Evidence overview: Devices for remote continuous monitoring of Parkinson's disease

This overview summarises the main issues the diagnostics advisory committee needs to consider. It should be read together with the final scope and the diagnostics assessment report.

1 Aims and scope

Current practice for monitoring the motor symptoms of people with Parkinson's disease includes using validated questionnaires, history taking and clinical observation. It can be difficult to assess the symptoms of people with Parkinson's disease who have difficulty communicating, remembering or recording their symptoms. Examination at a single point in time, for example at a clinical appointment, may over- or underestimate symptom severity or incidence, given that motor fluctuations can vary over time.

Devices that can monitor and record symptoms of Parkinson's disease could identify people who could benefit from changes to their care. Motor symptoms include dyskinesia (involuntary movement), bradykinesia (slowness) and tremor; periods of immobility may indicate non-motor symptoms. By objectively measuring these symptoms over several days, the technologies may more accurately estimate a person's symptoms and help to inform medication decisions.

Levodopa is recommended as the first-line treatment for people with Parkinson's disease whose motor symptoms affect their quality of life. However, levodopa itself can cause motor-symptom side effects such as dyskinesia, particularly after long-term use (see section 3.1 of the final scope), so the benefits of treatment must be balanced against the potential harms. Dopamine agonists, monoamine oxidase type B (MAO-B) inhibitors or catechol-O-methyl transferase (COMT) inhibitors are offered as additional

treatment for people with dyskinesia or motor fluctuations despite optimal levodopa therapy. Clinical experts highlighted that decisions about changes to medication can be complex. Levodopa effectiveness can also wear off, with the duration of benefit getting shorter over time. Response fluctuations are characterised by variations in motor performance, with normal function during the 'on' period, and weakness and restricted mobility during the 'off' period. People with advanced Parkinson's disease can also have alternative treatments such as intermittent apomorphine injection, continuous apomorphine infusion and deep brain stimulation.

Better-informed treatment decisions could lead to improved quality of life. Reduced motor symptoms could also reduce falls and hip fractures. The technologies could also help improve communication between people with Parkinson's disease and clinicians when discussing symptoms and potential changes to care.

The technologies may also allow remote monitoring of Parkinson's disease. This could help to alleviate the stress and anxiety of attending clinical appointments. Measurement of symptoms could also reduce the length and number of clinical appointments, thereby freeing up NHS resources.

The aim of the assessment was to review existing evidence on the potential clinical and cost effectiveness of devices for remote continuous monitoring of Parkinson's disease.

Decision question

Do devices for remote continuous monitoring of Parkinson's disease represent a clinically and cost-effective use of NHS resources?

Populations

People with Parkinson's disease.

If data permits, the following subgroups may be considered:

- current treatment and treatment options
- people with advanced Parkinson's disease
- people with communication barriers
- family background.

Interventions

The assessment includes only wearable remote monitoring devices that produce results with no or limited input from the user. All technologies assess, at least, bradykinesia and dyskinesia. Five remote monitoring devices with regulatory approval (or in the process of seeking it) were included:

- Kinesia 360 motor assessment system (Great Lakes Neurotechnologies) – worn on the wrist and ankle
- KinesiaU motor assessment system (Great Lakes Neurotechnologies) – worn on the wrist
- PDMonitor (PD Neurotechnology) – 5 sensors worn on each of the wrists, ankles and on the waist
- Personal KinetiGraph (PKG) Movement Recording System (Global Kinetics Corporation) – worn on the wrist
- STAT-ON (Sense4Care) – worn on the waist.

The devices use algorithms to convert data into summarised reports. At scoping, the clinical experts highlighted that the technologies could be integrated into care pathways in several ways. The scope specified that the technologies should be assessed for use in addition to current care for monitoring motor and non-motor symptoms, when used:

- for all review appointments
- for a subset of review appointments (for example if motor fluctuations are not being adequately managed; details are in section 4 of the final scope)

- without a review appointment (that is, results being considered by a healthcare professional without an appointment with the person with Parkinson's disease; details are in section 4 of the final scope).

Further details on the technologies are in section 2.2 of the final scope.

Comparator

The comparator is the assessment of motor and non-motor symptoms using clinical judgement (based on information including clinical history and patient diaries), which may include rating scale tools and activity trackers.

The Unified Parkinson's Disease Rating Scale (UPDRS) and Hoehn and Yahr scale can be used to describe and assess symptoms related to Parkinson's disease. See the glossary for further detail.

Healthcare setting

Community or secondary healthcare.

Further details, including descriptions of the interventions, comparator, care pathway and outcomes, are in the final scope for devices for remote continuous monitoring of Parkinson's disease.

2 Clinical effectiveness evidence

The external assessment group (EAG) did a systematic review to identify evidence on devices for remote and continuous monitoring of Parkinson's disease. Find the full systematic review results in tables 51 to 66 on pages 179 to 224 in the diagnostics assessment report.

Overview of included studies

There were 84 publications that met the inclusion criteria for the systematic review. Of these, 7 related to studies that are ongoing and have no results. Of the 77 publications with results, 57 evaluated the PKG, 15 the STAT-ON, 3 the Kinesia 360, 1 the KinesiaU and 1 the PDMonitor. No study directly

compared 1 device with another. Seven studies were done in the UK. The EAG categorised the outcomes reported in the publications as:

- diagnostic accuracy: reporting whether the devices can predict symptoms and outcomes, or predict the need for medication change
- association between device outputs and clinical measures
- intermediate impact of monitoring: how devices affected clinical decision making; for example, changes in treatment or adherence
- clinical outcomes: how devices affect outcomes for patients
- patient and carer opinions
- healthcare professional opinions.

There is further detail on the EAG's definition of each category in section 3.2 of the diagnostics assessment report on pages 41 to 42.

Three studies reporting clinical outcomes compared device use with standard clinical practice: 1 on the PKG (Woodrow et al. 2020) and 2 on the Kinesia 360 (Isaacson et al. 2021 and Peacock et al. 2021). There is further detail on these studies in the device sections below.

The EAG commented that there is no clearly established reference standard for measuring Parkinson's disease symptoms beyond clinician and patient assessment. This is unlikely to be a perfect reference standard, and a possible benefit of the devices is that they may more accurately evaluate symptoms than patient recall or clinical opinion. The EAG highlighted that this cannot be easily determined from a diagnostic accuracy study.

Because of the substantial diversity in study populations, conduct and outcomes reported, the EAG did not combine data from studies in meta-analyses.

2.1 Kinesia devices

Two Kinesia devices were considered. The Kinesia 360 uses sensors worn on the wrist and ankle and the KinesiaU is a smartwatch.

Kinesia 360

Clinical outcomes

The EAG identified 2 randomised control trials (RCTs) on the Kinesia 360 (Isaacson et al. 2021, Peacock et al. 2021) The EAG was concerned about the risk of bias, assessed using the Cochrane risk of bias tool, for Isaacson et al. (2019) based on the randomisation process. It also assessed Peacock et al. (2021) as having a high risk of bias, mainly because of the randomisation process. The EAG commented that neither study used the device during routine clinical visits.

Isaacson et al. (2019) was a 12 week pilot RCT (n=39) done in the USA. It compared the Kinesia 360 plus standard care (using the Unified Parkinson's Disease Rating Scale [UPDRS] III; motor examination) with standard care alone for optimising transdermal rotigotine (a dopamine agonist) dosage. The Kinesia 360 was worn throughout the day on at least 2 consecutive days in weeks 1, 2, 3, 4 and 11. The study included people with Parkinson's disease whose motor symptoms were insufficiently controlled.

At week 12 there was a statistically significant reduction (that is, improvement) in UPDRS II score (activities of daily living) in the Kinesia 360 arm when compared with baseline scores; there was a slight increase in score in the standard care arm (-2.1 for the Kinesia 360, 0.5 for standard care; p=0.004). There were non-significant reductions in UPDRS III (motor examination; Kinesia 360: -5.3, standard care: 1.0, p=0.1) and no significant changes in Parkinson's Disease Questionnaire 39 (PDQ-39) in either arm. There was an increase in rotigotine dosing and number of dosage changes in the Kinesia 360 arm compared with the standard care arm.

Peacock et al. (2021) was suspended because of COVID-19 (n=25). It compared telehealth follow-up care using data from the Kinesia 360 with in-person follow-up care in Canada. The study included people with Parkinson's disease who had tremor or dyskinesia identified as a treatment target at their

most recent visit. The average change in PDQ-39 summary index score from baseline to completion (primary outcome) was -4.7 points in the telehealth group (95% confidence interval [CI]: -10.2 to +0.7) and +0.9 (95% CI: -3.6 to +5.5) in the control group. Secondary outcomes were not significantly different between groups. Repeat measurement of Movement Disorder Society – UPDRS (MDS-UPDRS) III (motor examination) was not completed because of the suspension.

The EAG also identified a cohort study (Pulliam et al. 2018) on a device that the EAG judged to be equivalent to the Kinesia 360 (index test not explicitly stated). Full details of the studies are in table 21 on page 72 of the diagnostic assessment report.

Diagnostic accuracy

The EAG identified 1 study on the accuracy of the Kinesia 360. Pulliam et al. (2018) reported sensitivities of 74%, 80% and 90% for dyskinesia, bradykinesia and tremor, respectively, for the Kinesia 360 when compared with clinical assessment by video. Respective specificities were 85%, 66% and 80%, and overall area under the curve (AUC)s were 86%, 82% and 89%.

Association outcomes

No studies reporting association outcomes with the Kinesia 360 were identified.

Impact of monitoring

No studies reporting on the impact of results on decisions about care were identified for the Kinesia 360.

User opinions

Peacock et al. (2021) reported patient opinions of the Kinesia 360 using telehealth or follow-up usual care. In the telehealth group, 54% of people reported feeling comfortable or very comfortable and 8% were uncomfortable or very uncomfortable using motion sensors; 46% would have preferred to be

in the usual care group, 8% would not and 46% were undecided. In the usual care group, 8% of people would have preferred to be in the telehealth group, 67% would not and 25% were undecided.

KinesiaU

Clinical outcomes

One study was identified for the KinesiaU (Hadley et al. 2021). The EAG considered the study at high risk of bias because of limited reporting and a lack of blinding. This cohort study (n=16) in the USA assessed the use of the KinesiaU for 3 days in people undergoing therapy changes for Parkinson's disease. In the trial, 14 people successfully completed their KinesiaU readings, and 2 were unable to because of user difficulty or technical issues. Of the 13 people who returned for follow up, 8 people had improvements with their various new therapies, and 5 people resumed their previous therapy (because of side effects or a lack of benefit).

Diagnostic accuracy

No studies reporting diagnostic accuracy for the KinesiaU were identified.

Association outcomes

No studies reporting association outcomes for the KinesiaU were identified.

Impact of monitoring

No studies reporting the impact of results on decisions about care were identified for the KinesiaU.

User opinions

In a cohort study (Hadley et al. 2021), 88% of people agreed that the KinesiaU system was easy to understand and use, but only 44% agreed that they would continue to use the system if it was available (25% were neutral and 31% disagreed). Full details of the clinical outcomes and user opinions included in the KinesiaU study are in tables 21 and 22 on pages 72 to 74 of the diagnostic assessment report.

2.2 PDMonitor

Clinical outcomes

The EAG identified 1 case series describing the use of the PDMonitor in 2 people with Parkinson's disease (Tsamis et al. 2021). The publication described alleviation of various symptoms with PDMonitor use. Because of the sample size, the EAG did not consider this study further.

Diagnostic accuracy

A cohort study (Kostikis et al. 2020; n=30) was reported only as a conference abstract. It found that the PDMonitor accurately detected and estimated the severity of bradykinesia in the arm, dyskinesia, gait impairment, wrist tremor, leg tremor and freezing of gait, compared with clinical assessment by an expert physician using UPDRS III (motor examination) and Abnormal Involuntary Movement Score (AIMS).

Association outcomes

The association outcomes reported in Kostikis et al. (2020) are described in the PDMonitor diagnostic accuracy section.

Impact of monitoring

Tsamis et al. (2021) also described an increase in treatment dosage as a result of PDMonitor use.

User opinions

No studies reporting patient, carer or clinical opinions about the PDMonitor were identified.

2.3 Personal KinetiGraph

Clinical outcomes

Clinical outcomes were reported in 10 studies, 4 of which were conference abstracts. The EAG stated that the key evidence on the clinical value of the

PKG came Woodrow et al. (2020). Individual patient data from this study is available.

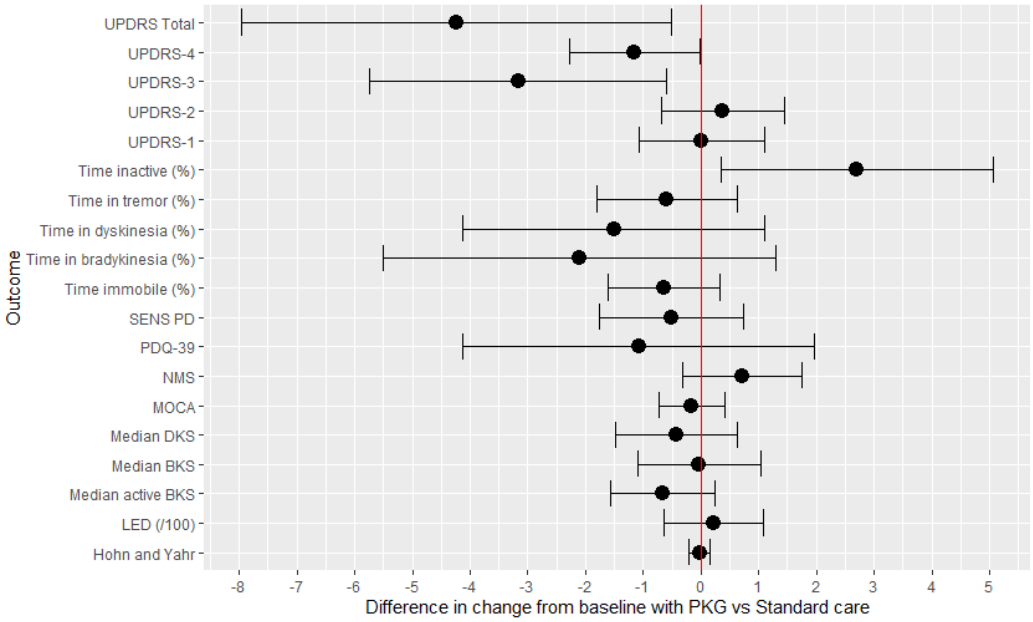
In Woodrow et al. (2020), PKG in addition to standard care was compared with standard assessment methods alone. The study was done in Australia and included 154 people with Parkinson's disease (75 in the PKG arm, 79 in the control arm). The PKG was used for people in both trial arms, but only the healthcare professionals in the intervention arm had access to the PKG outputs.

The EAG considered the Woodrow trial to be a quasi-randomised cluster trial because the clinics that were assigned to be in the PKG arm were done so generally based on their experience of using the PKG. People with Parkinson's disease were assigned to clinics based on location and convenience. The EAG assessed the risk of bias using the Cochrane risk of bias tool. The trial was judged to be at low risk of bias for most categories, with the main risk coming from the fact it was not randomised. The EAG commented that the patient characteristics in the 2 trial arms were not substantially imbalanced. The study measured changes in MDS-UPDRS scores between the first and last visit, based on up to 5 consultation visits (5 weeks between visits) until motor symptoms were considered to be 'in target' with no further treatment needed. The EAG stated that it is unlikely that UK clinical practice would follow a similar protocol.

The trial reported significant reductions (that is, improvement) in scores between the first and last visits in the PKG arm, in total UPDRS (-11.6; 95% CI 5.8 to 17.5), UPDRS IV (complications of therapy; -2.3; 95% CI 1.1 to 3.6), UPDRS III (motor examination; -7.9; 95% CI 4.6 to 11.2) and PDQ-39 (-6.1; 95% CI 0.3 to 11.8). There were no significant reductions in scores in the standard care arm. The EAG commented that the improvement was likely to be because the PKG use resulted in changes in care that reduced time with bradykinesia, dyskinesia and tremor (although none of these reductions reached statistical significance). The EAG commented that there was a

general trend favouring the PKG in outcomes, except for inactive time, which was higher in people using the PKG than in people not using one. The EAG highlighted that this outcome was not reported in the original trial publication. Differences in the change from baseline between the PKG and standard care arm are shown in figure 1. UPDRS I (mentation, behaviour and mood) and UPDRS II (activities of daily living) scores did not decrease in the PKG arm. Changes were not significantly different, and for UPDRS I virtually no difference was seen.

Figure 1. Results from the Woodrow trial: differences between PKG and standard care



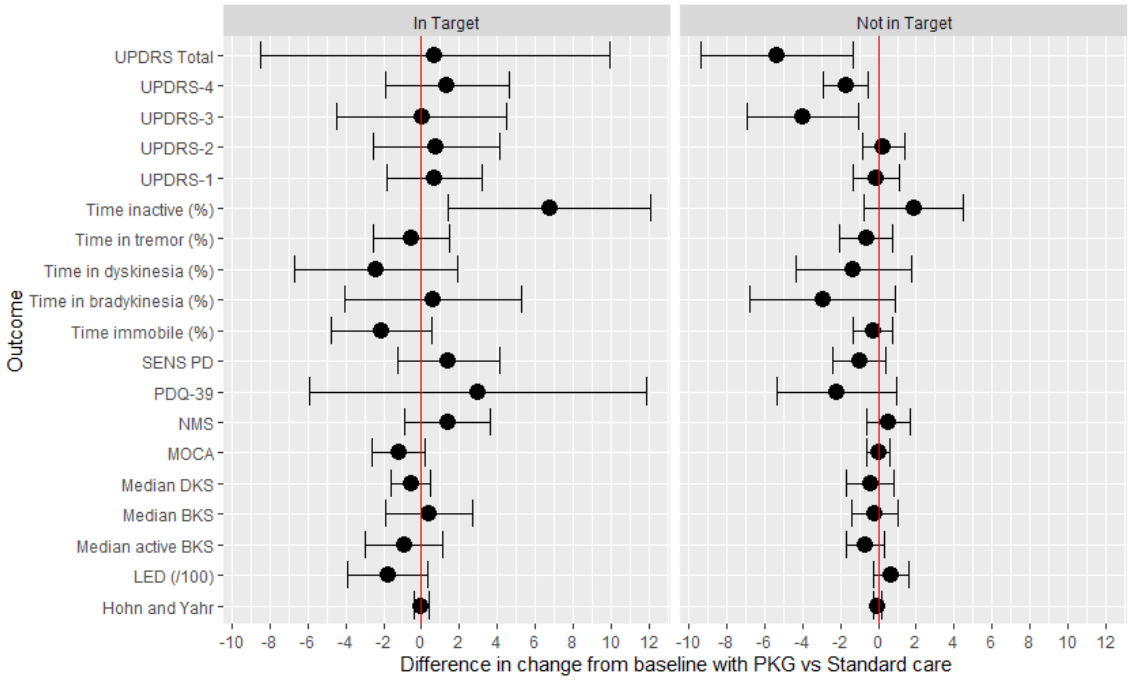
Abbreviations: BKS, Bradykinesia Score; DKS, Dyskinesia Score; LED, levodopa equivalent dose; MOCA, Montreal Cognitive Assessment; NMS, Non-Motor Symptoms; PDQ-39, Parkinson's Disease Quality of Life 39 Questions; SENS PD, Severity of Predominantly Non-Dopaminergic Symptoms in Parkinson's Disease; UPDRS, Unified Parkinson's Disease Rating Scale.

The EAG adjusted individual patient data for confounding factors (age, sex, Parkinson's disease duration, UPDRS III [motor examination] at baseline and number of clinic visits during follow up) and reanalysed the data. The EAG reported broadly the same outcomes as in the trial publication. Full detail on

the EAG's data analysis is in the appendix table 67 on page 225, and is summarised in figure 2 on page 54, of the diagnostic assessment report.

The EAG also did subgroup analyses for people with symptoms that were categorised using PKG results as 'in target' (defined as a PKG bradykinesia score under 26 and a dyskinesia score under 7) and 'not in target' at baseline. Figure 2 shows the changes in outcome from baseline when using the PKG versus standard care in these 2 subgroups. The EAG suggested that by reducing bradykinesia through changes in care, PKG use improved UPDRS score in people whose disease was not being adequately controlled. There was also a non-significant reduction in levodopa dose and dyskinesia in people whose condition was already well controlled. The EAG suggested that using the PKG may be useful in improving UPDRS score, by reducing bradykinesia, in people whose disease is not being adequately controlled. And that it may allow for levodopa dose reduction, and consequent reduction in dyskinesia, in people whose condition is already well controlled. Figures 2 and 3 on pages 54 and 55 in the diagnostic assessment report summarise the Woodrow trial results. The EAG concluded that there is reasonable evidence that using the PKG leads to genuine clinical improvements for some people, but that multiple clinic visits and PKG assessments may be needed before a controlled state is reached.

Figure 2 Impact of the PKG in the Woodrow trial, by 'in target' status at baseline



Abbreviations: BKS, Bradykinesia Score; DKS, Dyskinesia Score; LED, levodopa equivalent dose; MOCA, Montreal Cognitive Assessment; NMS, Non-Motor Symptoms; PDQ-39, Parkinson's Disease Quality of Life 39 Questions; SENS PD, Severity of Predominantly Non-Dopaminergic Symptoms in Parkinson's Disease; UPDRS, Unified Parkinson's Disease Rating Scale.

The EAG also assessed the impact of PKG use on change in medication, referral for device-assisted therapy (exact therapies were not reported) and 'in target' status (all dichotomous outcomes). The results suggested that people using the PKG were substantially more likely to have symptoms that were 'in target' at follow up (odds ratio [OR] 3.43; 95%CI 1.67 to 7.03) and to be referred for device-assisted technologies (OR 4.01; 95% CI 1.82 to 8.85). But the EAG said there was no clear evidence that people using the PKG were more likely to have a change in medication (OR 1.18; 95% CI 0.52 to 2.68). See table 9 in the diagnostics assessment report for further details.

The EAG highlighted that the trial included multiple uses of the PKG device over a short period of time (once every 5 weeks) so the clinical benefits of using the PKG less frequently might be different.

Additional clinical outcome studies

The 5 remaining clinical outcome studies (n=28 to 104) were judged to be at low risk of bias. They all had a non-comparative, cohort study design in which the impact of the PKG was measured over time. The EAG highlighted that it is unclear how much of the benefit can be attributed to the PKG specifically. Find more details in tables 6, 7 and 8 on pages 48, 52 and 57 of the diagnostics assessment report.

Diagnostic accuracy

Seven papers reporting diagnostic accuracy data for the PKG were identified. Three studies that reported accuracy in detecting bradykinesia, dyskinesia and tremor (Braybrook et al. 2016, Horne et al. 2016, Horne et al. 2015) reported sensitivities above 90% and specificities ranging from 83% to 92.9%. These studies used reference standards including UPDRS III (motor examination), AIMS and clinical examination. The EAG commented that the PKG was less accurate for measuring sleep disturbance (80% sensitivity, 86% specificity; McGregor et al. 2018) when using the Parkinson's Disease Sleep Scale 2 (PDSS 2) for people with Parkinson's disease, and polysomnography (a sleep study) in people without Parkinson's disease.

In 2 studies the PKG accurately identified levodopa responses (AUC 92%) and the need for device-assisted therapy (AUC 93%; Khodakarami et al. 2019a, Khodakarami et al. 2019b). These studies used clinical assessment including MDS-UPDRS (Khodakarami et al. 2019a) and levodopa response (measured by UPDRS III [motor examination], Khodakarami et al. 2019b) as reference standards. Another smaller study (n=26) showed slightly poorer accuracy for treatment classification (AUC 83.1%; Watts et al. 2021). The EAG concluded that, overall, the PKG showed high diagnostic accuracy, but its quality assessment identified substantial concerns. Full details on the diagnostic accuracy of the PKG are in table 4 on page 45 of the diagnostics assessment report. A summary of the Quality Assessment of Diagnostic

Accuracy Studies (QUADAS)-2 assessment is in table 3 on page 44 of the diagnostics assessment report.

Association outcomes

The association between PKG outputs and symptoms or other clinical outcomes was reported in 11 full publications and 10 conference abstracts. Evidence on the level of agreement between PKG outputs and clinical assessment varied (between 54.5% and 90%; Dominey et al. 2020, Farzanehfar et al. 2018, Krause et al. 2021). The EAG stated that PKG bradykinesia, dyskinesia and tremor scores were generally moderately correlated with UPDRS scores, whereas sleep and impulse control behaviours were less so. There is further detail on the PKG association studies on pages 46 to 48 in the diagnostic assessment report.

Impact of monitoring

The impact of monitoring on decisions about care was reported in 8 full publications, and 17 conference abstracts that were not discussed by the EAG. Five studies had a low overall risk of bias (Farzanehfar et al. 2018, Joshi et al. 2019, Krause et al. 2021, Nahab et al. 2019, Sundgren et al. 2021). One study did not clearly define the inclusion criteria, had high attrition, and did not state blinding methods clearly (Santiago et al. 2019). The 2 remaining studies were judged to have a high risk of bias (Dominey et al. 2020, Evans et al. 2020); these were the only studies done in the UK. Table 6 on page 48 of the diagnostics assessment report summarises the quality assessment of studies reporting on the intermediate impact of monitoring (changes in clinical management).

The proportion of people who had a change in clinical management as a result of the PKG varied considerably (between 31.8% and 79%; Dominey et al. 2020, Farzanehfar et al. 2018, Joshi et al. 2019, Nahab et al. 2019; Santiago et al. 2019, Sundgren et al. 2021). Adding a new therapy or increasing treatment dose were the most common changes in clinical management. Evans et al. (2020) assessed PKG use in virtual appointments

in the UK; 79% of virtual appointments were deemed successful (the clinician felt the outcome of the consultation was likely to have been the same as that in a face-to-face clinic).

User opinions

Patient or carer opinions were reported in 4 full studies (Dominey et al. 2020, Evans et al. 2020, Joshi et al. 2019, Nahab et al. 2019). Two studies found 75.3% (Joshi et al. 2019) and 100% (Nahab et al. 2019) of PKG users agreed that the PKG was most useful as a medication reminder. Patients and carers mostly agreed that the PKG was generally easy to use and found it helpful for discussions with their clinician.

The EAG commented that it was less clear if the PKG provided useful information on symptoms that the patients or carers could not provide themselves. Dominey et al. (2020) reported that this was the case for 59% of people and Nahab et al. (2019) reported 89%.

People were generally willing to use the PKG device again (Nahab et al. 2019). They also deemed virtual appointments successful when the PKG was used for remote management (Evans et al. 2020). Reasons for unsuccessful consultations included being in a complex phase of disease, problems with the PKG, needing a blood pressure reading and speech problems. Further details are summarised in table 11 on page 60 of the diagnostics assessment report.

Clinician opinions on the value of the PKG were reported in 4 full studies (Joshi et al. 2019, Nahab et al. 2019, Santiago et al. 2019, Sundgren et al. 2021). Between 4% (Nahab et al. 2019) and 41% (Santiago et al. 2019) of clinicians agreed that the PKG improved their assessment of symptoms or need for changes in therapy. Clinicians were largely in agreement that the PKG improved dialogue with patients (Joshi et al. 2019, Nahab et al. 2019, Sundgren et al. 2021).

Nahab et al. (2019) also reported that 89% of clinicians felt that the PKG improved their ability to assess the impact of therapy, whereas Joshi et al. (2019) reported that only 38% felt this way. The EAG highlighted that these clinical opinions were somewhat contradictory to the evidence on changes in treatment, which showed that many people did have a change in treatment because of PKG use.

2.4 STAT-ON

Clinical outcomes

No studies reporting on clinical outcomes were identified for STAT-ON.

The EAG identified an ongoing RCT (NCT04176302) in Spain. This trial is assessing the clinical efficacy of the STAT-ON compared with traditional clinical practice (primary objective), and whether it is not inferior to the 'on-off' diary recorded by the participants at home (exploratory objective). Secondary outcomes include changes in clinical management, number of clinical visits, user satisfaction and system usability.

Diagnostic accuracy

The EAG identified 8 accuracy publications, but noted that none of the papers explicitly indicated the index test as being the STAT-ON (although all were listed on the manufacturer's website). In some studies, the index test appeared to be part of the data analysis itself, and so it is unclear whether these index tests were part of the current device in commercial use. Most of the papers reported data from small samples, with overlapping authorship and a common source of trial data (most were from the REMPARK study). Overall, the EAG concluded that the risk of bias was unclear because of limited reporting.

Of the 8 studies, 4 papers assessed whether STAT-ON could assess 'on-off' times and reported sensitivities between 90.3% and 97%, specificities between 88% and 94% (Bayés et al. 2018, Pérez-López et al. 2016,

Rodríguez-Molinero et al. 2018, Rodríguez-Molinero et al. 2015). One study also reported positive and negative predictive values of 92% and 94% respectively (Rodríguez-Molinero et al. 2018). For identifying freezing of gait, the EAG commented that STAT-ON also had a high sensitivity (88.1% and 91.7%), with a lower specificity (80.1% and 87.4%; Rodríguez-Martín et al. 2017, Samà et al. 2018, respectively). One paper reported that the STAT-ON had 92.5% sensitivity and 89.1% specificity for detecting bradykinesia (Samà et al. 2017), and another found that STAT-ON had high accuracy for detecting trunk dyskinesia, but not for detecting dyskinesia in general (Pérez-López et al., 2016).

There is a summary of the QUADAS-2 assessment and details of the accuracy studies in tables 13 and 14 on pages 65 and 66 of the diagnostics assessment report.

Association outcomes

The EAG identified 3 studies that reported association outcomes. A small study (n=13) reported a significant correlation between STAT-ON and clinical assessment (0.70; 95% CI: 0.33 to 0.88) using the UPDRS. This was higher for trunk and legs scale sub-items (0.91, 95% CI: 0.76 to 0.97; Rodríguez-Molinero et al. 2019). A larger study (n=75) showed a moderate correlation between UPDRS III (motor examination) and STAT-ON outputs (rho -0.56), with good correlation between STAT-ON outputs and the gait item (14 items in total) in the UPDRS III (rho -0.73; Rodríguez-Molinero et al. 2017). Full details on the included association studies are in table 15 on page 67 of the diagnostic assessment report.

Impact of monitoring

No studies reporting the impact of results on decisions about care were identified for STAT-ON.

User opinions

In Bayés et al. (2018), 76.0% of people (n=33) reported satisfaction with STAT-ON. Rodriguez-Martin et al. (2021; only available as a conference abstract) also found that 76% of carers found it easy to use and 80.0% found it to be a 'good to very good' solution.

The STAT-ON was considered a 'useful tool' by most neurologists, and 81.5% considered it a useful tool for identifying people with advanced Parkinson's disease symptoms (Santos Garcia et al. 2020). A cohort study (n=39) reported in a conference abstract found that satisfaction with STAT-ON among users was high, and the system was found to be easy to use (Caballol at al. 2020). Full details on STAT-ON user opinion studies are summarised in table 16 on page 68 of the diagnostics assessment report.

3 Cost-effectiveness evidence

The EAG did a systematic review to identify any published economic evaluations of devices for remote continuous monitoring of Parkinson's disease. Find the full systematic review results on pages 80 to 85 of the diagnostics assessment report.

Systematic review of cost-effectiveness evidence

The EAG found no studies that met its inclusion criteria for any of the 5 devices, but retrospectively included 2 non-comparative studies that assessed cost savings and quality-adjusted life year (QALY) gains associated with the PKG (Lynch et al. 2018 and Rao et al. 2019). See pages 82 and 83 in the diagnostic assessment report for discussion of these studies. A company-sponsored cost-effectiveness study assessing the PKG was published after the EAG's cut-off date for its searches (Chaudhuri et al. 2022).

Chaudhuri et al. (2022)

This study used a Markov model to estimate the cost utility of the PKG and clinical assessment in the management of Parkinson's disease compared with

standard care. The population characteristics and clinical efficacy of the PKG compared with standard care were based on data from Woodrow et al. (2020). The assessment was done in the context of the NHS, with a discount rate of 3.5%. The model comprised 3 health states: (1) uncontrolled; (2) controlled; and (3) death, with the ability to transition in both directions between the uncontrolled and controlled health states. The uncontrolled and controlled states modelled different levels of symptoms, based on MDS-UPDRS data from Woodrow et al. This data was used to estimate health-related quality of life and costs for each state. There is a full description of the model in the diagnostics assessment report from page 83. People entered the model in an uncontrolled state and could transition between uncontrolled and controlled health states dependent on the improvement in MDS-UPDRS II (motor experiences of daily living) and III (motor examination) scores with either the PKG plus clinical assessment, or with clinical assessment alone (considered to represent current practice). The treatment effectiveness was assumed to be maintained for 5 years (based on a systematic literature review on the impact of levodopa–carbidopa intestinal gel in people with advanced Parkinson's disease), after which alternative rates of progression were modelled. A lifetime horizon of 22 years was used.

To estimate health-related quality of life, EQ-5D values were derived from MDS-UPDRS domain scores for II (motor experiences of daily living) and III (motor examination; from Woodrow et al. 2020) using a published mapping algorithm. The EAG raised concerns about this approach because it was based on a small, non-UK population and health states were converted into EQ-5D indices using weights from a pooled European population valued by a visual analogue technique, which does not align with the NICE reference case. The EAG also highlighted that the resulting estimates of health state utility values differed substantially from other alternative sources and approaches considered by the authors.

Chaudhuri et al. (2022) reported that the PKG, as an adjunct to clinical assessment, dominated standard care (£17,362 lower costs and 0.267 higher

QALYs per patient). Sensitivity and scenario analyses showed that the cost-effectiveness results were robust. But the EAG commented that most of these analyses had key structural assumptions affecting the long-term efficacy, health-related quality of life and costs over a lifetime horizon in the model (described below).

The EAG stated that the study appeared to be well conducted and was done from the perspective of the UK NHS, which makes it directly relevant to the decision problem. However, it highlighted several concerns about the work. Find a full description from page 86 of the diagnostics assessment report. In overview:

- Clinical effectiveness of the PKG was based on the MDS-UPDRS domain II (motor experiences of daily living) and III (motor examination) scores, so did not reflect the impact of the PKG on non-motor experiences of daily living (MDS-UPDRS I) or on the most severe motor complications (MDS-UPDRS IV).
- Using a lifetime horizon when there was no evidence that the benefits of the PKG equated to long-term changes in treatment, particularly as there are no disease-modifying treatments available, was concerning. The EAG stated that several strong assumptions were made to support the 6 to 12 months of benefits observed in Woodrow et al. (2020) being sustained over a longer time. These are described in the diagnostics assessment report on page 87.
- The EAG considered it unlikely that UK clinical practice would follow a similar protocol to that in Woodrow et al. (as used in the Chaudhuri et al. model) in which up to 5 consultation visits (5 weeks apart) occurred, with a PKG worn before each visit.
- Costs in the model for the controlled and uncontrolled states were based on converting the MDS-UPDRS scores to Hoehn and Yahr stages. The costs associated with the Hoehn and Yahr stages and the proportion of people in each stage by intervention strategy substantially affected the total costs. The EAG was unable to identify or validate the costs reported in

Chaudhuri et al. and stated that it was unclear how the MDS-UPDRS scores were converted onto the Hoehn and Yahr scale. Chaudhuri et al. (2022) stated that the MDS-UPDRS scores in the model were applied to derive average annual costs by Hoehn and Yahr stage, but no details were provided. The EAG stated that the cost savings for the PKG estimated by the model were driven by a lower likelihood, on average, that a person will end up in one of the more severe Hoehn and Yahr stages (3 to 5; 36.9%) compared with standard care (48%) over a lifetime horizon.

Economic analysis

The EAG developed a de novo economic model to assess whether devices for remote continuous monitoring of Parkinson's disease represent a cost-effective use of NHS resources. The model was limited to establishing the cost effectiveness of the PKG and the Kinesia 360 compared with standard care for people in the maintenance stage of Parkinson's disease (that is, symptoms are controlled with or without medication). This was because there was no comparative evidence on clinical effectiveness for STAT-ON, the KinesiaU, PDMonitor, or for any device in advanced Parkinson's disease.

The EAG modelled 2 monitoring strategies for the PKG and the Kinesia 360 in the base case:

- **one-time use**: remote monitoring (PKG or Kinesia 360) used once at the start of the model as an aid to clinical assessment
- **routine use**: remote monitoring (PKG or Kinesia 360) every 6 months.

The EAG explained that the one-time use attempts to approximate using the remote monitoring devices as in Woodrow et al. (2020) and Isaacson et al. (2019). A third strategy of recurrent use was considered in a scenario analysis, to attempt to replicate using the PKG in the NHS as reported in Dominey et al. (2020). In this analysis, the remote monitoring devices replaced the clinical appointments (at 6 months, 18 months and so on for

most people). Table 1 gives an overview of the standard care and device-based strategies modelled.

Table 1 Device use and clinical appointment (remote or face-to-face) schedules for each monitoring strategy

Base-case schedules	Baseline		6-months		12-months		18-months		Time horizon	
	RMD	SC	RMD	SC	RMD	SC	RMD	SC	RMD	SC
SC	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
PKG (one-time use)	Yes	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Kinesia 360 (one-time use)	Yes	Yes	No	Yes	No	Yes	No	Yes	No	Yes
PKG (routine use)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kinesia 360 (routine use)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
PKG (recurrent use scenario)	Yes	Yes	Yes	No*	No	Yes	Yes	No *	No	Yes
Kinesia 360 (recurrent use scenario)	Yes	Yes	Yes	No*	No	Yes	Yes	No *	No	Yes

*Intermediary consultations withdrawn, except for people with exceptional clinical need (assumed to be 21% in the base-case analysis).
Abbreviations: PKG, Personal KinetiGraph; RMD, remote monitoring device follow up; SC, standard care follow up.

The model assumed that remote monitoring provides some symptomatic relief by improving clinical assessment and therapeutic decisions (changes in MDS-UPDRS score from Woodrow et al. 2020 and Isaacson et al. 2019).

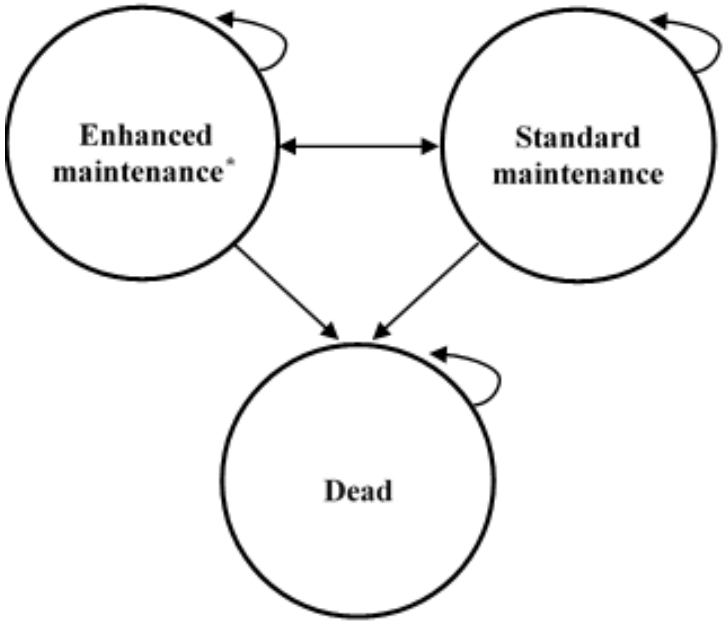
Model structure

The EAG's Markov model had 3 health states: enhanced maintenance, standard maintenance, and an absorbing death state. Figure 3 shows the possible transitions between these health states. For any remote monitoring strategy (that is, using the technologies being assessed), the cohort entered the model in the enhanced maintenance health state. People moved to the standard maintenance health state every cycle based on expected waning of symptom relief caused by device-guided changes to treatment if the device was assumed to be used only once ('one-time use'). When modelling routine

or recurrent use of the devices, people were assumed to remain in the enhanced maintenance state. Although the model allowed movement from the standard maintenance state to the enhanced maintenance state, this was not used in the EAG's analyses.

For standard care (the comparator), people entered the model in the standard maintenance health state and could not move to the enhanced maintenance state.

Figure 3 Markov model structure



*Only applicable to people having remote monitoring.

The 2 maintenance states were associated with specific MDS-UPDRS (I to IV) domain scores, which reflected the level of symptom control associated with remote monitoring (enhanced maintenance health state) and standard care (standard maintenance health state). Each of the maintenance states was associated with different health-related quality of life. This was based on changes in UPDRS score data from using the remote monitoring devices (from comparative studies) for the enhanced maintenance state compared with not using the devices (standard maintenance state).

The cycle length was 6 months and the EAG assumed a 5-year time horizon was most appropriate. This EAG considered that a lifetime model would need to account for transitions to more advanced stages of disease, and model advanced treatments such as deep brain stimulation, apomorphine injections and levodopa–carbidopa intestinal gel. Although the EAG considered it plausible the technologies could affect the transition to more advanced therapies, evidence was not available for modelling this. It further commented that the lack of disease-modifying or curative drugs suggests that any effects on costs and benefits would be transitory, and most likely only affect when the advanced treatments are started. The 5-year time horizon assumed that remote monitoring devices will be used for a maximum of 5 years (reflecting the approximate duration of the maintenance phase) with no lasting differences to costs and benefits after this time.

Population

The maintenance phase of Parkinson's disease was modelled, in which symptomatic motor and non-motor features of the disease are routinely managed. Patient characteristics were based on Woodrow et al. (2020), with a starting age of 67.8 years.

Comparator

The comparator was standard care in the UK NHS (with no use of the remote monitoring technologies).

Model inputs

The full list of model parameters is in table 31 on pages 119 to 123 in the diagnostics assessment report.

Consultation appointments

People having standard care were assumed to have consultation review appointments every 6 months, with 55% of consultations conducted face-to-face and 45% done remotely (2019/20 NHS reference costing schedule).

Use of review appointments alongside use of remote monitoring devices varied. The EAG also assumed that the repeated use of remote monitoring would increase the use of remote consultations (rather than face-to-face) to 79% (from 45%) (Evans et al. 2020). See table 1 for the frequency of review appointments for the different strategies of remote device use.

Efficacy

The impact of the remote monitoring devices was based on improvements in symptoms assessed using UPDRS scores, compared with not using the technologies. For the PKG this was based on Woodrow et al. (2020), with MDS-UPDRS scores converted to UPDRS scores. The EAG adjusted for potential confounding factors (age, sex, duration of Parkinson's disease, UPDRS III (motor examination) score at baseline and number of clinical visits during follow up). The EAG modelled data for all UPDRS domains from Woodrow et al. (scores worsen [increase] for domain II [activities of daily living]; unrestricted analysis, see table 2) and also, in a separate analysis, assuming that the PKG has no detrimental impact on UPDRS score (score for domains I and II set to 0; restricted analysis).

The impact of using the Kinesia 360 was based on Isaacson et al. (2019), which did not measure UPDRS domains I (mentation, behaviour and mood) or IV (complications of therapy).

Estimated changes in UPDRS (compared with standard care) are shown in table 2. Full details on the efficacy estimates are in table 26 on page 109 of the diagnostics assessment report.

Table 2 Efficacy estimates used in modelling

Intervention	Source	UPDRS I (mentation, behaviour and mood)	UPDRS II (activities of daily living)	UPDRS III (motor examination)	UPDRS IV (complications of therapy)
PKG Unrestricted and adjusted	Woodrow et al. (2020)	0	0.52	-2.84	-0.73
PKG Restricted and adjusted	Woodrow et al. (2020)	0	0	-2.84	-0.73
Kinesia 360	Isaacson et al. (2019)	0	-2.6	-4.3	0

Waning of impact of remote monitoring devices on symptoms

Assumed waning of the impact on symptoms resulting from device-guided changes to care (transition between the enhanced and standard health states) was based on assumptions, because there was no evidence on the efficacy of devices beyond the immediate term (up to 25 weeks after initial assessment). For the one-time strategies, every 6 months 50% of people were assumed to move from the enhanced maintenance to the standard maintenance health state. For routine and recurrent use of the devices, no waning in benefit was assumed (everyone remained in the enhanced maintenance health state).

Detail on the approach to modelling disease progression, mortality and adverse events is in the diagnostics assessment report from page 111.

Costs

The full list of costs used in the model is in table 31 on page 119 of the diagnostics assessment report. Differences in costs between standard care and alternative remote monitoring strategies were from costs associated with using the devices, changes in medication use, implementation costs and differences in clinical appointment use.

Monitoring devices

The 5 monitoring devices had 3 types of payment mechanism: (i) pay per use; (ii) subscription model; or (iii) outright purchase of the device. Modelled costs over 5 years are shown in table 3. There are more details on page 116 of the diagnostics assessment report.

Table 3 Remote continuous monitoring device costs

	Cost (exc. VAT)	Unit	Modelled cost per year	Modelled cost per 5 years (base-case time horizon)
Kinesia 360	£224	Monthly device subscription	£2,688**	£13,440**
KinesiaU	£64	Monthly subscription per patient	£768**	£3,840**
PDMonitor	£12,000	Outright device purchase	£12,000**	£12,000**
PKG	£225	Per use per patient	£450*	£2,250*
STAT-ON	£1,600	Annual device subscription	£1,920**	£9,600**

* Excludes initial assessments, **assumes one patient per subscription or device.

Consultation appointments

In line with Woodrow et al. (2020), it was assumed that several initial face-to-face consultations would be needed. For all remote monitoring strategies, people were assumed to have 2.57 initial face-to-face consultations, and people receiving standard care were assumed to have 2.17 visits. Face-to-face and remote consultations were assumed to cost £81.41 and £56.41, respectively (NHS reference costs 2019-2020). The base-case analysis assumed consultation costs sufficiently covered potential broader service factor costs such as device delivery, administration and relevant support.

Further details on the consultation costs for each of the monitoring strategies are in table 31 on pages 121 to 122 of the diagnostics assessment report.

Details of costs related to implementation and medication costs are on pages 117 and 118 of the diagnostics assessment report.

Health-related quality of life

Patient health-related quality of life

The EAG estimated QALYs from UPDRS domain scores using an algorithm from Chandler et al. (2020). This was identified from a review done by the EAG of decision models evaluating interventions used in Parkinson's disease (there is more detail on page 89 of the diagnostics assessment report). Health-related quality of life was only dependant on the UPDRS score and sex.

There are more details on page 114 of the diagnostics assessment report.

Carer quality of life

Because of a lack of data, the EAG did not include any impact of device use or change in symptoms in people with Parkinson's disease on carer quality of life. It highlighted that this may be an uncaptured benefit.

Base-case results

Cost effectiveness of remote monitoring using the Kinesia 360

The base-case probabilistic analysis results are shown in table 4. The deterministic results were similar (see tables 38 and 39 on page 128 of the diagnostics assessment report).

Table 4 Kinesia 360 probabilistic base-case cost-effectiveness results

Strategy	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER
One-time use					
Standard care	£21,886	2.760			
Kinesia 360	£22,651	2.780	£765	0.01977	£38,722
Routine remote monitoring					
Standard care	£21,951	2.756			
Kinesia 360	£34,061	2.936	£12,110	0.17973	£67,376

Abbreviations: ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year.

Cost of the device had the largest impact on the incremental costs compared with standard care. This was £672 out the total incremental cost for one-time use (about 88% of the cost difference) and £11,685 out of the total incremental costs for routine monitoring (about 96%). See table 40 in the diagnostics assessment report for more detail.

Kinesia 360 scenario analyses

In the scenario analyses for the Kinesia 360, the ICERs varied only slightly compared with the base-case ICERs. Recurrent monitoring produced very similar results to those of routine monitoring.

There are more details in table 41 on page 131 of the diagnostics assessment report.

Cost effectiveness of remote monitoring strategies using the PKG

The base-case probabilistic analysis results are shown in table 5. The cost-effectiveness estimates varied considerably for the restricted and unrestricted analyses (described in the efficacy section). The deterministic results were similar and are in table 34 and 35 on page 126 of the diagnostics assessment report.

Table 5 PKG probabilistic base-case cost-effectiveness results

Strategy	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER
One-time use restricted analysis					
Standard care	£21,886	2.760			
PKG	£22,225	2.765	£339	0.00504	£67,260
One-time use unrestricted analysis					
Standard care	£21,953	2.755			
PKG	£22,291	2.757	£338	0.00171	£197,475
Routine remote monitoring restricted analysis					
Standard care	£21,951	2.756			
PKG	£24,578	2.801	£2,627	0.04553	£57,702
Routine remote monitoring- unrestricted analysis					
Standard care	£21,875	2.756			
PKG	£24,527	2.771	£2,652	0.01509	£175,711

Abbreviations: ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year.

Cost of the device had the largest impact on the incremental costs compared with standard care. This was £225 out the total incremental cost for one-time use (about 67% of the cost difference) and £2,181 out of the total incremental costs for routine monitoring (about 83%). See table 36 in the diagnostics assessment report for more detail.

PKG scenario analyses

All scenario analyses for the PKG in the unrestricted analysis gave ICERs over £68,000 per QALY gained. In the restricted analysis, the ICERs for all but 2 of the scenarios were over £44,000 per QALY gained.

In scenario 1, PKG use was modelled as recurrent remote monitoring (see table 1). This reduced the ICER to £32,417 per QALY gained. In scenario 5, which modelled annual (as opposed to 6-monthly) routine follow up with the PKG (keeping the same assumed size of benefit as in the base case), the

PKG ICER decreased to £36,973 per QALY gained. The EAG commented that this was driven by reduced device costs compared with the base case.

There are more details in table 37 on pages 128 to 129 of the diagnostics assessment report.

Comparing the cost effectiveness of the Kinesia 360 and PKG

The EAG compared the cost effectiveness of the Kinesia 360, the PKG and standard care in a fully incremental analysis. For both one-time use and routine remote monitoring, the PKG was extendedly dominated by the Kinesia 360 and standard care. Details of the analysis are shown in tables 42 and 43 on page 133 of the diagnostics assessment report.

The EAG highlighted that caution must be taken when comparing the cost effectiveness of the PKG against the Kinesia 360 given the fundamental differences in the underlying evidence base that informed the expected improvements in patient outcomes. No study directly compared the performance of the devices. The EAG highlighted that the population in Isaacson et al. (2019) had more severe disease (people experiencing clinically significant motor symptoms that were not controlled by current therapy) than those in Woodrow et al. (2020).

Cost-effectiveness comparison of remote monitoring using the Kinesia 360, KinesiaU, PDMonitor, PKG and STAT-ON

The EAG produced an exploratory cost comparison analysis of all the monitoring devices. Because of a lack of data, the devices were modelled using the same parameters as for the PKG, changing only the cost of the device.

Costs were markedly different between the remote monitoring devices (see table 3). Device-related costs were the largest part of the incremental costs (across all devices and remote monitoring strategies). The EAG highlighted considerable uncertainty about costs, given the uncertainty about how the devices might be used in practice (for example, devices could be shared

between people, which would allow for an average device cost per person, but incur different costs such as administrative costs associated with the shared-use model).

Assuming equal efficacy for all the devices, analysis of the one-time use strategy showed that the KinesiaU dominated the other devices, and had an ICER of £53,331 per QALY gained compared with standard care for the restricted analyses, and £170,975 for the unrestricted analyses. For the routine remote monitoring strategy, the PKG dominated the other devices, and had an ICER of £57,877 per QALY compared with standard care for the restricted analyses, and £172,602 for the unrestricted analyses.

There are more details in table 45 and 46 on pages 134 to 135 of the diagnostics assessment report.

Sensitivity analyses

There was limited evidence to evaluate the effect of remote monitoring on patient outcomes beyond the immediate term, so the EAG assumed waning rates. To explore this, the EAG did a sensitivity analysis to assess the impact of alternative waning rates (0% to 90%). The results showed that the ICER was highly sensitive to the assumed waning rate for the Kinesia 360 and PKG (see table 6 for results for up to 60% waning). In Chaudhuri et al. (2022) no treatment waning for 5 years (the time frame of the EAG's model) was assumed.

Table 6 Sensitivity analysis of alternative efficacy waning rates

Pairwise ICERs (vs standard care)	0% efficacy decay rate	10% efficacy decay rate	20% efficacy decay rate	30% efficacy decay rate	40% efficacy decay rate	50% efficacy decay rate	60% efficacy decay rate
One-time use							
PKG - restricted analysis	£16,614	£20,816	£27,174	£36,352	£49,310	£67,856*	£95,853
PKG - unrestricted analysis	£49,548	£62,078	£81,039	£108,410	£147,054	£202,363*	£285,854
Kinesia 360	£6,570	£9,215	£13,218	£18,996	£27,153	£38,828*	£56,452
Routine use							
PKG - restricted analysis	£57,877*	£90,786	£140,284	£211,507	£311,919	£455,559	£672,355
PKG unrestricted analysis	£172,602*	£270,743	£418,357	£630,760	£930,214	£1,358,581	£2,005,114
Kinesia 360	£67,203*	£110,636	£176,285	£270,991	£404,667	£595,972	£884,745

*Base-case value. Grey shading indicates waning rates used in the base case models.

A further sensitivity analysis showed that the results were broadly insensitive to reductions in the cost of consultation appointments. The full sensitivity analyses results are on pages 135 to 138 of the diagnostics assessment report.

Differences between Chaudhuri et al. (2022) and the EAG model

Despite evaluating use of the PKG in the same context and using the same source of data for device impact (Woodrow et al. 2020), the cost-effectiveness results from the Chaudhuri et al. (2022) study differ drastically from those in the EAG's analyses. The EAG identified several possible reasons for the difference:

- Chaudhuri et al. modelling included cost savings from improvements in symptoms which were assumed to occur because of PKG use. This was

not included in the EAG's model. The EAG commented that its approach could be considered conservative, but it had concerns about the Chaudhuri et al. approach. The EAG also stated that it was not clear if reductions in MDS-UPDRS scores associated with remote monitoring would translate into meaningful reductions in NHS costs at the management stage of Parkinson's disease.

- QALYs were accumulated over the lifetime (up to 22 years) in Chaudhuri et al. (2022), compared with 5 years in the EAG's model.
- The frequency of use of the PKG assumed in the Chaudhuri model is higher than the EAG considered is likely in UK clinical practice. The EAG modelled one-time and routine bi-annual applications (over 5 years). Chaudhuri et al. (2022) modelled people with controlled symptoms having monitoring with PKGs twice, and people with uncontrolled symptoms having monitoring with PKG 3 to 4 times per year.
- Utilities associated with changes on the UPDRS scale, and the sources and methods used to calculate them, are markedly different between the analyses.
- Different UPDRS domains were considered; Chaudhuri et al. (2022) only considered changes in MDS-UPDRS domains II (motor experiences of daily living) and III (motor examination) whereas the EAG considered data from all UPDRS domains from Woodrow et al.
- Differences in model structure and the estimation methods used to derive efficacy estimates may also have contributed to differential findings.

For a full discussion, see the diagnostics assessment report from page 140.

4 Summary

Clinical effectiveness

Varying levels of data were identified for the different technologies. There were no studies that directly compared any of the devices.

The EAG considered that, because of the very different natures of the technologies assessed, such as the specific symptoms they measure, the position of the sensors on the body and the characteristics of people with Parkinson's disease that the devices have been assessed in, it should not be assumed that any clinical benefits observed for 1 technology would also be found with the other technologies.

The outcomes reported in studies also varied across the devices. Only 2 devices had studies that reported clinical outcomes comparing device use with standard clinical practice; 1 on the PKG (Woodrow et al. 2020) and 2 on the Kinesia 360 (Isaacson et al. 2021 and Peacock et al. 2021). The EAG commented that much of the evidence was either diagnostic accuracy or association studies. These were generally proof-of-concept studies to demonstrate that the devices could provide clinically viable measurements.

Kinesia 360

Evidence for the Kinesia 360 was largely from 2 small RCTs. The EAG commented that these suggested favourable results when using the Kinesia 360, with reductions in UPDRS scores, and improvements in quality of life (which were comparable to those seen for the PKG). One diagnostic accuracy study found that the Kinesia 360 had moderate to good accuracy for diagnosing dyskinesia, bradykinesia and tremor. The EAG concluded that there was a limited evidence base. It considered that the Kinesia 360 is a promising technology, but there is too little evidence at present to be confident about its clinical value.

KinesiaU

Evidence on the KinesiaU was from 1 small study (n=16). The EAG considered this to be too little evidence from which to draw any conclusions about the clinical value of the KinesiaU. It also stated that patient opinion of the KinesiaU system was not particularly favourable.

PDMonitor

The EAG considered that the small amount of evidence available meant that no conclusions about the clinical value of the PDMonitor could be drawn.

PKG

Overall, the EAG concluded that there is a good body of evidence to support the use of the PKG in practice. It commented that the PKG appears to be able to reliably measure bradykinesia, dyskinesia and tremor, and that the device provides information that leads to changes in clinical management for at least some people. There is reasonable evidence that using the PKG led to genuine clinical improvements when compared with standard management, in terms of reductions in UPDRS scores. This benefit seemed to depend on whether people's symptoms were 'in target' (condition was under control with current treatment) before PKG use. People with symptoms 'not in target' saw improved UPDRS scores, but those with symptoms 'in target' did not. People with symptoms 'in target' may, however, benefit from reduced levodopa dosing and consequent reduction in dyskinesia, but the EAG stated that this was inconclusive. In addition, multiple clinic visits and PKG assessments may be needed before people's Parkinson's disease reaches a controlled state.

STAT-ON

No data was found on the impact of STAT-ON on clinical outcomes or changes in clinical management. But an ongoing RCT may provide such data in the future (estimated completion date is December 2022). Most identified data was on diagnostic accuracy, and the EAG judged that the risk of bias in these studies was unclear.

Overall, the EAG concluded that the diagnostic accuracy evidence suggests that STAT-ON is a promising technology. However, the EAG considered that the lack of clinical evidence for STAT-ON means that it is not currently possible to determine if its diagnostic value will translate into real clinical value for patients.

Cost effectiveness

Limited clinical evidence meant that only monitoring strategies for the PKG and Kinesia 360 could be assessed for cost effectiveness. For the one-time use monitoring strategy, the EAG's base case resulted in ICERs of about £67,000 per QALY gained for the PKG using restricted analysis (unrestricted analysis ICERs were much higher), and around £38,000 per QALY gained for the Kinesia 360. For the routine use strategy, the ICERs for the PKG were about £58,000 per QALY gained, and for the Kinesia 360 the ICERs were about £67,000; again, the unrestricted analysis ICERs were much higher.

Scenario analyses for the Kinesia 360 resulted in only minor differences in ICERs. For the PKG, the largest change in ICER came from modelling a recurrent use scenario. This reduced the PKG ICER to £32,417 per QALY gained for the restricted analyses, or £96,675 per QALY gained for the unrestricted analysis. Limiting PKG use to once per year also substantially reduced the ICER to £36,973 per QALY gained for the restricted analyses, and £110,260 per QALY gained for the unrestricted analysis.

Both the Kinesia 360 and PKG increased costs compared with standard care in the EAG's model. The major contributor to this was device cost.

A recently published model (Chaudhuri et al. 2022) produced very different cost-effectiveness results for the PKG.

The main drivers of cost effectiveness in the EAG's model were:

- the direction and size of changes on the UPDRS scale associated with remote monitoring strategies
- the persistence in changes to UPDRS over time (effect waning)
- the number of devices requested (PKG) or length of subscription (Kinesia 360).

5 Issues for consideration

Clinical effectiveness

Limited data and replication across studies

There was limited replication of data on key outcomes. The EAG highlighted concerns about the robustness of the clinical effectiveness data and emphasised that many of the review conclusions were based on individual studies.

The PKG had the largest number of studies (including the largest comparative study). It is uncertain if data generated using this device would be a reliable indicator of how well the other technologies would perform, and the EAG cautioned against this. The only other comparative studies identified (for the Kinesia 360) were small studies.

Generalisability of data

Almost all of the studies were in people having pharmacological therapy, primarily levodopa. The EAG highlighted that there is little evidence on the possible benefits of the devices for people on other therapies, such as non-pharmacological or more advanced therapies (such as deep brain stimulation) that may only be used for people with advanced Parkinson's disease. The EAG did not think it safe to assume that any clinical benefits observed would necessarily apply to these other patient groups.

There was limited data comparing using the devices with not using them. The use of the devices and the level of assessment in the comparator arms described in the identified studies may differ to NHS practice, so estimates of device impact may differ to what would be seen in the NHS.

Subgroups

There was no evidence specifically related to people with communication barriers or difficulties, specific family backgrounds, or for socio-economic

status. The EAG stated that it is unclear how the technologies might perform in these populations.

Incidence of Parkinson's disease may vary by family background, so any benefit of the devices may be higher for some groups. People with Parkinson's disease from the black, Asian and minority ethnic groups may have an atypical pattern of Parkinson's disease, so devices that can help monitor the condition (in additional to clinical judgement) could be particularly beneficial.

At scoping, people with communication barriers were identified as a group who may particularly benefit from the technology, as they may have difficulty describing symptoms.

Cost effectiveness

Limited evidence upon which to base cost-effectiveness estimates for most devices

The EAG identified no evidence to reliably establish the clinical value of the KinesiaU, PDMonitor or STAT-ON, so these devices were not assessed in the base case. Data used to model the cost effectiveness of the Kinesia 360 was from a small pilot RCT.

Size and duration of the impact of device use

The size and persistence of improvement in symptoms from using the devices to guide decisions about care were major drivers of cost effectiveness, and both were uncertain. The EAG commented that there is no evidence on the long-term use or repeated use of the technologies. It is currently uncertain for how long the observed clinical benefits will persist, or how frequently the technologies should be used to maintain clinical benefit. The size of benefit (difference in UPDRS score) was based on studies that may not represent NHS care, in terms of how often the devices may be used and how well the comparator arms may represent current care in the NHS.

How the devices would be used in the NHS

Cost-effectiveness estimates varied depending on how the devices were modelled as being used (one-time, routine, recurrent). The EAG commented that using different configurations to those modelled in its analysis may substantially change the results. How the devices may be used in the NHS is uncertain. For example, as described in the scope, in the University Hospitals Plymouth NHS Trust, the PKG is now used largely in place of review appointments.

The impact of use, and cost effectiveness, may vary depending on whether the devices are used at set intervals or in response to issues with medication flagged by people with Parkinson's disease, their carers or clinicians. Greater benefit of the PKG was reported in the Woodrow study for people whose symptoms were identified as 'out of target'. At scoping, it was identified that the devices may only be used when considered particularly beneficial to help decisions about care, such as if motor fluctuations are not being adequately managed. The EAG did not model this population because of concern that people who may be prescribed a PKG because they were identified as having potential medication issues may differ from this population as defined in Woodrow et al. (2020) who were identified after use of the PKG. However, the EAG's model does allow use of data from this subpopulation, which improves the cost-effectiveness estimates for the PKG. For example, for one-time use of the PKG (adjusted data and restricted analysis), the ICER decreased to about £49,000 per QALY gained (from about £67,000 per QALY gained) if using data from the 'out of target' population. For recurrent use (restricted analysis), the ICER decreased to about £23,000 per QALY gained (from about £32,000 per QALY gained). If using unrestricted analysis UPDRS values, the ICER is about £28,000 per QALY gained.

Device costs

Device cost was the largest component of higher costs compared with standard care. This depends on how often the devices would be used in the

NHS (which is uncertain) and also the cost per use. Any costs associated with interoperability (the ability of computer systems or software to exchange and use information from the devices) that could be incurred through adoption of the devices were not included in the assessment.

Impact of device use on resources

The EAG commented that costs associated with consultation appointments may be underestimated in the model. People with Parkinson's disease interact with a variety of healthcare professionals; a UK survey reported engagement across 18 healthcare professions. The authors calculated the average NHS consultation cost for Parkinson's disease to be £443 per year (2015), compared with the standard care consultation cost in the EAG's analysis of £140 per year (from face-to-face and remote specialist nurse consultations). The EAG commented that a broader consideration of the healthcare professions involved with patient consultation may provide more information on potential cost savings from the devices allowing more remote appointments or reducing the need for consultations.

Chaudhuri et al. (2022) estimated that the PKG would be cost saving by £17,362, whereas in the EAG's modelling, the PKG was estimated to be cost incurring. Modelling of resource costs differed between the models. The impact of device use in the EAG's model reduced consultation appointment costs. In Chaudhuri et al. (2022), improvements in symptoms caused by device use reduced costs (see table 23 in the diagnostics assessment report). The EAG's model did not incorporate any cost savings associated with improvement in symptoms. The EAG raised concerns about the approach used in Chaudhuri et al. to estimate these costs. It was unable to verify or validate the approach used. It stated that it was also unclear whether the observed symptom improvements resulting from use of the technologies could realistically translate into large-scale changes in healthcare use.

Lack of cost-effectiveness estimates for people with advanced Parkinson’s disease

The cost effectiveness of the devices for advanced Parkinson’s disease was not modelled because of a lack of long-term evidence. Using the devices to inform use of device-assisted therapies such as deep brain stimulation (for advanced Parkinson’s’ disease) was identified as a potential use in the scope. The lack of evidence on the impact of the devices on people with advanced Parkinson’s disease also meant that the EAG did not think it appropriate to model a longer time frame.

Restricted and unrestricted UPDRS scores

The restricted analysis assumed that the PKG would have, on average, no detrimental impact on UPDRS domains I (mentation, behaviour and mood) and II (activities of daily living). The cost-effectiveness estimates from the unrestricted analysis, in which the UPDRS II score got worse (increased; see table 2) with the use of the PKG, based on data from Woodrow et al., were very different to those from the restricted analysis.

Potential uncaptured benefits

The impact of the devices, either directly or because of improving symptoms, on the carers of people with Parkinson’s disease was not included in the model. This may have underestimated benefit.

The benefits of remote monitoring may be larger for people with difficulties attending consultations or those accessing care from services at full capacity.

6 Equality considerations

NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others.

Parkinson's disease predominantly affects older people and is more common in men than women. Many people with Parkinson's disease may be protected

under the disability provision of the Equality Act because their condition is likely to have long-term adverse effects on their ability to do normal day-to-day activities. People who are frail or have cognitive impairment or both may struggle to use the technology. The technology is not suitable, or may not work as well, for people who have restricted movement, for example people who are bed bound or who use wheelchairs.

People with Parkinson's disease from black, Asian and minority ethnic groups may have an atypical pattern of Parkinson's disease that is not often recognised by healthcare professionals (NICE Medtech innovation briefing 258 on the PKG). Recent findings from a UK study suggest differences in the phenotype of Parkinson's disease in people from black, Asian or other minority ethnic groups, with a greater burden of non-motor symptoms, motor disability and a higher rate of cardiovascular comorbidities (Sauerbier et al. 2021). Incidence of Parkinson's disease may vary by family background.

Clinical experts highlighted that this technology may offer additional value to people experiencing problems communicating their symptoms. This may include people with language barriers, people with recall problems and people who live alone who may not notice changes in their symptoms. This could benefit people with cognitive disorders and people who do not speak English as a first language. An expert warned that the technology should not replace high-quality interpreters.

Clinical experts highlighted that it is important for training and other user-support resources to be accessible for people with hearing loss or visual impairment.

Improved remote management of Parkinson's disease may improve health outcomes for people in more rural or remote areas. Wider availability of remote appointments may also allow greater access to care for people who are less able to afford travel to in-person appointments.

7 Implementation

IT issues

Interoperability issues (the ability of computer systems or software to exchange and make use of information from the devices) and capacity issues in NHS Trusts may be a potential barrier to implementing technologies for the remote and continuous monitoring of the symptoms of people with Parkinson's disease.

Clinical opinion

Clinical opinion on the usefulness of the devices may influence the choice to use them. A perceived limit in the advantages of using the technology, coupled with reading the reports increasing workload, may deter use.

Differences between monitoring devices

The 5 devices vary in terms of where they are worn on the body, their outputs and how reports are generated. Some devices may be better suited for specific populations (for example, for people with symptoms restricted to their lower limbs, or people with restricted internet access).

8 Authors

Vera Unwin
Topic lead

Thomas Walker
Technical adviser

Glossary

Bradykinesia

Slowness of movement. A diagnosis of Parkinson's disease is considered when bradykinesia plus either tremor or rigidity is present.

Dyskinesia

Dyskinesias are involuntary, erratic, writhing movements of the face, arms, legs or trunk.

Tremor

Tremor tends to occur in the hands and is often described as 'pill-rolling'. But it can also appear in other parts of the body, including the lower lip, jaw or leg. Some people report an internal tremor, a shaking sensation inside the chest, abdomen or limbs that cannot be seen.

Levodopa

Levodopa is recommended as the first-line treatment for people with Parkinson's disease whose motor symptoms impact their quality of life (NICE guideline on Parkinson's disease in adults (NG71). Levodopa can itself cause motor-symptom side effects.

Deep brain stimulation

Deep brain stimulation is a type of surgery used to treat advanced symptoms of Parkinson's disease. A pulse generator (a device like a heart pacemaker) is placed under the skin around the chest or stomach area. The generator is connected to 1 or 2 fine wires that are inserted into the brain, which deliver high frequency stimulation to targeted areas of the brain. These change electrical signals in the brain that cause the symptoms of Parkinson's disease.

Unified Parkinson’s Disease Rating Scale (UPDRS)

The UPDRS is a rating tool used to gauge the severity and progression of Parkinson’s disease. The UPDRS scale consists of the following 6 segments: I) mentation, behaviour, and mood; II) activities of daily living; III) motor examination; IV) complications of therapy (in the past week); V) modified Hoehn and Yahr scale; and VI) Schwab and England activities of daily life scale.

Parkinson’s Disease Questionnaire 39 (PDQ-39)

The 39-item questionnaire offers a patient-reported measure of health status and quality of life. It assesses how often people with Parkinson's disease experience difficulties across 8 dimensions of daily living including relationships, social situations and communication.

Movement Disorder Society – Unified Parkinson’s Disease Rating Scale (MDS-UPDRS)

The MDS-UPDRS is a revised version of the UPDRS. It assesses both motor and non-motor symptoms associated with Parkinson's disease. The MDS-UPDRS has 4 parts: I) non-motor experiences of daily living; II) motor experiences of daily living; III) motor examination; and IV) motor complications.

Gait

A gait is a pattern of limb movements made during walking.

Hoehn and Yahr scale

The Hoehn and Yahr scale describes 5 stages of Parkinson’s disease progression, based on symptoms and the level of clinical disability.

Polysomnography

Polysomnography is a systematic process used to collect physiological parameters during sleep. Polysomnography uses measurements to evaluate underlying causes of sleep disturbances.