified. One member of staff talked about how she had expected individual results to be shared with residents and possibly staff, for example in the form of a print out.

Incentives for care home participation: consider what might be effective incentives to encourage participation of care homes

A 'how to' step by step manual for SPA and care home staff using the technology, from the start

Workload: whilst staff interviewed reported the workload was manageable, some also said if the numbers of participants they were expected to complete observations increased, they would need another person suitably qualified and experienced to do the observations, and trained to upload the data

Also consider:

- Exploring ways to support multi-disciplinary team (MDT) advanced care planning, including ceilings of treatment (i.e. the limits of active treatment for a particular condition that are based on discussions with the relevant medical personnel and MDT), within care homes to help avoid unnecessary emergency admissions, including ways to have this information accessible via a portal/online for the NHS, care homes, and GPs to all access
- The role of advanced care planning and the relationship with the GP could be key to preventing unnecessary emergency hospital admissions/A&E attendance. It is important to bring families and residents along with these plans and discussions
- Key performance indicators could be considered for long-term monitoring. For instance, monitoring of readmission over the whole period of time a person would remain in a care home. For example, examining readmissions within 24 hours could be valuable.
- It could be important in preventing admissions for people at the end of life if a rapid response was introduced to the current pathway.

6.4.2 Impact evaluation findings:

6.4.2.1 Stakeholder benefits

6.4.2.1.1 Did the NHS get better products or processes as a result of collaboration/testing/learning?

SPA established a new link with care homes, enabling them to discuss participating residents' health. It is unclear from qualitative findings what the impact of this was in a measurable way. Staff views on whether benefits had occurred as a result of this project were mixed. Staff working in care homes with an established relationship with GP and/or established care planning processes, reported they felt it had little, if any, impact on their residents or A&E attendance. Staff from two care homes that reported difficulties accessing GPs as needed reported they felt the project had supported their care homes to monitor early warning signs of ill health in some residents. One care home manager thought demonstrating that information could be shared digitally between care homes and the NHS (i.e. SPA) was positive.

6.4.2.1.2 What have the benefits to innovation partners been of engaging with the NHS as part of the Test Bed programme?

A benefit to Inhealthcare was the company being able to iteratively develop and improve the technology based on feedback from stakeholders, as it was reported they felt they had enhanced 'the care home

module' (i.e. the existing technology they had already developed, applied to care homes) as a result of this process. Also, an Inhealthcare staff member said they hoped this was a future relationship to take forward, that if they could prove that the digital care homes concept has a big impact on residents care it could be expanded to new areas or other care homes in Sheffield.

6.4.2.1.3 Patient experience: What were the impacts of the intervention on patients' experience?

Interviews with care home residents and relatives indicated the project was acceptable to them, none objected to having weekly observations, and some were willing to have observations more often, particularly if tailored to their individual health needs. It seemed residents and the relatives interviewed regarded this project as offering reassurance that they or their relatives were being well looked after, and their health being monitored. However, it is important to note that some residents stated they would wish to be admitted to hospital if they became unwell. Reasons for this were unclear but interviewers wondered if residents felt hospital would offer more specialised medical care, or more social contact, than their current care home. Some residents interviewed would have liked to see the results of their observations.

Under what conditions and for whom were improvements in health outcomes most significant (and least significant)?

Some care home staff suggested that observations would need to be more frequent to be of benefit, as people can start become unwell at any point in time, not just once a week. Some staff suggested that all residents, regardless of condition or needs could benefit from more regular observations, as standard practice was that people with nursing needs had vital signs observations done monthly, and people with 'residential' needs did not have such observations done unless they started to become unwell.

Some care home staff, a GP and a resident discussed how structured vital signs observations (or the NEWS score) were not necessarily needed to identify early signs of ill health. Rather consistency of staff caring for the person and social contact with staff who knew a resident well could help identify signs of early ill health. Some staff also questioned whether the NEWS score assessment could lead to benefits for those at the end of life or with long term conditions or severe cognitive impairments, as their scores on this assessment would regularly raise alerts, and this is to be expected.

There were some examples given by staff of when the project had resulted in a resident receiving treatment for early warning signs of ill health by staff working in homes where they reported difficulty accessing a GP as needed. No such examples were reported by staff in homes where they also reported an established and positive relationship with a GP and an established process of advanced care planning. Some staff reported they felt the project had raised awareness about what staff should be looking for when monitoring the health of residents, and it had led to increased understanding of vital signs measurements. Some staff reported that the link with SPA was of potential benefit for residents.

These findings indicate two important hypotheses that require further testing. Firstly, the relationship between the home and their GP/GPs could be important in defining the perceived value of the project for that home. The second hypothesis is that, whilst some respondents (early on in the study) could not see the benefits of the project working as intended nevertheless appreciated that the additional awareness raising and unintended consequences for improved resident care were important benefits.

6.4.3 Evidence of impact

Economic analysis: DCH intervention cost and achieving cost-neutrality

When accounting for the implementation and operational costs of the DCH intervention as part of the DCH study (i.e. across 67 residents and 7 care homes), the total estimated intervention cost in the first year is £66,840 which equates to an equivalent annual cost of £64,172 (A more detailed table of unit costs associated with this calculation is presented in *appendix 6*).

Table 17 shows the unit costs of non-elective inpatient long stays, short stays, and A&E visits as reported in the NHS Reference costs for 2016/17.¹⁸ Based on the estimated equivalent annual DCH intervention

cost of £64,172, it is estimated that across the 67 residents in the DCH study, the intervention would need to avoid 21.5 long stay non-elective inpatient contacts per year at £2,984.71 per contact to achieve cost-neutrality; which is equivalent to a decrease in 0.32 long stay non-elective inpatient contacts per resident/year (see *Table 17*). For short-stay non-elective inpatient stays or emergency medicine contacts (e.g. A&E visits), a decrease in 1.55 or 6.46 contacts per resident/year would be needed to achieve cost-neutrality (see *Table 17*). If the decision maker wanted to re-coop the technology costs via hospital cost-savings over the first year of implementation, the intervention would have to avoid 0.33 long stay non-elective inpatient contacts, or 1.62 short-stay non-elective inpatient stay contacts, or 6.72 emergency medicine contacts per resident in the first year, as examples (see *Table 17*). For the purpose of these examples, these estimates are assuming that avoiding contacts are independent of each other, whereas in reality various types of healthcare contacts could be avoided to achieve cost-neutrality (e.g. avoid both non-elective inpatient and A&E visits). Based on the statistical analysis of hospital contacts conducted for this DCH study, there is no clear evidence to suggest whether these reductions in hospital contacts are achievable; further evidence is required.

Table 17: Number of hospital contacts needed to avoid for DCH intervention cost-neutrality

Parameters	No. residents	Based on total cost (first year)	Based on total EAC
Total DCH intervention cost (£)	67	£66,839.58	£64,172.07
Non-Elective Long Stay cost (£)	1	£2,984.71	£2,984.71
No. to avoid per year (cost-neutrality)	67	22.4	21.5
No. to avoid per resident/year (cost-neutrality)	1	0.33	0.32
Non-elective Short Stay cost (£)	1	£617.11	£617.11
No. to avoid per year (cost-neutrality)	67	108.3	104.0
No. to avoid per resident/year (cost-neutrality)	1	1.62	1.55
Emergency medicine cost (£)	1	£148.36	£148.36
No. to avoid per year (cost-neutrality)	67	450.5	432.5
No. to avoid per resident/year (cost-neutrality)	1	6.72	6.46

Footnote. All hospital contact costs sourced from National Reference cost for 2016/17.³² EAC = Equivalent Annual Cost; see also Table 10 in appendix 6.

Number of emergency contacts

There were 67 residents who used the intervention across the seven care homes. The number per home ranged from 5 to 16.

The start date for the intervention varied by care home, from 5 June 2017 to 11 January 2018. As the date for the end of the study period was 29 May 2018, this gave between 137 and 357 days for the intervention period.

The baseline period was treated as exactly one year before the intervention period; therefore, this varied by care home also. 36 of the residents (54%) were in the care homes before the baseline period started; 12 (18%) became resident during the baseline period; 13 (19%) only became resident after the baseline period finished; and for six residents we did not have the date of their becoming a resident.

The total numbers of emergency contacts per resident across the baseline and intervention periods are shown in the following table (including all participants). Across all the participants there were 81 emergency contacts in the baseline period (44 A&E, 37 inpatient), and 60 in the intervention period (35 A&E, 25 inpatient).

Table 18: Emergency contact frequency

Baseline period			Intervention period		
Emergency contacts	Frequency	%	Emergency contacts	Frequency	%
0	36	53.7%	0	39	58.2%
1	8	11.9%	1	11	16.4%
2	10	14.9%	2	10	14.9%
3	3	4.5%	3	1	1.5%
4	7	10.4%	4	4	6.0%
5	2	3.0%	5	2	3.0%
6	1	1.5%			

The underlying periods for this differ by care home, however, due to the staggered start dates of the intervention. A more useful way of showing the data is as a rate of emergency contacts over a 12-month period. The following table shows the mean, standard deviation, median and interquartile range for each period, both including all cases, and then only including the 36 cases who were resident in the care home at the start of the baseline period (labelled as consistent cases).

Table 19: Emergency contact rate per 12 months

Contacts per 12 months			
Baseline period		Intervention period	
All cases		All cases	
Mean (SD)	2.4 (3.3)	Mean (SD)	1.8 (2.9)
Median (IQR)	0 (0.0-3.9)	Median (IQR)	0 (0.0-2.7)
Consistent cases		Consistent cases	
Mean (SD)	1.0 (2.1)	Mean (SD)	1.7 (3.0)
Median (IQR)	0 (0.0-1.9)	Median (IQR)	0 (0.0-2.4)

This reveals an interesting effect. When all cases are considered, there appears to be a slight decrease in the number of emergency contacts during the intervention period compared with the baseline period. However, when only considering participants who were resident in the care home at start of the baseline period, there appears to be a slight increase in emergency contacts over the intervention period compared with baseline. This highlights the importance of considering like-for-like cases. Often, entry into the care home may be precipitated by a care episode, which may involve an emergency contact. Therefore, the higher rate of contacts in the baseline period when all cases are considered is possibly connected to the very reason that some participants entered the care home in the first place. This suggests that to get a meaningful comparison, the use of consistent cases (those who were care home residents across the whole of both periods) is indeed the more sensible approach.

To test whether there was a difference in the rate of emergency contacts between the intervention and baseline periods, we ran a repeated measures Poisson mixed effects model, using intervention period as the offset.

This confirmed what the descriptive statistics had suggested: that there is a slight but non-significant increase in emergency contacts in the intervention period compared with the baseline period. Specifically, the incidence ratio for events occurring in the intervention period compared with the baseline period was 1.59 (95% confidence interval (0.62, 4.06)), $p = 0.329$.

The conclusion, therefore, is that there is no difference that can be inferred between the rate of emergency contacts while using the intervention compared with baseline. The observed number shows a slightly increase (a rate of 0.6 per year higher in the intervention period), but the confidence interval here is (-0.4, 1.6), therefore including zero and meaning no firm conclusions could be drawn.

Even if there were an increase in emergency contacts during the intervention period, there could be a logical explanation for this that has nothing to do with the intervention itself. These residents are typically old and are often frail or experiencing multiple comorbidities; the extra year of age in the intervention period could well be associated with greater health problems generally, which might cause the higher rate of emergency contacts. Without an appropriate comparison group, it is impossible to know whether this would be the reason.

As a secondary analysis, we examined the rate using the same analysis but including all cases. Again, this was in accordance with the descriptive statistics: there is a slight but non-significant decrease in emergency contacts in the intervention period compared with the baseline period. Specifically, the incidence ratio for events occurring in the intervention period compared with the baseline period was 0.74 (95% confidence interval (0.45, 1.23)), $p = 0.245$.

6.4.4 Types of alerts and responses

Out of 55 alerts 11 (20%) were confirmed as incorrectly entered and there were 3 incidents of SPA staff not being able to make contact with the home. It is likely that care home staff will become more proficient at taking and entering readings with further experience and SPA checking back with them, thereby recognising when mistakes have been made.

Of the remaining 41 contacts with care homes, following alerts being triggered, 8 cases were either escalated to GPs or had a GP visit planned for the next day. One of these GP visits was recorded as resulting in a medication change:

Telephone conversation with Nurse re the patient alerting this am. X informed she has re took the patients observations which remained the same. X has contacted the patients GP who has took patient off her Atenalol medication. GP has no concerns. Patient is alert and does not want to be admitted to hospital.

Another GP visit resulted in pre-emptive medication being put in place for a palliative care case.

In two cases, alerts were attributed to contextual factors that could be resolved (i.e. high temperature in room and lying position). In 16 cases care home staff reported having no concerns about the alert. Other alerts resulted in care home staff reporting continuing monitoring.

6.5 Limitations

It is important to be aware of the following limitation of the evaluation:

- Qualitative findings are not intended to represent the views of all staff involved in project, and may not therefore be transferrable to other projects or settings
- Interviews were conducted between January and March 2018, whilst care homes only began monitoring between October 2017 and January 2018. This means that respondents had little experience of the project. We have anecdotal reports that care home Managers have different opinions now that the project has been running for some time.

- Residents approached for interview were all recommended by the care home staff for interview and so may not represent the views of other residents
- The lack of a comparison group means that the slight change between baseline and intervention periods in rate of emergency contacts cannot be interpreted meaningfully. Even if a significant effect had been found, it would not have been possible to attribute this to the effect of the intervention itself
- Various methods were attempted to overcome the lack of a comparison group using routinely collected data, but with no direct link between hospital and care home data which formed part of usual care, the only way to get the data required for a robust quantitative analysis plan would be for consenting patient groups which was beyond our financial, resource and time constraints for this particular study
- If a robust quantitative analysis plan is required, allocating resources to consenting a comparison group should be given thorough consideration. It is also worth noting that larger sample sizes and a longer time horizon to capture the hospital outcomes of interest would also be desirable.

6.6 Conclusions and implications

- There were no conclusive findings regarding the effectiveness in reducing non-elective hospital admissions.
- Some care home staff felt NEWS criteria for this population was not suitable as scores were regularly raised and caused alerts for those with long term conditions or at end of life.
- Views of tailoring of alert parameters were mixed. Some staff thought this might be helpful but others, including GPs had reservations, as they were concerned that tailoring might lead to missing signs and symptoms in particular patients, due to over reliance on a protocol or score.
- Some care home staff felt more regular vital signs observations for all residents could be beneficial. Others also commented that regular continuity of care could also be beneficial to identify if someone becoming unwell
- Concerns were raised that it needn't necessarily be the NEWS score that triggers an alert. However, it should be noted there is ability in the technology to add notes from carers to note signs of deterioration alongside the quantitative input and these notes also trigger an alert at SPA. These notes have been used, but it was not clear that care home staff understood that this automatically triggered alerts. It could be preferable to provide opportunity for notes that create alerts and ones that do not and are only for retrospective information.
- Differences between the care/nursing homes indicate variation in processes and outcomes. Key differences seem to be access to or quality of relationship with GPs, advanced care planning process (especially with residents with dementia/cognitive impairment), mix of nursing/residential needs and mix of nursing non nursing staff
- The relationship is unclear between types and frequency of alerts and the subsequent responses, and appropriateness of the clinical support. Often when alerts were generated action advised from SPA was usual care. However, for staff from two care homes, which reported difficulty accessing GP care at times, advice was reported as helpful.
- Doing observations and using the technology was considered acceptable and feasible by care home staff. There were some reports of residents with severe dementia objecting to observations. However, residents interviewed found it acceptable and were happy for observations to take place.

6.7 Recommendations

- There were variable views regarding the potential benefits of the project for care homes, which seemed to be related to the characteristics of the home. Further work should explore assumptions about possible characteristics of care homes that may benefit from and be interested in this intervention.
- Consider how to support advance care planning (including 'ceilings of treatment') as several staff discussed how this can help minimise hospital admissions, if patients and families are in agreement.
- Consider the importance of a same day response to prevent admissions. A systematic acute response is missing from the current pathway, but it is not clear whether this would be beneficial.
- To determine whether the use of the intervention is actually associated with a reduction in emergency contacts, as well as other outcomes, it would be necessary to test its use alongside data from a matched comparison group. Although a randomised control trial would be ideal, it is recognised how difficult to achieve this might be in practice, so an alternative method such as propensity score matching could be used instead (i.e. create sets of participants for treatment and control groups in an observational study, using participant characteristics). These results could be compared against the economic analysis presented in this report.
- It should be noted the exact intervention cost if the DCH intervention was to be rolled out further is dependent on the number of residents involved, and number and size of the care homes connected to the platform in terms of potential residents that require monitoring; the unit costs associated with each of these aspects are presented in the scientific report (<https://eprints.whiterose.ac.uk/>) to enable future studies and/or decision makers to calculate the potential cost of the DCH intervention for their own circumstance.
- Consider how to update staff and stakeholders throughout the course of the project and about when findings will be available and how they will be shared: Several staff wanted to know whether it had led a decrease in admissions.
- Consider how to share data with residents. There was an appetite from residents for knowing their 'results' from the observations, even if no problems were identified. One member of staff talked about how she had expected individual results to be shared with residents and possibly staff, for example in the form of a print out.
- Explore ways to support Multidisciplinary Team (MDT) advanced care planning (possibly with technology support), including ceilings of treatment, and ways to have this information available via a portal/online for the NHS, care homes, and GPs to be all able to access.
- In order to enable an assessment of 'effectiveness' or 'efficacy' (e.g. reducing emergency admissions) counterfactual information is required; due to the lack of linkages between care home and hospital data, this information would need to come from a consenting patient cohort rather than using anonymised patient records. Therefore, thorough consideration should be given to allocating resources to recruit and consent a patient cohort to just use usual care to form this counterfactual information to enable this assessment of effectiveness. However, it should be recognised that this can be extremely challenging, as there are no direct benefits to recruits.

7 Diabetes

Table: Diabetes Project summary

The project	Diabetes
Healthcare challenge	
Health care challenge being addressed by the project	Self-management of insulin regimens and improved access for patients and clinicians to frequent and accurate data about the individual's insulin use.
Project rationale	(see project logic model in <i>appendix 2.5</i> for details)
Rational and logic underpinning the project	<p>The aim of this project was to support improved self-management of insulin regimens and improve access for patients and clinicians to frequent and accurate data about insulin use.</p> <p>Shared systematic monitoring and reporting of time from last dose of insulin will help patients to better manage insulin levels, avoiding missed or 'double' doses, healthcare professionals will also have access to this data to facilitate overall management. See appendix for logic model.</p>
Intended outcomes	<p>Insulin is managed in a more controlled manner leading to better management of HbA1c and blood sugar levels. This should impact on long-term health status and use of health-care services.</p> <ul style="list-style-type: none"> • Primary outcomes - changes in HbA1c or changes in daily blood glucose meter readings • Secondary outcomes: <ul style="list-style-type: none"> - Number of incidents of hypoglycaemia, including required health services (self-reported or service-level records) - Improved health-related quality of life (EQ-5D-5L) - Increased patient activation level (PAM 13 scores)
Timing of project	(see timeline <i>appendix 3.5</i> for details)
Dates of project (including Evaluation dates if different) [numbers of participants]	<p>Phase I: Jun 2016- Jan 2017 (Insulcheck 'Classic' pilot) [12]</p> <p>Phase II (App and health record connectivity & 'Insulcheck Connect' bluetooth device): Jan-Aug 2017 - development and planning Sept-Nov 2017 - recruitment and baseline data collection Nov 2017 - project shutdown early Dec-Mar 2018 - project review and qualitative interviews Apr-July - analysis and reporting [Total 41; 20 intervention, 21 non-intervention]</p>
Technology and service delivery model	
Technology	<p>Devices developed by Innovation Zed Ltd.</p> <p>Insulcheck Classic: A device that fits onto an insulin injector pen, which records the date and time of the last injection.</p> <p>Insulcheck Connect: A Bluetooth Smartphone app, which records information (date and time of the injection) from the Insulcheck Connect pen device that fits on the top of an insulin injector pen.</p>

	Information is sent via Bluetooth to the app developed by PPP Test Bed innovation partners Inhealthcare Ltd. This app provides the individual and clinicians with a platform to access this information.
Service delivery model	The implementation team made contact with clinicians conducting hospital outpatient diabetes clinics. Information was given to patients and the implementation team were available to assess suitability, discuss the intervention and to on-board patients with the technology and download the app to their phone. Patients taking part in the project were asked to consent to clinicians having access to the data recorded through the sensor and the app.
Those involved	
Target population	Patients aged 18 and over with insulin dependent diabetes who attend either of the Diabetes Clinics at Sheffield Teaching Hospitals NHS Trusts.
Those involved (stakeholders)	Patients who use insulin to manage their diabetes, carers, GPs, Insulcheck and Inhealthcare, specialist diabetes services at Sheffield Teaching Hospitals NHS Foundation Trust, PPP Test Bed PMO, Healthwatch Sheffield, SchARR, NHSE. Telehealth Team
The evaluation	
Evaluation questions	<ol style="list-style-type: none"> 1. Does the use of the Insulcheck Connect device and phone application improve diabetes control as measured by routine HbA1c testing or blood glucose meter readings? 2. Does the use of the Insulcheck Connect device and phone application reduce incidents of hypoglycaemia? 3. What are the potential economic consequences of the intervention? 4. What are the barriers and facilitators for implementation and evaluation of the intervention? 5. Are there people for whom the technology is more suitable and acceptable?

7.1 The technology and intervention

This project began as a pilot with a small population of people that needed to inject insulin to manage their diabetes and who also required daily visits from a community nurse for administration of their injections. This was considered a useful route to access people that might benefit most from the technology. The technology (Insulcheck) is a simple clip on attachment that is designed to fit a standard insulin injection pen without modification of the pen or its operation. Once attached it automatically detects when an injection is performed. Logging the time and displaying the elapsed time since the last injection for the user.

However, following this pilot, the innovator released a new model (Insulcheck Connect), which was intended to log and send information about times of injections via Bluetooth to the user's smart phone, and potentially electronic medical records as well. Two companies, Inhealthcare and Innovation Zed (developers of Insulcheck) collaborated with the PPP Test Bed to deliver this project. Each company had developed a technology independently of each other. The Insulcheck Connect trial device required an app and a web platform which were developed by Inhealthcare. A pilot phase had involved examining the acceptability of the InsulCheck Classic device (no Bluetooth) alone and evaluation data collection tools, which did not involve sharing or syncing data with electronic patient record systems with the main goal to improve patients' self-management.

Innovation Zed informed the PPP Test Bed team that they were developing an app and a device with Bluetooth connectivity which would improve patients' self-management and also provide data to support clinical decisions. It transpired that this technology would not be ready and available within the time span of the PPP Test Bed project. So, in line with the NHSE's Test Bed's aim to use combinatorial technologies to improve care pathways the PPP Test Bed introduced these two companies to each other; in order for them to find a solution within the timescale available. Inhealthcare already had an app that had been developed for another purpose which they adapted to work with the Insulcheck Connect trial device and also provided a web platform.

7.2 Engagement report provided by Healthwatch Sheffield

People who injected insulin and those who had family and friends with insulin pens thought the device would be useful. The focus on prevention and empowering patients to self-manage were thought to be valuable aspects of the project, as was avoiding the potentially serious consequences of having a double dose of insulin within a short period of time. It was thought that the device and app may bring users peace of mind in this respect, but a pen with an inbuilt mechanism preventing insulin being injected twice within a fixed time was thought to be a better option.

It was widely acknowledged that patients who lacked digital skills or didn't own a smartphone or have internet access, either through choice or due to cost, would not be able to participate and use the technology offered.

The device only fitting with a limited number of insulin pens and people feeling checked up on by clinicians were put forward as reasons why the project may not be desirable to potential participants. The value of the device and app was thought to be of limited without the inclusion of a reminder of some sort to prompt participants to take their insulin, and it was thought that insulin users would benefit from the dose of insulin administered being recorded, but this was not offered by the technology. The technology was considered not to reflect that patients might find monitoring certain information useful in managing their Diabetes, whereas clinicians might want to access different information to inform their patients care.

Recruiting patients following their hospital clinic appointment was thought to be inconvenient for some patients and the suitability of giving them a considerable amount of information following an appointment was questioned. A drop-in session at a central clinic was suggested as an alternative setting to onboard patients with the technology.

Teenagers and people who struggled with remembering were viewed as being likely to benefit from the technology. Newly diagnosed patients were suggested as ideal candidates to use the technology, as it would be part of how they learnt to manage their condition from the beginning. However, it was acknowledged that some patients might not appreciate this at a time when they are coming to terms with their diagnosis.

An unintended benefit of the technology was thought to be that it could help overcome a language barrier between patient and clinician. An example was given of a similar device helping an insulin user share information with their clinician which they couldn't express in English

7.3 Evaluation Focus

The main project was a promising intervention, in terms of the quality of evidence that the evaluation was expecting to produce and the combinatorial element (involving two innovator partners). However, shortly following initial recruitment, the Insulcheck Connect device and associated technology did not prove to be market-ready. The PPP Test Bed PMO team took the decision that no further recruitment should take place at least until the problems identified with the technology had been satisfactorily addressed. Any devices that had been issued were recalled and no further recruitment took place. Therefore, the evaluation focus is on describing the planned evaluation design and a process evaluation aimed at understanding how the project had developed and why it stopped early.

7.4 Methods

7.4.1 Process evaluation methods

Face-to-face, telephone interviews and one focus group were used to collect qualitative data from key stakeholders and service users. Interviews were guided by interview guides. The interview guide was a practical data collection tool called an experience map, which facilitated collecting data within a comparative framework. Participants were encouraged to look back and describe their experiences of being involved in the programme.

Eight participants were interviewed for this project. Two innovators, one diabetes clinician and five people with diabetes were interviewed. In addition, one focus group with six participants including programme managers, engagement lead, and evaluation team members was conducted. We used thematic analysis^{2,3} to generate themes and discuss the themes supported by participants' accounts.

7.4.2 Impact evaluation methods

The evaluation plan for the main project was to conduct a service evaluation of the two diabetes outpatient clinics in Sheffield to provide contextual evidence and baseline data for technology implementations. A decision was made by the clinical team responsible for both clinics that the technology was to be implemented at Sheffield Teaching Hospital's Northern General Hospital site (NGH), and a decision whether to implement at Sheffield Teaching Hospital's Royal Hallamshire Hospital site (RHH) would be delayed until after more was known about the implementation process. This provided the opportunity for a comparative evaluation of services at both sites.

Patients at NGH were recruited to use the technology. Participants at both hospitals were consented to anonymously sharing their hospital records and completing patient-reported outcome measures (PROMs) and consumer service receipt inventories (CSRI) for the purposes of the evaluation.

PROMs questionnaires and self-reported descriptive, diabetes and health status data were completed either face to face or over the phone with trained health care professionals or Healthwatch Sheffield staff provided by the PPP Test Bed programme or were self-completed and posted back. PROMs were completed at the recruitment stage and the intention was to repeat three months later at follow up (this was to include self-reported hypoglycaemia events in prior three months).

Participants were also supported to complete consumer service receipt inventories (CSRI) by trained health care professionals or Healthwatch Sheffield staff, which were to be used to inform an economic analysis, on recruitment and had been planned at three months. The most recent HbA1c reading and follow up reading were also recorded on recruitment and the intention was at follow up as well.

The key outcome measure for the evaluation was changes in HbA1c. There was an option of exploring glucometer readings, which are less useful data, as we were informed that this might be more regularly collected and recorded.

7.5 Key findings

7.5.1 Process evaluation key findings

7.5.1.1 Design and Set Up

7.5.1.1.1 What was the process to design the project?

It is relevant to note that the products which formed the basis of the Test Bed projects were selected as a consequence of extensive clinical and patient engagement which took place within the early stages of the programme overall. The Insulcheck technology was by far the top product that clinicians and users felt could provide value to the NHS and patients.

The pilot study was using the Insulcheck Classic in a population of patients in whom it was considered could most benefit – i.e. those who required daily visits for administration of insulin. This was to test out

feasibility, but it also identified that the number of potential recruits in this category was likely to be small.

In order to look at the value of this technology in the general diabetic population the setting was changed to recruitment at secondary care clinics. It was at this time that the innovator made Insulcheck Connect available and they were not keen for further testing of the classic model. There were no indications of any issue with market-readiness at the start of Insulcheck Connect trial. Problems only became apparent once feedback started to be received about the product reliability, and hence for safety reasons recruitment ceased and products were recalled.

7.5.1.1.2 What changes had to be made during implementation to ensure effective delivery of the intervention, and why?

During the pilot project recruitment was slower and lower than that expected. Reasons for low recruitment were that due to existing cognitive challenges, people did not or could not undertake new 'learning' and they did not want to disrupt the routines they had already established. Other reasons were that they had carers that managed their medication, the device did not suit their pen or management regime (e.g. more than one pen) and participants did not have the dexterity required to use the device. Slow recruitment was also a result of competing clinical pressures of the staff involved and a limited number of potential participants within the cohort (i.e. patients within the community nurse team's care who used an insulin pen).

The decision was therefore made to seek a setting with potential for higher recruitment rates for the main study (hospital outpatient clinics). During this transition the new device (Insulcheck Connect) was introduced and the supporting app and data platform was developed. However, the intervention was halted, and therefore effective delivery did not take place. The implementation process was challenged by the nature of cross-organisational working. Further, the difficulties in building effective working relationships between the two innovator companies impeded the process of co-design for a combinatorial technology, which had a significant negative impact on the implementation process and co-production of outcomes.

7.5.1.2 Partnership

7.5.1.2.1 Did the partnership of the NHS with innovator firms work as intended and why?

Differences in expectations about the aims and intentions of the national Test Bed programme created some lack of alignment leading to different interpretations of the project. While it was assumed that NHSE wanted a trial-based project evaluating cost and effectiveness of the intervention, the innovators in this case expected a design and development project for their innovation to be tested out with end-users in real world situations.

These difficulties were addressed by agreeing to approach the project in two-stages. The first stage was mostly focused on the delivery of the project according to the innovators' plans; to test implementation, recruitment, usability, acceptability and data flows. There was an agreement that the partners would work together to try to design an evaluation to assess effectiveness and cost benefits. Whilst all partners worked collaboratively to try to design an effectiveness evaluation, this proved very difficult in practice. Additional PROMs were introduced at 2 time points to try to assess improvements in health related quality of life. However, alongside other measures (required for stratification for personalisation of the intervention), this created burden on staff and service-users, and a number of service-users had already been recruited by this stage (precluding baseline data collection) which contributed to low response rates.

As an approach to rapid implementation, this 2-stage process was not successful in being able to provide good quality effectiveness findings. It is perhaps uncontroversial to assert that interventions designed in collaboration with service-users, service-providers and evaluators from the start have a better chance of providing satisfactory outputs for all partners.

7.5.1.2.2 Has this partnership/ different engagement with the NHS resulted in improved technology pull-through?

From the perspective of one of the innovators the project was a health service initiative and its purpose was to examine new technologies that have potential to help improve patients' pathway and new ways of data collection. The expectation was that the technology might be rolled out further with opportunities for the companies to get their innovation used and procured by NHS. However, Insulcheck benefitted from the project through the development of their product (Insulcheck Connect), which is now in production and available on the open market.

7.5.1.3 Implementation and uptake

7.5.1.3.1 What were the barriers to effective delivery (and uptake of technology/ services) and how were they overcome?

In addition to the inter-organisational relationship building, the implementation process was challenged by the lack of readiness of the technology.

7.5.1.3.2 What were the facilitators of effective delivery (and uptake of technology/ services) and how were they ensured?

1. Willingness of different stakeholders including the programme managers, innovators, patient groups and evaluation team to engage in cross-disciplinary partnership
2. Opportunities to obtain feedback from the potential users of technology and use the constructive feedback to inform the next cycle of technology development

7.5.1.3.3 To what extent is the intervention likely to be scalable and why?

Due to the challenges involved in the process of implementation, the project was not scaled up. Recruitment difficulties, were gradually being addressed, and would require further investigation to assess potential levels of uptake. The disconnection between clinical services and the project implementation staff was reported as a significant barrier. They were often not easily visible by patients (having to set up in different areas of the building) and the coordination of on-boarding with routine clinical interactions required further improvement.

7.5.2 Impact evaluation findings

7.5.2.1 Stakeholder benefits

7.5.2.1.1 Did the NHS get better products or processes as a result of collaboration/testing/learning?

The intention of the NHS was to redesign Diabetes care pathways and services using innovative technologies not just to evaluate a piece of technology. Technology innovations were only components of the service redesign.

7.5.2.1.2 What have the benefits to innovation partners been of engaging with the NHS as part of the Test Bed programme?

Engaging with the NHS as part of the PPP Test Bed programme was a great opportunity for innovation partners to get their technologies tested and evaluated by potential users including patients and health professionals in real situations. As a result of this partnership, their innovation (Insulcheck Connect) was improved and they came up with the next iteration of technology ready to be deployed to the market.

7.5.3 Quantitative key findings

The recruitment target was 150 per hospital. 20 subjects were recruited to the RHH and 21 to the NGH before the study was closed in November 2017 due to concerns as to the reliability of the technology.

Recruitment, at NGH in particular, was challenging and initially slow for a variety of reasons including:

- Identifying and contacting potential participants (clinicians signposting patients, clinic physical layout and patient arrival and departure processes)
- Issues with the technology (functionality and connectivity)

- NHS firewall and difficulty with WiFi connectivity despite portable WiFi box
- Accessing the app when on-boarding patients
- Not all patients used a pen which was compatible with the Insulcheck device
- Some patients did not have a smartphone
- Issues with dexterity to attach and use the device and app
- Requirement to add time to visit to the hospital (which people had not accounted for), or make an additional visit

Due to project cessation, only baseline data were collected. Where the PROMs questionnaires were available they were generally well completed with only a few individual questions not answered.

7.6 Limitations

Due to the project being halted we only have baseline data available and no formal economic evaluation has been conducted.

7.7 Conclusions and implications

Whilst the pilot study was conducted with a cohort that could have benefitted most from the technology, there were a number of difficulties associated with this population's relatively high care needs and disabilities.

This project suffered from a common complication in dealing with the evaluation of technology. Namely, that digital technology cannot stand still and needs to constantly evolve and adapt in a rapidly changing context. This resulted in the programme theory for how the technology was to result in benefits being quite different in the pilot and the main project, owing to developments in the technology.

It also suffered from a common implementation issue, in that the technology is only in small part of the project implementation, and the context of implementation is equally as important. In this case the pilot and the larger scale project were delivered in contexts that were very different. For these reasons the pilot and main project constituted fundamentally different projects and there was therefore limited learning that could be translated from one setting to the other.

The main difficulty with the implementation of this project was that the second iteration of the technology was not considered reliable and concerns were raised about its readiness for deployment, which caused the project to be halted, shortly after inception. However, there were benefits for the innovators in term of product testing and development opportunities, which have led to a product which is now available on the open market. There was also a large amount of learning that was derived from the recruitment process, including recognising the percentage of patients with compatible injection pens, and overcoming problems with NHS firewalls and Wi-Fi.

7.8 Recommendations

Different settings: Whilst it can be a useful stage of implementation to try out different contexts, cohorts, services and patient pathways, it should be recognised that moving to another setting will require a complete re-design of the intervention and evaluation. Additionally, transferrable learning between settings can be limited. Therefore, it is recommended to allow time and resources for these processes, and to treat each new setting as a new project.

Technological developments: Innovators should consider the implications of changes to their technology. Whilst this information is not always available, it can be useful to systematically and regularly check with innovators whether there are plans or aspirations for development of the technology. When there are likely to be technological developments it is useful to establish the extent to which the service-provider or implementation team are to be involved in the co-design or testing. This can be beneficial for co-production so that the technology is designed to meet the service requirements. Importantly, this can also provide the opportunity to engage with patients and the public to introduce user-centred design principles.

Readiness of technology: As a result of the problems around readiness of the Insulcheck Connect technology, independent laboratory testing of products was implemented for all products as part of the Test Bed.

8 Test Bed 'Plus'

8.1 3Rings Project

3Rings is an 'internet of things' intervention. It consists of a smart mains electricity plug, which allows appliances to be plugged into it. The 3Rings plug is designed to give daily reassurance to families or remote carers by monitoring the routine use of an electrical appliance, (e.g. a kettle) by an individual living alone and considered to be at risk. It indicates changes in behaviour that could be a cause for concern (www.3rings.co.uk). It lets family members know by email or text that the kettle has been used. If the kettle hasn't been used, then they are alerted via email or text that there might be a problem, and are asked to respond and resolve the alert.

The 3Rings project ran between June 2017 and April 2018 and the Sheffield Hallam University (SHU), Advanced Wellbeing Research Centre was commissioned to investigate and pilot the introduction of 3Rings plugs into the homes of people in the PPP Test Bed Region who have dementia (PwD) and live alone. The project partner Westfield Health was the commercial provider of the 3Rings plug technology. The Alzheimer's Society and Age UK identified potential participants for the pilot project.

The following information was provided by Sheffield Hallam University for the PPP Test Bed Team for inclusion in this report:

8.1.1 Aims and objectives

- Report on the feasibility of the use of a monitoring device for people with dementia and their family carers
- Evaluate the potential of digital monitoring to identify patterns and routines of daily living
- Achieve a theoretical framework for further testing of benefits and individual outcomes in relation to well-being, frailty and carer burden?

8.1.2 Methods

In this pilot the effectiveness of the 3Rings plug was measured quantitatively and qualitatively by SHU as follows:

- Study design: A non-randomised pilot study over 6 months with background data gathering and surveys.
- Recruitment: 31-person sample of people with dementia living alone in the Testbed Region (South Yorkshire) and the same number of Family Carers.
- Ethics and Informed Consent gained via Sheffield Hallam University Ethics Committee.
- Surveys: The Edmonton Frail Scale; The Zarit Burden Interview; The Short Warwick-Edinburgh Mental Well-being Scale (SWEMWBS).
- Analysis: Descriptive statistical analysis of responses to alerts and surveys.

8.1.3 Main Findings

31 people used the plug (equivalent to 0.75% of the target population). Participants came from: Barnsley (3 people); Doncaster (4 people); Rotherham (3 people); and Sheffield (21 people). Each person had a Family Carer (FC) who took part. 30 carers lived in South Yorkshire, and 1 further away.

Findings 1: 1% reach into the population of people with dementia living alone

- There are 12,046 people who are diagnosed with dementia in this region.
- It is estimated that 4,015 (about a third) of these people live alone.
- As 31 people took part, this means that we reached 0.77% of the total population of people who have dementia and live alone in this region.

Findings 2: Living well within a usual routine

- 22 people had fewer than 2 alerts each month, indicating that their day-to-day routine was stable over a 4-month period.
- 8 people had between 3 and 10 alerts each month.
- Altogether there were 266 alerts where family responded and said:
 - 215 there were no problems,
 - 46 were false alerts,
 - 5 action was needed
 - 0 emergencies.

Findings 3: Changes in Frailty and Burden scores over 4 months

Edmonton Frail Scale. The scores indicate that of the people with dementia in this study:

- 7 people improved and became less frail (23%)
- 10 people stayed the same (33%)
- 13 people became more frail and (43%)

Zarit Burden Interview. The scores indicate that of the Family Carers in this study:

- 18 people felt less burdened (60%)
- 2 people stayed the same (7%)
- 10 people felt more burdened and (33%)

Findings 4: Changes in Well-being survey scores (SWEMWBS) over 4 months

The scores indicate that of the people with dementia in this study:

- 13 people had increased levels of well-being (43%)
- No one stayed the same
- 17 people had lower levels of well-being (57%)

The scores indicate that of the Family Carers in this study:

- 10 people had increased levels of well-being (33%)
- 5 people stayed the same (17%)
- 15 people had lower levels of well-being (50%)

8.1.4 Conclusions:

The qualitative and quantitative evidence derived from this feasibility and usability study demonstrated that PwD and their carers continued using the 3Rings device over six months. In line with the study aims, we can conclude that the 3Rings technology has potential for digital monitoring by identifying patterns and routines of daily living. There were only a small number of pre-post measures, which cannot demonstrate any statistically significant findings, and no comparative data. Therefore, only descriptive statistics are available. These showed maintained or reduced carer burden in two-thirds of carers (33% increased burden). The majority of PwD maintained or decreased frailty, whilst 43% became more frail. However, whilst these results should be treated with extreme caution, it is clear that predicted benefits and individual outcomes linked to well-being were not realised in this small sample; wellbeing reduced for 50% of carers and 57% of PwD.

8.1.5 Next Steps:

The potential of the 3Rings device to be embedded in daily lives of people with dementia was tested in this study. The insight gained from this study in terms of its usability and acceptability by target users is promising in order to further examine and test the theories behind the technology in future studies. Theories underpinning the 3Rings are increased carer resilience and functional habituation to improve well-being by the use of digital monitoring technology and preventing a “crisis” by knowing routine and

recognising when something is out of the ordinary including the pattern of routine behaviour and putting remedies in place. If the assumptions about potential benefits are demonstrated to work for the target population then the 3Rings technology has potential to facilitate policy locally by reducing admissions to care homes and building family and community resilience.

8.2 Digital Health Training project

The Good Things Foundation partnered with the PPP Test Bed PMO to deliver the Digital Health Training project between May 2017 and June 2018. The following summary is derived from the final report for the project, which is available on the PPP website or from Good Things Foundation.

The Good Things Foundation was commissioned to investigate and pilot the training needs of health professionals which would need to run alongside the introduction of health technologies for patients. The Good Things Foundation is a social change charity, helping people use digital technology to lead happier and healthier lives. The Good Things Foundation works with the Online Centres Network of over 5,000 community based organisations across the country to design and deliver programmes that address social challenges, from unemployment and poverty to poor health and wellbeing, using digital approaches.

8.2.1 Aims of the project

To answer the question “How can we develop health professionals’ capacity to improve digital health skills for people living with long-term conditions?”

8.2.2 Methods

The Foundation used the following methods to develop a small scale pilot project:

1. Semi-structured Interviews with health professionals (n=10); roles included Receptionists, Nurses (community and hospital-based), Consultants, GPs and Practice Managers. Staff were asked about: perceptions of digital technology (work and personal use); what they thought patients felt about digital technology; how they respond to different ways of learning or training
2. Patient focus group with the PPP Testbed Advisory Group (TAG); findings from the interviews with health professionals were discussed
3. Co-creation workshop - the health professionals, TAG and staff from the Good Things Foundation were brought together to further identify where efforts in training health professionals should be focused.
4. Development of the first training package to address six core themes that emerged from the interviews, focus group and workshop
5. Test and learn: To learn about what did and didn’t work, the training was delivered to three cohorts of health professionals. (7 attendees workshop 1, 17 attendees at workshop 2, 11 attendees at workshop 3). Each workshop was iteratively developed, based on feedback from attendees and observers of the workshops
6. Evaluation: To aid understanding of how the workshop approach had resulted in behaviour change, an online survey was administered to each and every participant of the three workshops one calendar month after the workshop took place. (Response rate: 57% of 35 workshop attendees overall [20 total responses])

8.2.3 Headline findings:

1. There is a lack of digital skills confidence and willingness to engage amongst health professionals
2. To promote adoption by as many members of staff as possible, digital health training must be recognised at and promoted by all levels of the NHS hierarchy and should be incorporated into practitioners’ continuing professional development (CPD)
3. Peer support amongst colleagues at all levels is key to raising awareness of digital health before it can be translated into conversations with patients
4. To engage health professionals in digital health, the benefits to patients must be clearly communicated when promoting any related training
5. Allowing time to iterate the delivery of the training throughout this pilot enabled the team to make changes that had a positive effect on health professionals’ subsequent behaviour change

6. It is crucial to acknowledge the concerns of staff who are reluctant to engage with digital and to have those people as an equal part of any training
7. This pilot has shown that there is scope to embed digital health training for health professionals with a range of backgrounds and confidence when it focuses on core, baseline themes, rather than specific technologies. The workshops which have been delivered have been well received and there is benefit to scaling this approach across the Trust and beyond
8. Including patients or people with lived experience in the training workshops supports learning as it provides real context and a counterpoint to common misconceptions
9. This project has shown that delivering digital health training that focuses on sharing experiences and peer support achieves behaviour change. It is worth noting that the pilot as delivered has shown more success in enabling changes with how people speak with colleagues compared to changes with how they speak with patients

8.3 Predictive Analytics

8.3.1 Using machine-learning to classify patient readmissions within 30 days

The predictive analytics project sought to classify patient readmissions to predict, on the day of discharge from a frailty ward, whether a patient would be readmitted to the same or another frailty ward within 30 days. This type of analysis has been attempted before by others, with modest success, but in most cases the assessment has been made using hospital based data only. The objectives for this activity were twofold: first, demonstrate that an analysis tool could be developed that would be competitive with the state of the art based on hospital data alone, and second, to go beyond the state of the art by combining this with an objective assessment of the patient's social/home circumstances, to further improve the accuracy of the prediction. The social care service in Sheffield, as part of its care package planning process, has a very structured assessment which covers areas such as mobility, social isolation and other aspects of the patient's daily living which, when combined with the specific clinical conditions or interventions, might better help identify those at risk of readmission.

The PPP Test Bed has access to data from 1st January 2008 to the 14th September 2017. The project looked at patients with more than 1 admission to the frailty wards in the hospital. This resulted in 5021 patients with a total of 7542 hospital admissions, of which 2823 were readmissions within 30 days of discharge. There were 4.3 million data points associated with these admissions.

Analysis of real world data is very challenging as it is highly sparse, i.e. most patients have a unique pathway through their encounter with different tests and procedures which makes direct comparisons between them very difficult. The data has approximately 6,700 individual types of test or observation, some of which are numeric, some text, introducing problems with free-text fields, and some have no data associated with them at all. Not every event has even a majority of these data items associated with it.

After discarding other strategies such as time-series classification, the most promising strategy involved using a naïve-Bayesian strategy to classify patients according to the observation types, and then feeding this into a decision tree machine learning strategy, known as a Random Forest, together with other data collected about the event. This included age of patient at the event, length of current stay, number of stays in hospital in the previous lead period, and socio-economic about their local region.

Naïve-Bayes has a strong history in e-mail spam filtering, using the frequencies of words to predict either spam or not spam, and random forests are considered the pre-eminent decision tree technique. This is amplified by creating many (100's) of trees with different parameters in order to select the best. This strategy also uses a five-fold cross-validation strategy by running the trees five times on a random four-fifths of the data and use the remainder for validation. This is important as it means the success we get in this model is likely to be replicated on other data.

In this project we defined an event as the last day of a patient's previous admission, as patients can have multiple events. The dataset we used was unbalanced, with fewer events having readmissions within

thirty days than having readmissions after thirty days. The ratio between the two groups was 35:65, so it would be expected by chance to identify an event with a readmission within thirty days in 35% of cases.

Using the machine learning system to analyse the two years' data prior to the event, we found that if the machine learning system identified that the patient would be readmitted within thirty days, then the probability that the result is true is over 70%, doubling the success rate over blind chance. In the opposite case, the machine learning system also improves the chance, from 65 to 73%, of correctly identifying that the event will lead to a readmission after thirty days.

This project is ongoing beyond the duration of the PPP Test Bed and for sustainability is now embedded into the machine learning capability of the Trust's Scientific Computing team. Knowledge transfer and sharing of software tools between IBM, one of the innovator partners in the Test Bed, and this team, has led to new capability in predictive analytics for the host Trust, with the opportunity to spread analytics capability across the system.

9 Programme-Wide Evaluation

Table 20: Whole Programme summary

The project	Sheffield City Region PPP Test Bed
Healthcare challenge	
Health care challenge being addressed by the project	<p>“New treatments for a growing and aging population mean that pressures on the service are greater than they have ever been. But treatment outcomes are far better – and public satisfaction higher – than ten or twenty years ago. The NHS needs to adapt to take advantage of the opportunities that science and technology offer patients, carers and those who serve them. But it also needs to evolve to meet new challenges: we live longer, with complex health issues, sometimes of our own making.” (Next steps on the NHS Five Year Forward View, 2017)³³</p> <ol style="list-style-type: none"> 1. Innovations are not tested in combination with each other or dependent infrastructure 2. There is little evidence of ‘real-world’ implementation 3. Innovations are simply added, rather than used to re-think service delivery <p>These reasons are suggested to indicate an unexploited opportunity to <i>combine</i> different technologies, and test these together with innovations in service delivery in real world settings.</p>
Overview of project	Building infrastructure and testing innovative combinations of technology and pathway development with the aim of improving patient outcomes and experience at the same or lower cost than current practice.
Rationale of project	
Rational and logic underpinning the project	<p>Testing out methods for re-designing care pathways and new ways of working; integrating technology, co-ordinated across providers and around patient needs.</p> <p>Resulting in a wide range of possible outcomes including health, economic and social benefits.</p>
Intended outcomes	<p>Development of infrastructure and methods for testing, implementing and evaluating technological innovation in the NHS.</p> <p>Developing protocols and procedures for information gathering to feed into the data warehouse/predictive analytics project.</p>
Timing of project (see timeline <i>appendix 3.6</i> for details)	
Dates of project (including Evaluation dates if different)	Start February 2016 intended to be completed in March 2018, but extended until June 2018.
How the project changed over time	<p>Change</p> <p>The project focused on interventions in Sheffield City only, rather than the region as originally intended (except for 3Rings project which was region-wide), and relaxed the focus on people with a number of long term conditions (as many technologies were focused on a single health issue). The programme changed and evolved as it adapted to emerging situations and performance pressures and learning was operationalised.</p> <p>The delivery personnel also changed over time.</p>
Technology and service delivery model	
The technology deployed	Various technologies
Model of service delivery (including pre-Test Bed pathway)	<p>The PPP Test Bed delivery model involved bringing together NHS services, commissioners, innovators, evaluators and patients and the public to develop infrastructure, processes and systems for health technology implementation, service re-design and evaluation.</p> <p>Pre-Test Bed there was little dedicated infrastructure in the region for identifying technology requirements, linking this need with technology solutions and implementing and evaluating in revised patient pathways and delivery systems.</p>

Those involved	
Target population	Originally intended: Sheffield City Region, people with 3 or more long-term conditions (including mental health problems)
Those involved (stakeholders)	Very wide range of stakeholders
The evaluation	
Evaluation questions	<ul style="list-style-type: none"> - What is the value of a dedicated programme to identify, implement and test new health technologies and technology enhanced patient pathways to key stakeholders (including the NHS, service-users, commissioners, innovators, and higher education)? - What are the challenges for testing health technology and how can they be addressed? - What are the enablers for testing health technology and how can they be operationalised?

9.1 Background

The following quote is taken from the invitation to innovators, jointly published by the Office for Science, Office for Life Sciences, the AHSN Network and NHS England in 2015. This sets out the perceived problem that the Test Beds were intended to address.

“There is no shortage of innovation in the NHS or the health sector more widely. However, this innovation has not diffused as quickly as, or had the impact that, has been seen in other industries—particularly in reshaping how clinical services are delivered. This is despite the NHS having natural advantages over many other health systems including universal coverage of a diverse population, national standards and relatively rich healthcare data.

Test beds seek to address three important problems that have constrained the impact of innovation in the NHS. First, innovations are often implemented in isolation from each other—and from the infrastructure on which they depend. For example, new technologies have all too often been piloted without complementary changes to existing working practices and information systems, curtailing the value they can release.

Second, there is a comparative paucity of robust evidence about the effects of innovations in real world as opposed to experimental or research settings. This can lead to a divergence between the benefits found in experimental settings from those in clinics or hospitals. It also means we tend to understand too little about the economics of innovations before they are introduced.

Third, innovations are often introduced in an accretive way in the NHS; that is, on top of existing working practices and infrastructure, some of which these new innovations may even be designed to replace. As a result innovations often simply add cost, with little or no gain in value.

For these reasons there remains a large unexploited opportunity to combine different technologies, testing these together with innovations in how services are delivered in real world settings. This is particularly the case with digital innovations that empower patients to manage their own health and care, as well as those that exploit health data in new ways. These new generation technologies have the power to transform the care of long term conditions like hypertension, diabetes or mental health.”

(NHSE (2015) Real world testing of ‘combinatorial innovation’ A global invitation to innovators, p.2)³⁴

9.2 Technology and intervention

The PPP Test Bed proposal included general descriptions of the technology to be implemented and the potential populations. The following is taken from the 3rd Draft implementation plan 14.04.16.

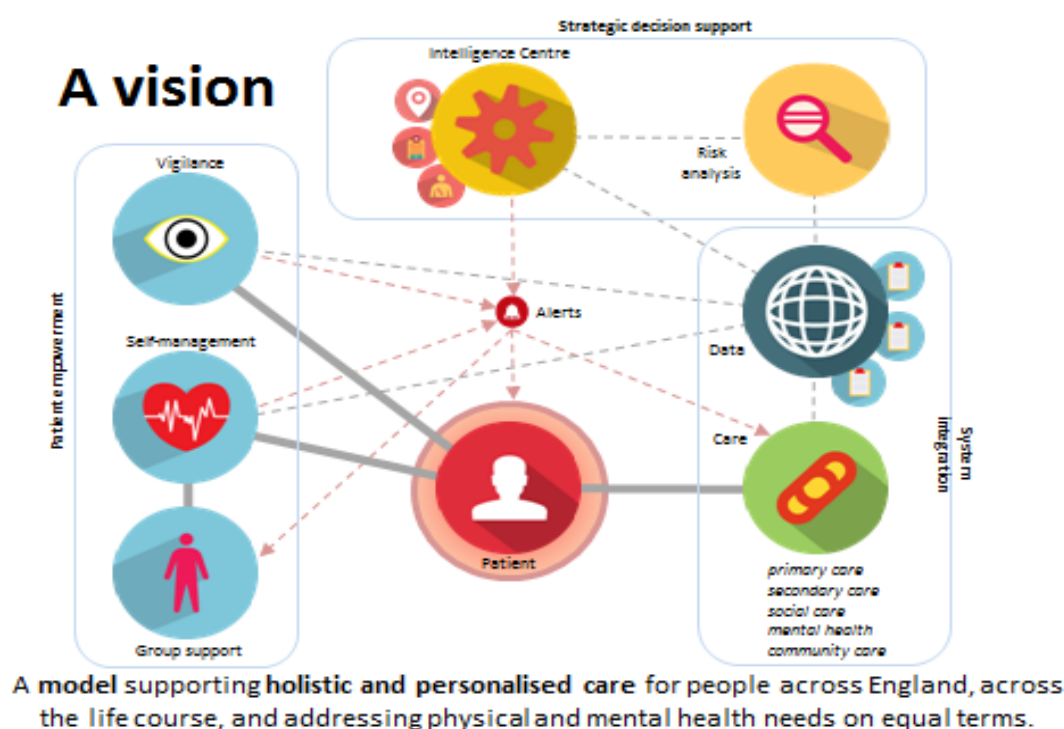
Programme Aim: The programme will create the perfect patient pathway to bring substantial benefits for patients suffering from multiple long-term health conditions, such as diabetes, mental health problems, respiratory disease, hypertension and other chronic conditions.

The **PPP Test Bed** aims to bring new benefits to patients with multiple long-term conditions through the combination and integration of innovative technologies and pioneering service designs, keeping them well and independent and avoiding unnecessary hospital attendances...

The technologies to be implemented in the programme were chosen through the NHS England led 'matchmaking events', where innovators were introduced to the implementation teams. A selection of potential innovations/innovators were selected to match the plans and ambitions of the original proposal. Fifteen innovator partners began working with the programme. There were a total of 29 partners of the PPP Test Bed Consortium, including 8 Trusts (Working Together Partnership), 8 Commissioners, 2 Universities, 15 companies, 87 GP practices (Primary Care Sheffield), 1 local authority, Healthwatch Sheffield, 1 Academic Health Science Networks (AHSN) and 1 NIHR Collaboration for Leadership in Applied Health Research (CLAHRC YH).

The vision of the programme was to integrate data systems across primary, community, secondary and social care and mental health services; to create pooled data resources. Digital monitoring technology would be used alongside self-management tools. An intelligence centre would provide predictive analytics, and strategic decision support. All of these systems would be based around holistic patient needs with the facility for real-time alerts to prompt rapid service decision-making. The original vision is visually displayed in the following diagram (*Figure 7*).

Figure 7: The PPP vision



It was from this collection of partners, ambition and vision and early stakeholder consultation events that decisions were formed about the individual technologies/innovators that would form projects within the programme.

9.3 Rationale for the intervention

Also extracted from the 3rd draft implementation plan, the five high-level objectives of the PPP Test Bed were:

1. Provide an ongoing platform for testing, refining and scaling-up innovations.
2. Re-design pathways, bringing combinatorial technologies and system transformations to support holistic and personalised care.
3. Embed the culture of transformation and improvement in NHS and other health and care organisations.
4. Support co-ordinated decision making across health and care, informed by real-time data and predictive analytics.
5. Evaluate the combination of new technologies and service re-designs producing robust and objective results that can be shared and disseminated.

9.4 Evaluation Focus

This programme wide evaluation aimed to explore the challenges and key elements for successful implementation of the PPP Test Bed programme. If the five objectives and six work-streams from the early planning phase are combined, 8 key themes emerge that describe areas of development and what the PPP Test Bed was intended to achieve:

1. Programme Management & Governance
2. Provide an ongoing platform for testing, refining and scaling-up innovations.
3. Re-design pathways, bringing combinatorial technologies and system transformations to support holistic and personalised care.
4. Technology Deployment
5. Patient Involvement and Engagement
6. Evaluate the combination of new technologies and service re-designs producing robust and objective results that can be shared and disseminated. [O5]
7. Embed the culture of transformation and improvement in NHS and other health and care organisations.
8. Support co-ordinated decision making across health and care, informed by real-time data and predictive analytics (Intelligence Centre).

9.5 Methods

Face-to-face and telephone interviews were used to collect qualitative data from a purposive sample. Interviews were guided by interview guides. The interview guide was a practical data collection tool called an experience map, which facilitated collecting data within a comparative framework. Participants were encouraged to look back and describe their experiences of being involved in the programme. Interviews were audio recorded and notes were taken during the interview. Informed consent was obtained before the interview, for the interview to be recorded and for the participant's job title to be used within the evaluation.

Thematic analysis^{2,3} was used to generate themes and discuss the themes supported by participants' accounts.

22 participants were interviewed specifically about the whole programme. Data from another nine interviews was also used within the analysis. These data were from interviews where the participant discussed both specific projects and the programme as a whole. This makes the total number of interviews analysed 31.

These included PPP Test Bed PMO team members, members of the evaluation team, and innovators, and two follow up interviews to explore specific areas of interest. One focus group was held with seven members of the evaluation team. Our aim was to get a range of perspectives.

9.6 Key findings

9.6.1 Process evaluation key findings

9.6.1.1 Design and Set Up

9.6.1.1.1 What was the process to design the project?

The different expectations of innovators indicate that there was a lack of alignment about the aims of the national Test Bed programme initially. For instance:

- One innovator expected the Test Bed to be a forum for them to test the movement of a combination of technologies from early-stage testing into a 'real-world' NHS setting
- Another innovator expected large-scale promotion and adoption of their technology, to mainstream its use in the NHS
- Yet another innovator expected a large-scale randomised trial; to demonstrate generalizable evidence of effectiveness
- Other innovators were invested in the development of the digital control centre. However, there were difficulties in developing the specification and potential areas of use for this

The programme had a short timescale, especially considering the testing of novel technologies and the need to redesign pathways and integrate the requirements of a large number of stakeholders. For the NHS locally, the key value of the Test Bed programme was the testing of methods and infrastructure for the identification, implementation and evaluation of new technologies. This was facilitated by the inclusion of a wide variety of technologies in a wide range of settings.

Early pressures from the funders to recruit large numbers of service users had an effect on the design of the programme; promoting the seeking of opportunities for large numbers of recruits. However, owing to the experimental nature of the programme, recruitment was unexpectedly slow in some projects and more rapid than expected in others. This resulted in a programme which needed to be agile and flexible. These interrelated factors created a complex situation where the design of the programme evolved and shifted over time to be more responsive to the emerging context.

9.6.1.1.2 Were the governance arrangements for the intervention effective and why?

The overall management and structure for governance arrangements was effective in responding to the needs of the programme. Many practical governance issues were realised as they emerged and were dealt with retrospectively. As discussed earlier, the governance structure did not change a great deal. However, these relatively small changes were very important. The focus on project activities to bring partners together was effective and created a core group of evaluators, implementers, and patient representatives who worked closely together, solving problems and creating efficient working practices to design and deliver projects. Project-specific innovators, technical and clinical experts and service delivery teams then worked with this core set of partners as required on a project-by-project basis.

There were challenges agreeing collaboration agreements with innovators and partners; this took longer than expected. Centralised legal and information governance advice was requested from the national Test Bed team. Funding was provided for legal support for the collaboration agreements and IG issues. Central support was provided but this took time to become established and of value. When advice was provided, it was not always proportional to the problem or accounting for important local contexts.

9.6.1.2 Implementation and uptake

9.6.1.2.1 What were the barriers to effective delivery (and uptake of technology/ services) and how were they overcome?

1. When collaboration began, it was clear that there were differing opinions about the intentions of the national Test Bed programme that were, in part associated with a lack of clarity in the aims and objectives of the programme and how stakeholders interpreted these differently

2. Across stakeholder groups, there was not an agreed perception about the type of evidence that the evaluation was supposed to provide. There were broadly two types of evidence considered; evidence from statistical quantitative inquiry (e.g. difference in difference techniques) and economic evaluation requested by the national Test Bed team, and evidence of acceptability and usability of the technology interventions being the main concern of some innovators. This was further compounded by delays to the commissioning of the national evaluation team, part way through the programme.
3. Third, some technology innovations were at the development stage and were not completely ready to be deployed as fully functioning products or solutions (e.g. Asthma, Diabetes projects). The programme created an opportunity for these prototype technologies to be tested out with real users in real life situations using pragmatic methods of data collection. However, this was in contrast with the intentions of some stakeholders (e.g. Kinesis and NHSE) that expected economic and cost effectiveness evaluation.

9.6.2 Impact evaluation findings

9.6.2.1 Stakeholder benefits

9.6.2.1.1 What have the benefits to innovation partners been of engaging with the NHS as part of the Test Bed programme?

Some innovators learned an enormous amount from the PPP Test Bed. Their technology was modified along the way and in parallel with they were able to build market-ready products. The learning from the PPP Test Bed informed technology design and as a result all innovators developed improved iterations of their technology.

9.7 Conclusions and implications

9.7.1 Programme design

The main characteristics of the programme were interdisciplinary working and cross-organisational partnership between NHS providers, commissioners, University, industry, and patient groups to implement different innovative health technologies in NHS settings as a new way of thinking and working with different stakeholders. However, some respondents considered the original aims to be overly ambitious, and the findings revealed changes to the original ambitions, aims and intentions of the programme.

Due to encountering barriers to the aims and objectives of the programme, the initial design around the coordinated and integrated implementation of combinatorial technologies, for multiple health conditions, rolling out to Sheffield City Region and a big data sharing centre were revised. The main programme aims changed to test out individual interventions supported by technology in real world situations with potential target users, whilst continuing the original ambition in a scaled-down fashion.

The challenge regarding progress of the command centre was a result of lack of commissioner commitment given the potential scale of investment and evidence required for benefits. The Digital Care Homes project was developed as a demonstrator of potential for the control centre. A large challenge was not being able to pay innovators when requests were made of them (e.g. making data available to feed into centralised NHS system (command centres or similar)). This limited the extent to which they were able to respond to requests.

In addition, the variable and emergent expectations from different stakeholders and the existence of different drivers, values and different organisational cultures of working led to different understanding of the programme, and different perceptions of how the programme should be designed, implemented and evaluated.

These interrelated factors created a complex situation where the design of the programme evolved and shifted over time to be more responsive to the emerging context. Moreover, the various partners developed improved ways of working together effectively over time to coordinate activities and design

mutually agreeable interventions. This inevitably involved developing processes, shared knowledge and relationships to manage conflict and attain compromise.

For instance, the implementation team were motivated to deliver the intervention to large numbers of users in ways that would be acceptable to service providers, some innovators were concerned with having their technology adopted by the NHS, and the evaluators were keen to gather high quality evidence. There were also variable responses to the advice of the public and patient feedback, which was not always in accordance with innovator's aims. There were also tensions between the types of (research) evidence that was being requested by the funders and the evaluation evidence that was considered attainable and locally useful to inform decisions around adoption and spread.

9.7.2 Programme set up

Due to the short timescale of the programme, the set-up was faster than would have been ideal. It was aided by regular communication between different stakeholders and an established management structure and governance systems. Flexibility was a necessary feature of the set up and implementation. However, it took time for partners to understand the implications of changes on each other's work and to manage to collaborate efficiently.

Many of the projects were characterised by various issues related to lack of readiness; whether this was in terms of market readiness of the technology, readiness for good quality data collection, or organisational readiness for implementation. More time for planning and achieving a mutually agreed state of readiness amongst all partners prior to roll-out would have been beneficial.

9.7.3 Programme implementation

Despite all the challenges, the implementation of the programme was associated with a great amount of learning that reflects the emerging and unpredictable nature of the implementation process and its unintended outcomes. There were a number of outputs designed to address the uncertainty and challenges of the programme. For instance, the implementation team developed a formal process for identification, implementation, evaluation and uptake of innovations. The evaluation team produced an evaluation design checklist and a guide to differentiating research from evaluation approaches.

Two key areas that affected implementation were management and governance of the programme and projects and recruitment of sites and individuals. The following are the lessons learned as reported by the PPP Test Bed PMO.

In April 2018 the PPP Test Bed PMO reported a series of management lessons learned. The speed and scale of the programme required a rapid and agile approach, but which also worked within appropriate frameworks and regulations. The pace of change required was not always compatible with information governance (IG), organisational governance and ethics approvals. IG is everybody's concern (not just for the experts) but there can be knowledge gaps. Tight timelines also required definite deadlines to get decisions made (clinical and scope) and clear decisions about what is in scope or out of scope. Solutions had to be created and 'work arounds' designed. There were difficulties in managing a programme structure to ensure that it was agile but did not derail the evaluation by making critical changes. Innovators found a greater challenge than anticipated in providing "market ready" devices, which added delays to timelines.

Turnover of staff and a time-limited programme required additional support, flexible approach to resource use and knowledge transfer plans to reduce impact of personnel rotation. Ultimately, high staff turnover, secondments and fixed term contracts were reported to negatively affect the project delivery. Regarding record-keeping, a central repository allowed recording of information historically to show the journey of the programme and the importance of programme rigour and discipline regarding the completion of action logs and risks assessments was recognised. There was a trade-off in managing depth vs breadth with the ambition of the programme and the number of projects.

It was a positive experience, having enough space to display the technology and to bring people for demonstrations (there was an 'Innovation Hub' space provided for the programme at RHH).

The PPP Test Bed PMO reported their recruitment lessons learned. They found that often patients who might benefit most from support are least engaged or able to use technology. They also discovered that people with long-term conditions have developed ways of managing, and are often resistant to change, or have difficulty adapting to new ways.

A great deal of mistrust or suspicion about how data might be used by innovators was experienced. It was considered useful to spend time face-to-face with service-delivery staff to help them to understand the intervention and create relationships with them; this can help to gain their support to increase patient recruitment. Word of mouth recommendations and advice should not be underestimated as a mechanism to increase recruitment numbers.

9.7.4 Partnership working and engagement

Partnership working was characterised by interdisciplinary and cross-organisational working. An array of issues had to be overcome for the successful implementation of the programme including communication difficulties and clarity around arriving at shared terminology, lack of knowledge of other disciplines, the physical location of stakeholders, and limited time to build up relationships.

Whilst the core partners involved in the delivery of the programme had the opportunity to develop predictable partnership working, partnership working with the innovators and service delivery teams involved in specific projects was very individualised and context reliant depending on the type of the setting in which the technology was deployed and the views and values of those who were involved.

The PPP Test Bed PMO reported their lessons learned regarding the functioning of partnership working. A strong collaborative environment is required for testing: changes in innovator/partner commitment need to be monitored and have rapid contingency plans in place. In order to accomplish this, it is important to understand and manage partner motivations on an ongoing basis; stakeholder management requires collaborative rather than technical leadership. It is also important to keep sharing the vision; it is not static and is needed to ensure partners understand their role.

The process to agree content and sign collaboration agreements created delays and could have been carried out sooner in the process. Not all partnerships have been equal, this is to be expected and managed accordingly, but can create frictions between partners.

Innovators felt there was a lack of clarity on procurement frameworks, which contributed to difficulties managing innovators' expectations. Related to this, better due diligence with innovator partners can help in setting clearer expectations and requirements, especially to work collaboratively.

Better links were required at central level with NHS Digital and other organisations. Improved awareness of other initiatives and similar activities across the NHS would help to identify synergies and apply economies of scale.

9.7.5 Engagement report provided by PPP Test Bed PMO

The PPP Test Bed PMO reported their engagement lessons learned. These included ensuring that stakeholders are engaged as an ongoing process using week by week planning and activity. The use of the voice of the patient to influence the programme scope and feedback to innovators was important. In some cases, it was considered that it might have been better for users/patients to select the technology not clinicians/organisations. This comment relates to the push vs pull uptake of technology. Engagement workshops were reported to be helpful for the identification of unmet needs. However, the evaluation team noted that in examining the specifications arising from these events, the limitations regarding the available technology and feasibility of implementation meant that solutions were not necessarily an ideal fit for identified problems.

Backfilling clinical time is essential to support clinical engagement. The early identification of clinical leaders/champions can make a big difference to the projects. However, ongoing difficulties were identified in getting clinical steer and decisions. The importance of co-production approaches was recognised, promoting engagement to 'learn and listen', not to 'tell and sell'. Ultimately, the Test Bed

programme did not operate in isolation, engagement with the wider health and care system was crucial to find the links, alignment and identify pathways.

9.7.6 Engagement report provided by Healthwatch Sheffield

The report provided by Healthwatch Sheffield concluded the following key points that were derived from consultation with service-users and carers, project champions and the Advisory Group.

Inconsistency in patients' digital skills and attitudes to technology were considered to influence who would participate in using patient-technologies and one of greatest challenges was thought to be effectively implementing change within services where limited staff capacity and resistance to change were likely barriers. A patient's mental health was identified as influencing their ability to self-manage and viewing care of patients with multiple health conditions holistically rather than focusing on one condition in isolation was thought to be important, but it was felt that both factors could have been more fully considered in the design of the projects.

There were concerns regarding who exactly could access patient data and whether it would be stored securely, and people questioned whether innovators could be trusted with their data. In general, people had no issue with the idea of data being shared within the NHS as this was thought to benefit them and potentially the wider patient population. The reliability of a care-coordination centre was questioned, and the importance of having something in place as a back-up was emphasised. Understanding the language of digital and presenting digital health information in a way that everyone could fully understand was viewed as being a major challenge in the programme. Personalised verbal explanations were thought to be vital in addition to written materials when explaining how to use a device or describe what would subsequently happen to a patient's data.

It was widely accepted that technology could bring improvements in health and care, but fear of technology was thought to be a major factor in preventing people from engaging with technology to manage their health. Fear of breaking technology and not knowing what to do if something goes wrong were also thought to be barriers to using technology. The importance of personalised face-to-face training and providing ongoing technical support were thought to be useful in helping patients overcome these barriers. There was a common perception that older people don't want to use technology and lacked the required skills. However, some reported using technology regularly. Additionally, when discussing digital health with older people who said they didn't possess digital skills, there was a willingness to learn if the right support was available. Older people supporting people of a similar age to learn digital skills was thought to be a useful approach, and patient stories of success were thought to play an important role in encouraging patients to try using digital to manage their health, regardless of age.

9.7.7 Key principles for partnership working and engagement

The main learning points for partnership working and engagement and can be summarised as the following principles:

1. **Use of co-production from the start:** Applying a co-production approach by engaging all stakeholders especially innovators and patient groups in partnership with health professionals, managers and evaluators from the start.
2. **Identify problems and match with available technology:** Ongoing effort, promotion, identification of technology and engagement with service providers, patients, commissioners, innovators, researchers and other stakeholders to identify opportunities for innovation
3. **Appreciate the difference between "trial-technology" and "production technology":** Understand whether or not technology is ready for wide-scale implementation before being included in projects
4. **Evaluation evidence:** Create clarity and mutual understanding about the evaluation approaches and the type of evidence required
5. **Consider programme sustainability:** Considering sustainability, that is, to maintain all the interdisciplinary and cross-organisational learning and use this for future implementation

9.7.8 Evaluation integration

A number of key challenges to evaluation within a Test Bed programme were recognised, and whilst several of these were addressed or mitigated throughout the life of the programme, others resulted from elements of the overall design of the programme that were not amenable to change and were therefore not easily addressed.

- Speed, and number of implementations
- Constant evolution of interventions based on what was able to be successfully implemented and what wasn't in real world testing (requiring minor, major or complete re-design of evaluation proposals)
- Required time and relationship building:
 - Variable focus on programme theory and expected main outcomes
 - Integrated co-design is important to achieve best results, and requires a shared understanding of limitations and possibilities of evaluation
 - Reasonableness of expectations for evaluation outputs (especially regarding effectiveness and cost effectiveness)
 - Stakeholders' variable understanding of value and usefulness of different types of evidence
 - Efficiency of governance, oversight and data security processes across organisations
 - Little control over data collection and data management
- Restricted to evaluation rather than research methodologies
- Challenges in obtaining counterfactual evidence
- Level of preparedness of innovators (too prepared with inflexible plans for evaluation, or not well prepared to engage in evaluation)
- Costing of novel technologies that are not available for a specific price on the open market is not always possible or useful

The PPP Test Bed evaluation was characterised by multiple parallel projects and rapid, constant change and seeking opportunities for implementation. This was largely driven by the short timescales of the programme and the pressure from funders to recruit large numbers of participants over a short time. This resulted in a constant re-design process, which was time consuming, used up a great deal of evaluation resources and did not result in any tangible outputs other than a large number of evaluation plans. This process also required very high intensity of communication by email and phone, as well as several weekly and fortnightly teleconferences and meetings.

The majority of challenges for the evaluation were mitigated through close-working, frequent communication and relationship-building between the evaluators, implementation team, service providers and innovators. This took time to establish. Co-location was tried, but this did not result in improved levels of communication. Outside of set meeting times members of the evaluation team felt isolated, they did not have access to colleagues or resources that would help them perform effectively in their role and had to make additional journeys. It was felt that this was creating inefficiencies and was discontinued after a few weeks.

The reliance on collaboration between individuals to address evaluation challenges is an important resource that is built on learning and professional development, linked with effective and trusting working relationships. However, this is a fragile resource, which requires time to develop. The existence of the Test Bed programme management office and implementation team, working alongside a stable evaluation team and a lay advisory group reduced the need to develop these working relationships for each project and vastly increased the efficiency of the programme; making it possible to deliver a large number of projects over a short time.

The speed of implementation and development and short life-span of the projects required evaluation approaches, rather than research, approaches, as the Test Bed was largely carrying out pragmatic testing of delivery approaches. However, there was an expectation from the funders to obtain rigorous effectiveness and cost-effectiveness evidence that is more compatible with a research approach. There

was an incompatibility between the short life-span of projects that were novel and previously untested and the quality of evaluation findings that were expected. Whilst this did provide a motivation for the evaluation team to explore all possible approaches to demonstrate effectiveness and cost effectiveness of interventions, this created tensions and unrealistic expectations. The source of this tension can be seen in the original Test Bed specification, especially when the short length of the programme, pressure from the funders to recruit large numbers of users and the novelty of the interventions are considered:

“The primary aim of the programme is to improve patient outcomes and experience of care at the same cost as, or at a lower cost than, current practice, while helping the economy grow”

Demonstrating improved outcomes and cost savings is a massive challenge for new service-delivery models that are rapidly set up, innately changeable, have short lifetimes, are trying to maximise numbers of users and include novel technologies and combinations of technologies. A key challenge is assuming what might have happened without the intervention (business as usual), when useful counterfactual data are not available and the intervention is evolving. Another key challenge is applying costs to interventions that are ultimately tests, and do not represent how the service-change might be implemented in the longer-term, and where technologies are not available on the open market and therefore have no unit costs. An additional concern to consider is the potential burden of evaluation activities on service-users at the early stages of implementation, prior to longer-term monitoring or audit procedures being developed.

Some innovators were less prepared than others to have their technology evaluated, whilst some entered the programme with firm plans about what they wanted to achieve regarding the production of evidence. Both of these extremes required different approaches and both required time to provide an evaluation response that was agreeable to all partners.

The evaluation team experienced little control over data collection and management; particularly for quantitative data we were reliant on service deliverers, innovators, and data controllers. This created a number of problems relating to completeness, questions around quality, ability to interpret findings and timeliness. Qualitative data collection was often reliant on gatekeepers providing access to respondents. However, owing to good working relationships and frequent communication this was largely unproblematic.

The PPP Test Bed PMO recognised much of the same challenges and ways to address them; indicating how the programme partners developed a shared culture and agreement about effective approaches over the course of the programme.

The PPP Test Bed PMO report their ‘evaluation lessons learnt’ in April 2018 as;

- Co-design the programme and evaluation at the start
- Challenge of academic evaluation required versus real time testing and changes required
- Stakeholders require clear understanding of the purpose of evidence
- Discuss the evaluation process at planning stage to consider critical requirements.
- Integrate the evaluation team with the management team.
- Consider ‘real world’ testing methods, with less academic approach.
- Keep control over data collection, reducing dependence on innovators and clinical teams
- Difficulties integrating service improvement techniques to testing due to resource pressures
- Flexible evaluation methodologies required to adapt approach to real world scenarios

9.7.9 Dissemination

The Test Bed programme had broad reach and a large and varied group of stakeholders (for instance senior national NHS decision-makers, regional commissioners, clinicians and service providers, service-users and carers etc). Whilst some projects depended on dissemination and communication more than others, it was recognised as an important function for the operation of the programme. For instance, awareness-raising amongst clinicians and other stakeholders was a useful way to encourage dialogue and

identify areas where technology solutions can address specific service-delivery or patient needs (e.g. use of QTUG in acute balance clinics).

The PPP Test Bed PMO report their 'dissemination lessons learnt' in April 2018 as:

- Consider the audience from innovators to patients- one size does not fit all
- There is work in progress to lever the spread opportunities from the AHSN
- Challenge the NHS speak – the patients/public voice needs to be valued and heard
- Celebrate small wins along the programme journey
- The programme requires dedicated communications support from the start to avoid delays and maximise impact
- Benefitted from opportunities to speak at events/conferences and spread the learning

9.7.10 Programme scale up and spread

The ambition of the original programme design resulted in a programme which was equally ambitious in its scale. There were a large number of different technologies implemented in numerous settings over a short amount of time. In addition, the programme developed surrounding infrastructure to support commissioning, data sharing, and capacity development. In this regard, the programme demonstrated the potential for this type of inter-organisational, system-focused infrastructure to support and drive the agenda for implementation and testing of health care technology in general as well as carrying out a large number of specific projects in primary, community and secondary care, third sector, social care and private nursing settings.

The short-term funding for the programme has meant that the PMO has mostly been disbanded. However, one of the projects (QTUG in acute balance clinics), is set to continue, providing appropriate ongoing funding can be secured. The Digital Care Home project is a legacy project from wave 1 test bed which is seeking to use qualitative feedback and recommendations from wave one to implement co-design with care home community and test and evaluate over a longer time period, with a larger. Some of the learning will be about usability and acceptability of digital technology in care homes.

Some of the learning and ambitions of the Test Bed programme will be carried forwards into planning for digital strategy in the regional Integrated Care System (ICS). However, it is important to recognise that many of the challenges for the programme were addressed through the individual professional development, learning and relationship-building for collaborative co-production with key partners. A great deal of the programme resources were expended in developing these collaborative partnerships. Developing this infrastructure from scratch was a time-consuming and ultimately costly venture, which will require replication for further projects.

For successful scale-up and spread, it is important to recognise differences between organisational technology, individual person technology and circumstances in which technology can be considered a hybrid of the two. Organisational technology, such as QTUG (digital falls risk assessment) can be classified as organisational technology, which is purchased and implemented as the result of an organisational decision and the key influence of patients is whether or not it is considered acceptable and useful in its application. Personal technology, such as the SOS UK emergency contact app, requires uptake by individuals; it needs to be attractive so that spontaneous uptake occurs. Some of the technology within the PPP Test Bed can be conceptualised as a hybrid of these two approaches, for instance Insulcheck (insulin monitoring) and Teva (asthma medication monitoring) formed an optional part of an established long-term condition management service, but also required uptake from individuals. These differences can be understood to involve the variable reliance on push or pull factors for successful uptake.

9.7.11 Impact

As the result of this collaboration and partnership, the key stakeholders experienced significant amount of learning and managed to understand the cultural gap between different disciplines and sectors and shed light on some of the institutional barriers to the implementation of complex health interventions.

The partners gained unique knowledge and skills needed to undertake such interdisciplinary practice in the real world, rather than in an artificial research-focused setting.

Real-world testing has not to date happened effectively in NHS. This programme has shown how to approach real-world testing, issues, challenges, and how it could support commissioning decisions.

The programme management and governance infrastructure was developed over the lifetime of the programme, specifically in terms of incorporating processes and making decisions with an understanding of the complexity of technology implementation when:

- combined with pathway re-design,
- incorporating co-design with a range of important stakeholders
- and supporting the production of appropriate evidence to support development, spread, sustainability etc.

This learning resulted in, for instance; actionable tools (e.g. considerations of evaluation issues); appropriate forums and means of engagement (e.g. seeking and promoting the views of the users' groups); systematic processes for technology assessment, documentation and administration processes (e.g. memorandum of understanding, data sharing agreements); creative approaches to data sharing and integration.

With regard to the eight key themes that emerged from the five objectives and six work-streams identified in the early planning stage, the process evaluation concludes that the programme has succeeded in making considerable developments in all of these areas:

1. Programme Management & Governance
2. Provide an ongoing platform for testing, refining and scaling-up innovations
3. Re-design pathways, bringing combinatorial technologies and system transformations to support holistic and personalised care
4. Technology Deployment
5. Patient Involvement and Engagement
6. Evaluate the combination of new technologies and service re-designs producing robust and objective results that can be shared and disseminated
7. Embed the culture of transformation and improvement in NHS and other health and care organisations
8. Support co-ordinated decision making across health and care, informed by real-time data and predictive analytics (Intelligence Centre).

9.8 Overall conclusions and implications

Despite the technologies that were tested being considered market-ready, and had often been used in other settings, there was a considerable amount of feedback that was produced. Recommendations for product improvement and development were produced from engagement with the implementation team, front-line staff, service-users and evaluators. As innovators were not aware that their products could be improved or adapted in the ways that were suggested; this would indicate that everyday technology deployment does not always result in useful feedback to innovators. This provision of feedback for improving and adapting digital technology could be considered a crucial function for NHS infrastructure similar to the PPP Test Bed. The unique combination of organisations and functions integrated into a core team seems to be a critical condition for this mechanism to operate.

The NHS should consider the creation of this type of infrastructure to look at a pipeline of technologies appropriately selected on an ongoing basis, which provides a conduit from development to commissioning and implementation in order to create evidence on where and how value is most likely to be created for the NHS.

In the 'background' to this section we summarised how the Test Beds were originally intended to test implementation of combined (combinatorial) innovations, in 'real-world' settings, to understand how

health services might be transformed (rather than just added to). Briefly exploring these concepts, indicates why they might have been chosen as the focus for such a national programme, and helps to contextualise and explain the findings of this evaluation.

Real-World' testing: There is a wealth of evidence concerning the lack of engagement of traditional research approaches with the 'real-world' of service delivery, which leads to difficulties in successfully translating knowledge. Indeed, there is compelling evidence that when promising innovations are scaled-up without the application of an extensive understanding of context offered by recent methodological advances (e.g. in Systems Thinking and the application of Complexity theory in health services research), then any early benefits rapidly dwindle.^{35, 36}

Combinatorial Innovation: The phrase 'combinatorial innovation' is widely accepted to have originated from the work of Hal Varian in 2003,³⁷ in which several works were drawn together to form a coherent conceptualisation of the potential benefits of combining innovations. In this work, the speed of the development of the internet is related to combinatorial innovation and a link is made between combinatorial innovation and periods of economic boom. Other sources list examples of combinatorial innovations such as Gutenberg's printing press, clockwork radio, wheeled suitcases and alarm clocks. However, it could be argued that all innovation builds on elements of established technology and is by its very nature combinatorial.

A key difficulty, when assessing the influence of combinatorial innovation is that we are always dealing with historical examples, and therefore examples that were successful and transformative. The history of innovation is littered with failure, accidental success (discovery) and years of toil (invention), which it could easily be argued are criteria that were not considered as acceptable outcomes by people invested in the national Test Bed programme. The remark attributed to Thomas Edison sums up the effort required to arrive at successful invention:

"I said: 'Isn't it a shame that with the tremendous amount of work you have done you haven't been able to get any results?' Edison turned on me like a flash, and with a smile replied: 'Results! Why, man, I have gotten a lot of results! I know several thousand things that won't work.'"³⁸

Disruptive Innovation: The transformative rather than additive (accretive) potential of innovation is related to the concept of disruptive technology/innovation. These concepts of accretive and disruptive innovation are often considered as opposite poles of implementation approaches. However, it is more likely that both accretive and disruptive elements are factors in implementation of innovations. A key benefit that is cited for a focus on disruptive innovation is reduced cost. However, whilst examples of disruptive innovations that manage to save costs are available; these are likely to be an artefact of the higher likelihood of innovations being scaled-up (and promoted as successes) if they manage to prove cost savings. Equally, there are examples of disruptive innovation that prove to be very costly and unsuccessful; a high-profile example being the NHS National Programme for IT.

Driving forces for the potential cost saving of disruptive innovation include more transparency about costs and the move to Accountable Care Organisations.³⁹ However, the key problem with disruptive innovation in health care is that when complex systems are disrupted, the effects are; unknown, non-linear, uncertain, emergent, adaptive, dynamical and co-evolutionary.⁴⁰ Rather than being able to be defined as a success or failure, there are likely to be variable outcome patterns that are heavily dependent on the context (and contextual factors that change over time).

Accretive innovation, on the other hand, develops from current understanding and has some knowledge to draw upon regarding reasonable assumptions about effects and outcomes. These factors have significant implications for the ability to evaluate disruptive innovation using methods of evidence production that are widely understood and accepted in health service delivery. Trials methods and traditional economic evaluations rely on assumptions about what might have happened without the intervention, which, in a world of increasing attempts at disruption, are increasingly forced to compare one system that has been disrupted in a certain way to a system that has experienced a different type of disruption. New methodologies are developing to cope with these difficulties (e.g. synthetic data & mixed-

method theory-driven evaluation). However, the types of evidence produced might not be well-understood by audiences, and these evaluation approaches need to operate in a rapidly changing environment and can therefore be risky and resource intensive.

It is not understood why these three elements were prioritised within the national Test Bed programme, but we can observe some of the effects.

Firstly, the programme had a very short time-scale, and Real-World testing cannot take place without a large amount of planning and design. This was only possible following the: appointment of staff, development of partner relationships, signing of collaborative agreements, working out solutions to data management and information governance issues, evaluation design, ethics and governance approvals, finding agreeable and appropriate service delivery settings. Combined with the speed of implementation required and the pressure to produce large recruitment figures early on, the PMO had to take opportunities for implementation as they arose and the programme developed a characteristic whereby elements of innovations were being implemented whilst other elements were being designed (or ‘flying a plane, whilst trying to build it’).

These factors introduced instability and reduced the time available for evaluation. The ‘Real-World’ testing that was carried out was therefore heavily focused towards project design, developing efficient ways of collaborative working and exploring implementation approaches.

Secondly, the testing of Combinatorial Innovation was a keystone of the Test Bed programme. However, this approach was translated to the Test Bed sites as requiring the combination of more than one of the Test Bed innovator’s products. This factor was a distraction for the programme, and resulted in novel technologies being combined together, which multiplied the complexity of projects added risk and created organisational conflict and relationship management problems. Examples of where combinations worked to best effect (in terms of demonstrating potential for sustained use) were when single technologies were combined with existing infrastructure to re-design pathways and methods of care, for instance the QTUG assessment being introduced in balance clinics and the care home monitoring linked to the Single Point of Access (SPA) team. The difficulty of combinatorial innovation should not be underestimated and historically has been built on combining established concepts in novel ways, rather than combining novel innovations together. The focus of the National Test Bed programme on the latter was problematic and emphasised the care that is required for this type of endeavour, as described by a practitioner of combinatorial innovation:

“It is important for us new practitioners of combinatorial innovation in health to bear in mind that for the technique to work, you need to have a way to separate out the combinations that bind from those that don’t. This might translate roughly into doing a thorough assessment and making sure you are asking the right questions as to what works.”⁴¹

Thirdly, the national programme focus on disruptive, rather than additive innovation encouraged increased complexity and therefore increased characteristics of complex systems. The purpose of evaluation of complex systems is to improve understanding so that; chaotic elements become complex, complex elements become complicated, and complicated elements become simple. Through this process it is possible to begin to increasingly understand and predict effects. The disruption of complex healthcare systems is therefore incompatible with the production of the types of evidence that were being requested by the national programme (effectiveness and economic efficiency); this created tensions regarding the differences between evaluation approaches that were considered locally appropriate and those that were considered necessary to satisfy the national programme (i.e. the types of evidence considered appropriate for press releases, Ministers and senior civil servants). This challenge was alleviated to an extent with the commissioning of an experienced national evaluation team that had an understanding of the feasibility issues involved.

A final point to note is the extent of the gap between testing a technology and testing an innovation in a complex system. For instance, combinatorial innovation does not simply require putting two technologies together in a compatible configuration, but implementing a whole system within other complex systems

and incorporating a sustainable business model. The innovator 'match-making' process resulted in a range of technologies selected prior to recognising problems to address and environments to fit into. The PPP Test Bed conducted extensive stakeholder consultations, including workshops using the CLEAR IDEAS methodology⁴² to reconcile available technologies with service provider and patient requirements. However, with only a limited range of innovations, with very specific functionality to ultimately choose from, there was a recognisable gap between identified problems, suggested solutions and available technology.

9.9 Recommendations

Programmes like the PPP Test Bed should be implemented using systems change approaches, where complexity concepts of unpredictability and emergence should be considered and applied. Therefore, a combination of complex, 'real-world' evaluation approaches and traditional research approaches for the assimilation of required evidence should be considered in the context of developing business models and understanding the requirements of commissioners and decision-makers. This will require a combination of rapid evaluation and longer-term research for the production of appropriate evidence at appropriate times.

For the successful implementation of programmes such as the PPP Test Bed, partners need allocated time and resources to build the relationships and to reach a mutual understanding of each other's intentions and expectations and methodological approaches. They need to share their experiences and develop a co-production approach where the process is shaped and the outcomes are produced as a result of engagement and integration.

Strategically, health and social care commissioning and service delivery partnerships in collaboration with higher education, service-user groups and strategically important innovators could decide to sustainably support collaborative infrastructure for the selection, implementation, testing and scaling of health technologies. This could improve the efficiency of this type of work and provide a context for collaborative learning and capacity development.

The key issue to address is the identification of areas where there are technologically supported solutions to identified problems, which have strong programme logic and a potential for improving patient outcomes at the same or lower cost. Some consideration should be given to testing mechanisms for this problem identification process, as the first critical stage to designing effective patient pathways. This requirement could form the focus of a combined approach involving the NHS and NIHR.

References

1. NICE. *QTUG for assessing falls risk and frailty (Medtech innovation briefing [MIB73])*. 2016. URL: <https://www.nice.org.uk/advice/mib73> (accessed 22nd November 2017).
2. Braun V, Clarke V. Using thematic analysis in psychology. *Qualitative Research in Psychology* 2006;**3**:77-101. <https://doi.org/10.1191/1478088706qp0630a>
3. Coffey A, Beverley H, Paul A. Qualitative Data Analysis: Technologies and Representations. *Sociological Research Online* 1996;**1**:1-12. <https://doi.org/10.5153/sro.1>
4. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, *et al*. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011;**20**:1727-36.
5. Al-Janabi H, Flynn TN, Coast J. Development of a self-report measure of capability wellbeing for adults: the ICECAP-A. *Qual Life Res* 2012;**21**:167-76.
6. Keetharuth AD, Brazier J, Connell J, Bjorner JB, Carlton J, Buck ET, *et al*. Recovering Quality of Life (ReQoL): a new generic self-reported outcome measure for use with people experiencing mental health difficulties. *The British Journal of Psychiatry* 2018;**212**:42-9.
7. Hibbard JH, Mahoney ER, Stockard J, Tusler M. Development and testing of a short form of the patient activation measure. *Health Serv Res* 2005;**40**:1918-30.
8. Hill KD, Schwarz JA, Kalogeropoulos AJ, Gibson SJ. Fear of falling revisited. *Arch Phys Med Rehabil* 1996;**77**:1025-9.
9. Mulhern B, Rowen D, Brazier J, Smith S, Romeo R, Tait R, *et al*. Development of DEMQOL-U and DEMQOL-PROXY-U: generation of preference-based indices from DEMQOL and DEMQOL-PROXY for use in economic evaluation. *Health Technology Assessment (Winchester, England)* 2013;**17**:v.
10. StataCorp. Stata Statistical Software: Release 15. In. College Station, TX: StataCorp LLC; 2017.
11. Gillespie LD, Robertson MC, Gillespie WJ, Sherrington C, Gates S, Clemson LM, *et al*. Interventions for preventing falls in older people living in the community. *The Cochrane Library* 2012.
12. Scuffham P, Chaplin S, Legood R. Incidence and costs of unintentional falls in older people in the United Kingdom. *J Epidemiol Community Health* 2003;**57**:740-4.
13. Greene BR, Redmond SJ, Caulfield B. Fall risk assessment through automatic combination of clinical fall risk factors and body-worn sensor data. *IEEE journal of biomedical and health informatics* 2016;**21**:725-31.
14. Barry E, Galvin R, Keogh C, Horgan F, Fahey T. Is the Timed Up and Go test a useful predictor of risk of falls in community dwelling older adults: a systematic review and meta-analysis. *BMC Geriatr* 2014;**14**:14.
15. Public Health England. *A Return on Investment Tool for the Assessment of Falls Prevention Programmes for Older People Living in the Community*. 2018. URL: <https://www.gov.uk/government/publications/falls-prevention-cost-effective-commissioning> (accessed 08 July 2018).
16. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. *Value Health* 2010;**13**:509-18.
17. Curtis L, Burns A. Unit Costs of Health & Social Care 2017: PSSRU; 2017.
18. NHS Improvement. *NHS reference costs 2016 to 2017*. 2017. URL: <https://improvement.nhs.uk/resources/reference-costs/> (accessed 22nd November 2017).
19. Kind P, Hardman G, Macran S. *UK population norms for EQ-5D*; 1999.
20. Barnes CB, Ulrik CS. Asthma and Adherence to Inhaled Corticosteroids: Current Status and Future Perspectives. *Respir Care* 2014;**60**:455-68. <https://doi.org/10.4187/respcare.03200>
21. Engelkes M, Janssens HM, de Jongste JC, Sturkenboom MCJM, Verhamme KMC. Medication adherence and the risk of severe asthma exacerbations: a systematic review. *Eur Respir J* 2014;**45**:396-407. <https://doi.org/10.1183/09031936.00075614>
22. Meißner A, Schnepf W. Staff experiences within the implementation of computer-based nursing records in residential aged care facilities: a systematic review and synthesis of qualitative research. *BMC Med Inform Decis Mak* 2014;**14**:54.
23. McCall K, Keen J, Farrer K, Maguire R, McCann L, Johnston B, *et al*. Perceptions of the use of a remote monitoring system in patients receiving palliative care at home. *Int J Palliat Nurs* 2008;**14**:426-31.
24. Kruse CS, Mileski M, Alaytsev V, Carol E, Williams A. Adoption factors associated with electronic health record among long-term care facilities: a systematic review. *BMJ Open* 2015;**5**. <https://doi.org/10.1136/bmjopen-2014-006615>
25. Kruse CS, Mileski M, Vijaykumar AG, Viswanathan SV, Suskandla U, Chidambaram Y. Impact of Electronic Health Records on Long-Term Care Facilities: Systematic Review. *JMIR Medical Informatics* 2017;**5**:e35. <https://doi.org/10.2196/medinform.7958>

26. Care Quality Commission. *The state of adult social care services 2014 to 2017*. 2017. URL: https://www.cqc.org.uk/sites/default/files/20170703_ASC_end_of_programme_FINAL2.pdf (accessed 07 July 2018).
27. NICE. *Guide to the methods of technology appraisal*. National Institute for Health and Care Excellence; 2013. URL: <https://www.nice.org.uk/process/pmg9/chapter/foreword> (accessed 07 July 2018).
28. NICE. *Guide to the methods of technology appraisal*. 2013. URL: <https://www.nice.org.uk/process/pmg9/chapter/foreword> (accessed 22nd November 2017).
29. Drummond M, Sculpher M, Claxton K, Stoddart G, Torrance G. *Methods for the Economic Evaluation of Health Care Programmes (4th Edition)*: Oxford University Press; 2015.
30. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. *Methods for the economic evaluation of health care programmes*: Oxford university press; 2015.
31. Pattison A, Vernon M. THE MODIFIED EARLY WARNING SCORE (MEWS) DOES NOT PREDICT MORTALITY IN COMMUNITY DWELLING NURSING HOME RESIDENTS. AGE AND AGEING, abstract no. 81, p. II32-II.
32. NHS Improvement. *NHS reference costs 2016/17*. 2017. URL: <https://improvement.nhs.uk/resources/reference-costs/> (accessed 07 July 2018).
33. England N. Next steps on the NHS five year forward view. 2017. In; 2017.
34. NHS England. *Real world testing of 'combinatorial innovation': A global invitation to innovators*. 2015. URL: <https://www.england.nhs.uk/wp-content/uploads/2015/03/test-bed-prospectus.pdf> (accessed 30 June, 2018).
35. Ioannidis JP. Contradicted and initially stronger effects in highly cited clinical research. *JAMA* 2005;**294**:218-28.
36. Parry GJ, Carson-Stevens A, Luff DF, McPherson ME, Goldmann DA. Recommendations for evaluation of health care improvement initiatives. *Acad Pediatr* 2013;**13**:S23-S30.
37. Varian HR. Innovation, components and complements. *University of California, Berkeley, October* 2003.
38. Dyer FL, Martin TC. Edison: His Life and Inventions. In: New York: Harper & Brothers; 1910:pp.615-6.
39. Zimlichman E, Levin-Scherz J. The coming golden age of disruptive innovation in health care. *J Gen Intern Med* 2013;**28**:865-7.
40. Patton MQ. Developmental evaluation: Applying complexity concepts to enhance innovation and use: Guilford Press; 2010.
41. Videler T. *Combinatorial Innovation: What can we learn from science?*; 2015. URL: <https://www.linkedin.com/pulse/combinatorial-innovation-what-can-we-learn-from-science-videler> (accessed 30 July, 2018).
42. Birdi K. *CLEAR IDEAS: a clearer view on innovation*. 2010. URL: <http://clearideas.sheffield.ac.uk/> (accessed 30 July, 2018).