Can remote patient monitoring be the new standard in primary healthcare, post-COVID-19?

Authors: Ashley Elizabeth Muller*, Rigmor C. Berga,b, Patricia Sofia Jacobsen Jardima, Trine Bjerke Johansena, and Sari Susanna Ormstad

*aDepartment of Reviews and Health Technology Assessments, Norwegian Institute of Public Health (NIPH), PO Box 222, Skøyen, N-0213 Oslo, Norway

bDepartment of Community Medicine, The Arctic University of Norway, University of Tromsø (UiT), Tromsø, Norway

*Corresponding author
Can remote patient monitoring be the new standard in primary healthcare, post-COVID-19?

Abstract

Background: One lesson from the current COVID-19 pandemic is the need to optimize healthcare provision outside of traditional settings, and potentially over longer periods of time. An important strategy is remote patient monitoring (RPM), allowing patients to remain at home while they transmit data and receive follow-up services.

Methods: We conducted an overview of the latest systematic reviews that had included randomized controlled trials with adult patients with chronic diseases. We summarized results and displayed these in forest plots, and used GRADE to assess our certainty of the evidence.

Results: We included four systematic reviews that together reported on 11 trials that met our definition of RPM, each including patients with diabetes and/or hypertension. RPM probably makes little to no difference on HbA1c levels. RPM probably leads to a slight reduction in systolic blood pressure, with questionable clinical meaningfulness. RPM probably has a small negative effect on the physical component of health-related quality of life, but the clinical significance of this reduction is uncertain. We have low confidence in the findings that RPM makes no difference to the remaining five primary outcomes.

Discussion: Most of our findings are consistent with reviews of other, broader definitions of RPM. The type of RPM examined in this review is as effective as standard treatment for patients with diabetes/hypertension. If this or other types of RPM are to be used for “long covid” patients or for other chronic disease groups post-pandemic, we need to understand why RPM may negatively affect quality on life.

Introduction

The COVID-19 pandemic has transformed remote care provision, in what Wilhite and colleagues have declared “the telemedicine takeover” [1]. Provision has been rapidly implemented and scaled up in many places, as healthcare settings have needed to reduce face-to-face contact to ensure
social distancing, and triage care provision to accommodate pandemic-specific responsibilities. Other settings are remotely monitoring suspected COVID-19 patients [2], while several countries collect anonymized symptom data on a large scale through apps. As we build an evidence base about the chronicity of “long covid” or “post covid” (see …[3] for a brief discussion), remotely monitoring COVID-19 patients post-discharge, or those with risk factors for persistent symptoms, could be a way for healthcare providers to respond quickly with tailored care.

Already before the current pandemic, healthcare systems faced the challenge of managing, rather than curing, the chronic diseases with which people are living longer. Management includes preventing the impairment of functioning and helping people maintain as good a quality of life for as many years as possible [4]. This requires frequent assessments of the patient’s health status in order to monitor treatment regime, prevent deterioration, and prevent the development of additional diseases, injuries, and complications. Yet, neither providers nor patients have the capacity to meet face-to-face to conduct such frequent assessments.

One solution is for patients to be able to transmit health data without seeing providers, and for this data to be sent and evaluated often enough to initiate interventions or treatment adjustments before the patient's health status becomes acute. Strategies that allow patients to remain at home while they transmit data and receive follow-up services can be collectively referred to as remote patient monitoring (RPM). Watson et al. [5] have recently argued that more than traditional rural access and video communication, RPM may be uniquely situated to address care needs in the context of COVID-19.

No generally accepted definition of the term RPM exists. The terms telemedicine, telehealth, and eHealth are often used interchangeably. For example, three recent Cochrane Reviews use different terms for similar interventions: In Kew et al. [6], remote monitoring refers to interventions that allow patients to share data using information and communication technologies and healthcare providers to respond, and is situated as “a form of ‘telehealth’, otherwise referred to as ‘telecare’, ‘digital health’, ‘mHealth’ or ‘telemedicine’” (page 7). McLean et al. [7] writes that “telemedicine” implies health care is being delivered, and instead uses telehealthcare to mean the electronic transfer of patient data and the receipt of provider feedback. In Flodgren et al, [8], interactive telemedicine specifically means providers respond to patient data transmission in real time, while remote monitoring services also include provider feedback, but not in real time (page 7). Two 2021 overviews of systematic reviews used telehealth to cover all patient-provider interaction, including remote patient monitoring, except when conducted over the phone or using non-interactive web-
sites \cite{9,10}. Mordaunt \cite{11} pointed out in a recent commentary that a lack of precise definition, particularly in a systematic review, seriously hampers quantitative meta-analyses as well as qualitative summaries of effect.

The Norwegian Directorate of Health (DoH) is interested in implementing a specific type of RPM that occurs within the primary healthcare services. We worked closely with the DoH to describe this type of RPM, as described in the review protocol \cite{12}. First, the patient answers questions about their own health condition using a digital device, and/or takes measurements of metabolic data related to their diagnosis using digital devices. This health data is then transmitted to a provider. In step two, data evaluation, the patient's data is received and evaluated by the provider. Alternatively, the assessment may be automated (i.e. evaluated by a program, as in a “traffic light” system), and the program forwards data it evaluates as high-risk to healthcare providers for further follow-up. In the third and final step, follow-up, a provider follows up with the patient if the patient’s health data indicates a concern.

**Research question**

The relevance to Norway, and countries with similar healthcare system needs, of the types of RPM implemented in previous studies is unclear. The DoH needs knowledge of the effectiveness and cost utility of RPM, and particularly which patient group may benefit most. This review therefore sought to answer the following question: what is the effect of a specific type of RPM on clinical and healthcare utilization outcomes of certain groups of chronically ill patients?

**Methods**

We conducted an overview of reviews. A study protocol was developed by the research team, peer-reviewed, approved by the DoH, and published in Norwegian and in English \cite{12,13}.

**Inclusion and exclusion criteria**

We searched for overviews of reviews and systematic reviews, with the following inclusion criteria:

- **Population**: Persons who are 18 years or older, have a chronic disease (cancer, cardiovascular disease, chronic lung diseases, chronic musculoskeletal disorders, diabetes, hypertension, impaired vision/hearing, mental disorders, or osteoporosis), and are neither in the very early nor very acute phase of these conditions.
- **Intervention**: RPM according to the definition above; RPM provided in the primary healthcare services; RPM involving phones, mobile phones, videos, and portable/implantable devices; data sent regularly (at least twice per year).
• Comparison: Standard care that does not involve RPM; other type of RPM.
• Outcomes: Mental health (symptoms or diagnoses); diagnosis-specific physical health; physical functioning level; quality of life; consumption of health services (hospital admissions, emergency care, number of bed-days, outpatient consultations, nursing home stays, home care [both home nursing and practical assistance], and general practitioner consultations); health services costs.
• Year: Search for literature conducted 2015 or later.

We excluded reviews if participants were reported to have reduced cognitive function, as they may not be able to report their own health outcomes. Reviews utilizing internet-based RPM or RPM executed through mobile applications on phones or tablets were also excluded, as per DoH interest. Finally, reviews that explicitly excluded Norway or the part of the world in which Norway is located, e.g. reviews of low- and middle-income countries, were also excluded. We had no language exclusions a priori.

Literature search
An information specialist developed and conducted systematic searches for literature in MEDLINE, Embase, PsycINFO, Epistemonikos, Cochrane Database of Systematic Reviews, and Web of Science. We employed both subject headings (e.g. MeSH terms in MEDLINE) and free text terms characterizing the intervention and population. The complete search strategies and results are in Appendix 1.

Study selection
Two review authors independently assessed all titles and abstracts from the systematic literature search for eligibility using Rayyan [14]. Full-text publications were retrieved when one or both author(s) judged the review to likely meet the inclusion criteria. Full-text publications were then read by two authors independently using Covidence [15], with final inclusion based on consensus by the two authors. As anticipated in the protocol, few of the overviews and systematic reviews read in full-text described interventions thoroughly enough for us to determine eligibility. Therefore, when assessing a systematic review, we retrieved each review’s included primary studies and assessed eligibility after reading in full-text. If a systematic review included at least one primary study that met all our inclusion criteria, the entire systematic review was included in our review, along with only the primary study that met our inclusion criteria. (When reading an overview of systematic reviews, we did not proceed to primary studies, but read the full-text of each included systematic
Assessment of included systematic reviews

After a review was read in full-text and determined to meet our definition of RPM and the other inclusion criteria, we assessed its methodological quality using the NIPH’s checklist for systematic reviews (Appendix 3). Two authors independently assessed methodological quality and met to discuss conflicts. Any disagreements were resolved through discussion. Only reviews rated as having high methodological quality were included in our review; in practice, this required that a review met all items on the checklist.

Assessing risk of bias in included primary studies

We extracted and presented systematic review authors’ own risk of bias assessments of included primary studies. All the reviews with relevant primary studies included RCTs, and used the Cochrane Risk of Bias Tool. The authors of one review [16] modified this tool slightly by not reporting blinding of participants and personnel and other biases and instead reporting funding as a separate (risk of bias) criterion. We chose to report systematic review authors’ judgement on funding as part of the assessment of other biases.

Data extraction and synthesis

One author extracted data from the included systematic reviews, and another author double-checked data extraction for accuracy and completeness. If the systematic review(s) did not sufficiently report findings or characteristics, we proceeded to the RCT itself to extract the necessary information. As neither entire overviews of reviews nor systematic reviews met our inclusion criteria, we summarized only the data from the relevant RCTs the reviews contained.

Interventions lasted six months (four RCTs), nine months (three RCTs), or twelve months (four RCTs). When an RCT measured an outcome at multiple time points, the most recent measurement was used. In one RCT, data was collected three months after the completion of the intervention; the remainder of the RCTs collected outcome data at intervention completion. We presented normally distributed results for each primary outcome in a forest plot and reported raw mean differences, standardized mean differences, odds ratios, or rate ratios; non-normally distributed outcomes were reported as medians. However, to avoid misleading readers into thinking that our overview included a meta-analysis (inappropriate because this review is an overview of systematic reviews, and not a systematic review that exhaustively searched for and identified RCTs), we did not produce the summary statistic within forest plots or report these summary statistics in
the summary of findings table. Detailed results and forest plots for twenty-three secondary outcomes are available upon request.

Data for one outcome, HbA1c, was only available as adjusted for baseline values in both Dario et al. [17] and Egede et al. [18], and was presented alongside the remaining RCTs’ unadjusted outcomes. One RCT [19] contained more than two arms: usual care, high-intensity RPM, and low-intensity RPM. The high-intensity RPM differed only from the low-intensity arm in that the former included automated messages that were more tailored to each patient, compared to the latter. We analyzed only data from the high-intensity arm, as dividing this RPM into two comparisons would have duplicated the usual care group’s data.

Missing data
Several RCTs failed to report standard deviations. For the purposes of visualizing outcomes in forest plots, we borrowed standard deviations from RCTs with the most similar patient population. Wakefield et al.’s [19] missing systolic blood pressure standard deviations were borrowed from Magid et al. [20], due to both patient populations coming from the United States and having co-occurring hypertensive and diabetic patients. Schillinger et al.’s [21] standard deviations were used for Carter et al.’s [22] blood pressure and BMI, outcomes due to both patient groups being American, obese, urban, and with racial minorities overrepresented.

Assessment of certainty of the evidence
We assessed the certainty of the evidence for each of the seven primary outcomes using GRADE (Grading of Recommendations Assessment, Development, and Evaluation) [23]. GRADE is a method for assessing the certainty of the evidence in systematic reviews, and can be used even when meta-analytic pooled effect estimates are not available [24]. An assessment for each primary outcome is conducted using five criteria: systematic review authors’ assessment of primary study methodological quality, degree of inconsistency, indirectness, imprecision, and publication bias. The GRADE assessments were conducted using the software GRADEpro [25].

Protocol deviations
During our full-text review of RCTs, we decided to include patients with hypertension, although they were not one of the eight original chronic condition groups. This was because of the overlap between diabetes and hypertension among many of the included RCTs’ patients, and because hypertension is a common comorbidity with many of the other chronic conditions of interest. The only practical consequence of this protocol deviation was to allow the inclusion of one RCT, Magid
et al. [20], which recruited hypertensive patients and reported that nearly half also had either diabetes or renal disease.

Results

Results of the literature search

The literature search for reviews resulted in 3,373 unique records, as exhibited in Figure 1. We excluded 151 publications after full-text review, most commonly for not reporting on RPM as defined by the DoH, not being a systematic review, or not being of high methodological quality. We included four systematic reviews [16, 26-28]. Appendix 2 lists all publications excluded after full-text review, with reasons for exclusion and chronic disease category.
Description of the included systematic reviews

The four included systematic reviews searched for randomized controlled studies (RCTs), cluster RCTs, quasi-RCTs, controlled before-and-after studies, and interrupted time series studies of different types of remote communication or healthcare delivery. Every review defined their intervention of interest differently and with a slightly different name, as displayed in Table 1. Bittner et al.
searched for *telerehabilitation* services, explicitly including remote monitoring within this definition, for patients with impaired vision. Faruque et al. [16] searched broadly for *teledmedicine* interventions, defined as all electronic forms of communication, among diabetes patients. Kebede et al. [28] focused on diabetes type 2 patients using *digital interventions*, meaning any technology-based intervention. Posadzki et al. [27] searched for *eHealth* interventions among patients with long-term conditions. Faruque et al. [16] excluded studies that involved patients with gestational diabetes, and Kebede et al. [28] excluded studies with diabetes type 1 patients. Aside from this, there were no other disease-related exclusion criteria specified by the systematic reviews.

**Table 1 Description of included systematic reviews**

<table>
<thead>
<tr>
<th>Systematic review</th>
<th>Search date</th>
<th>RCTs*</th>
<th>Description of RPM, in the authors' words / interventions of interest</th>
<th>Chronic disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bittner et al. [26]</td>
<td>June 2015</td>
<td>0 included in this review; 0 analyzed by the authors.</td>
<td>Telerehabilitation: rehabilitation services delivered via information and communication technologies, and including monitoring and clinical evaluation services.</td>
<td>Low vision</td>
</tr>
<tr>
<td>Faruque et al. [16]</td>
<td>Nov. 2015</td>
<td>7 included in this review, out of 111 analyzed by the authors: Carter et al. [22], Nicolucci et al. [29], Rodriguez-Idigorioz et al. [30], Schilinger et al. [21], Steventon et al. [31], Stone et al. [32], Wakefield et al [19].</td>
<td>Teledmedicine: all electronic forms of communication between provider and patient (telephone, smartphone application, email, text messaging, web portal, “smart” device or glucometer).</td>
<td>Diabetes type 1 or 2</td>
</tr>
<tr>
<td>Kebede et al. [28]</td>
<td>June 2017</td>
<td>3 included in this review, out of 21 analyzed by the authors: Dario et al. [17], Egede et al. [18], Wild et al [33].</td>
<td>Digital interventions: technology based, such as m-health interventions, web-based interventions, interventions delivered through the use of a personal digital assistant, a tablet, a computer, the Internet, teledmedicine, videoconferencing, telehealth, or other forms of e-health.</td>
<td>Diabetes type 2</td>
</tr>
<tr>
<td>Posadzki et al. [27]</td>
<td>June 2015</td>
<td>1 included in this review, out of 132 analyzed by the authors: Magid et al. [20]</td>
<td>eHealth interventions: interventions that use devices featuring interactive wireless communication capability, operating web-based applications and with high portability (such as smartphone, computer and personal digital assistance tools), or interventions comprising self-care, self-management, self-care behavioral change or education dissemination.</td>
<td>Any long-term condition</td>
</tr>
</tbody>
</table>

*We included and extracted data only from the RCTs that met our inclusion criteria.*
Descriptions of the included RCTs from the reviews

With the exception of Bittner, a review that found no relevant studies, these reviews' interventions of interest were defined broadly, and therefore the RCTs they analyzed also reported on a wide array of interventions. Altogether, only eleven RCTs implemented an intervention that met our definition of RPM. These RCTs were conducted in Italy, Spain, the United Kingdom, and the United States. None of these RCTs occurred in more than one of the included reviews. Seven RCTs involved diabetes type 2 patients, one RCT included both type 1 and 2 diabetes patients, two RCTs included patients with both hypertension and diabetes, and one RCT included only hypertensive patients.

Patients' health status

The RCTs included patients with average ages ranging from 51 to 68. Six RCTs reported on participants’ co-/multi-morbidities. About four of ten patients in Rodriguez-Idigoras et al. [30] and Nicolucci et al. [29] reported dyslipidemia. Chronic obstructive pulmonary disease was reported by 6.3-14.2% in Dario et al. [17], Steventon et al. [31], and Stone et al. [32], and heart failure by 3.6-16.1% in Steventon et al. [31], Stone et al. [32], and Nicolucci et al. [29]. Stroke was reported by 2.4-4.7% of the patients in Dario et al. [17] and Nicolucci et al. [29].

Given these similarities in multi-morbidities, all eleven RCTs could be grouped together as involving “diabetes and/or hypertension” patients.

It is worth noting that psychiatric morbidities were not reported by any of the RCTs. They were exclusion factors of five RCTs in some manner, such as “reliance on psychotropic medication” [22], “mental conditions, depression, or high anxiety;… abuse of drugs or alcohol” [29], “alcohol or drug abuse/dependency, active psychosis or acute mental disorder” [18], “psychotic illness” [21], and “psychosis” [19].

Descriptions of the various types of RPM used in the included RCTs

In accordance with our definition, RPM referred to the three steps of digital data transmission, evaluation, and follow-up. Still, there were a variety of devices used to transmit data, three methods of data evaluation, and some variation in the method of follow-up response given to patients, as summarized in Table 2. In all RCTs, patients collected and transmitted up to three biometric measurements: blood glucose, blood pressure, and weight. Frequency of patient data transmission varied from three times a day to twice a month, with two RCTs individualizing frequency
according to clinical histories. Patient data was evaluated either manually, automatically by a monitoring center, or automatically by the RPM device itself. Follow-up was individualized medical care such as medication adjustment, discussion of adherence, counselling on behavioral changes such as diet, smoking, weight management, and physical activity, and support for other conditions. Care was often described as focusing on helping patients self-manage their conditions. In about half of the RCTs, patients were only followed up with by providers if their data had been evaluated (manually or automatically) as concerning. In the remaining RCTs, patients received scheduled follow-up regardless of data values; both scheduled follow-up and follow-up indicated by concerning data; or automated responses if data was not of concern, and personal follow-up if data was concerning.

Table 2 Description of RPM implemented in the included RCTs

<table>
<thead>
<tr>
<th>Author</th>
<th>Chronic disease</th>
<th>Data transmission</th>
<th>Data evaluation</th>
<th>Follow-up response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carter et al.</td>
<td>Diabetes type 2</td>
<td>Weight and blood pressure sent 1/week and blood glucose sent 3/day; using a laptop that was equipped with a wireless scale, a blood pressure cuff, and a glucometer</td>
<td>Manual review by telehealth nurse.</td>
<td>Nurse discussed data with patient over video conference during biweekly calls.</td>
</tr>
<tr>
<td>[22]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dario et al.</td>
<td>Diabetes type 2</td>
<td>Blood glucose measured (frequency not reported) with a glucometer connected to a telecare device that sent data to an eHealth center</td>
<td>Alerts automatically generated by eHealth center if data values crossed pre-specified thresholds.</td>
<td>If automatic alert was generated, eHealth staff contacted clinician. Clinician took subsequent action according to normal protocols and contacted patients by telephone or other unspecified methods. If an emergency, eHealth center contacted next of kin and emergency department.</td>
</tr>
<tr>
<td>Study</td>
<td>Condition/Type</td>
<td>Measurement Details</td>
<td>Method/Alerts</td>
<td>Additional Contact</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------</td>
<td>-------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Egede et al. [18]</td>
<td>Diabetes type 2</td>
<td>Blood glucose and blood pressure sent 1/day, using a commercial telehealth device that uploaded blood glucose and blood pressure to a central server</td>
<td>Manual review by nurse case manager If necessary, nurse contacted patients by telephone to make medication adjustments. Additional contact: Nurse case managers made weekly reminder calls to upload data.</td>
<td></td>
</tr>
<tr>
<td>Magid et al. [20]</td>
<td>Hypertension</td>
<td>Blood pressure sent 1/week over the patient’s usual telephone, using an interactive voice response phone system</td>
<td>Manual review by pharmacist If data values exceeded guideline-recommended treatment goals, pharmacists contacted patients to review medication adherence, adjust medications, and provide counselling on healthy therapeutic lifestyle changes, using the interactive voice response system or telephone. Pharmacists contacted GP in the case of medication adjustments.</td>
<td></td>
</tr>
<tr>
<td>Nicolucci et al. [29]</td>
<td>Diabetes type 2 and hypertensive (&gt;130/80 mmHg)</td>
<td>Blood glucose, blood pressure, and weight sent 2/month, using a weight scale, glucometer, and a sphygmomanometer, respectively, connected via Bluetooth to a device that transmitted data in real-time to a telehealth center.</td>
<td>Alerts automatically generated by telehealth center if data values concerning Telehealth center nurses forwarded alerts to GPs, who contacted patients. Additional contact: Telehealth nurses also contacted patients monthly to discuss results and barriers to compliance, using text messages, e-mail, or telephone.</td>
<td></td>
</tr>
<tr>
<td>Rodriguez-Idigoras et al. [30]</td>
<td>Diabetes type 2</td>
<td>Blood glucose measured using a glucometer and sent via patient’s usual telephone to a call center; no required</td>
<td>Alarms automatically generated by call center if data Call center staff contacted GP and patient by telephone. Unspecified “standard protocols” were followed.</td>
<td></td>
</tr>
<tr>
<td>Study Authors</td>
<td>Study Type</td>
<td>Blood Measures</td>
<td>Traffic Light System</td>
<td>Additional Support</td>
</tr>
<tr>
<td>---------------------</td>
<td>------------</td>
<td>----------------</td>
<td>----------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Schillinger et al.</td>
<td>Diabetes</td>
<td>Blood glucose and blood pressure sent 1/week, using telephone touchpads during an automated telephone call</td>
<td>Evaluated by an automated telephone support system</td>
<td>The telephone system either immediately responded with a narrated health education message, or the system alerted a nurse, and the nurse contacted the patient.</td>
</tr>
<tr>
<td>Steventon et al.</td>
<td>Diabetes</td>
<td>Blood glucose and blood pressure sent up to 5/week, with the frequency adjusted according to participants' individual clinical histories, using a freestanding telehealth unit or a television set top box that connected to a blood pressure monitor and glucometer or to weighing scales / pulse oximeters</td>
<td>Traffic light system: automatic evaluation at monitoring center</td>
<td>If &quot;red&quot;, monitoring center staff reviewed data 1/day and contacted the patient for further evaluation, to offer disease management advice, or to give referrals. Contact was made using the telehealth unit or other unspecified methods.</td>
</tr>
<tr>
<td>Stone et al.</td>
<td>Diabetes</td>
<td>Blood glucose, blood pressure, and weight sent 1/day, using a commercial home telemonitoring device that transmitted measurements to a central server</td>
<td>Traffic light system: automatic evaluation by the device</td>
<td>If &quot;red&quot;, nurse contacted patients and adjusted medication, over the telephone or using the home monitoring system. Additional support: monthly calls to provide individualized self-management counseling tailored to specific issues, based on data values.</td>
</tr>
<tr>
<td>Wakefield et al.</td>
<td>Diabetes</td>
<td>Blood glucose and blood pressure sent 1/day, using a commercial home telehealth device that sent and received</td>
<td>Manual review by nurse, 1/day</td>
<td>Tailored, automated responses sent based on data. Nurses reviewed data daily and contacted the patients if necessary.</td>
</tr>
</tbody>
</table>

Values outside normal range.
<table>
<thead>
<tr>
<th>Study</th>
<th>Type</th>
<th>Data Collection</th>
<th>Additional Support</th>
<th>Provider Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild et al.</td>
<td>Diabetes type 2</td>
<td>Blood glucose sent at least 2/week, and blood pressure, viewed by primary care nurse or family practice clinician, and weight sent at least 1/week; using Bluetooth-enabled blood pressure, blood glucose, and weight monitors that transmitted data via a supplied modem.</td>
<td>Provider changed treatment if necessary to comport with national guidelines for diabetes and hypertension management if necessary, and provided advice on lifestyle modification, information about medication effects; the method of communicating back to the patient was not specified.</td>
<td></td>
</tr>
</tbody>
</table>

Usual care as a comparator

In all RCTs, participants were recruited from existing general practitioner lists, healthcare networks, or other patient pools, indicating that they had already some minimum amount of contact or usual care with the primary health services. The usual care described in seven RCTs appeared to be quite minimal, such as an educational pamphlet, encouragement to contact providers, or yearly review of health status [18-20, 22, 29, 30, 33]. This was also the case for the two RCTs that specified that they compared RPM adjunct to usual care, with usual care alone [21, 31].

In the remaining two RCTs, usual care contained the same self-monitoring that the RPM group conducted, but without the benefit of digital transmission of this data or guaranteed provider knowledge of health status [17, 20]. In Dario et al. patients in usual care were supposed to measure HbA1c and bring paper logs to providers [17]. In Stone et al., usual care patients were supposed to measure HbA1c, blood pressure, and weight daily, and discuss these with diabetes nurse educators over the phone once per month [32]. Neither of these RCTs reported the actual frequency of self-monitoring, making it difficult to conclude the extent to which usual care in these RCTs differs from usual care in the remaining RCTs.
Self-monitoring may have been practiced by patients in usual care in Rodriguez-Idigoras et al. and Schillinger et al. [21, 30]. Six months’ self-monitoring was an inclusion criterion for Rodriguez-Idigoras et al., in order to recruit among patients already capable of complying with a monitoring regime, although there was no mention of the usual care group being expected to continue. Schillinger et al. specified that if usual care patients were already self-monitoring when enrolled in the RCT, they were encouraged to continue doing so; no estimates were provided of how common this was. In both cases, as in Dario et al. and Stone et al., any self-monitoring conducted by usual care patients would not have been digitally transmitted or evaluated by providers.

Risk of bias in the RCTs

The review authors’ own judgment of each risk of bias domain is presented as percentages across all 11 included RCTs in Figure 2. All to the majority of the RCTs were assessed as having low risk of selection bias, detection bias, attrition bias, reporting bias and other bias. The majority of the RCTs were assessed as having unclear risk of performance bias. In many instances, this was due to non-reporting of blinding of participants and personnel.

Figure 2 Risk of bias graph

Review authors’ own judgements about each risk of bias item presented as percentages across all included studies.
The review authors’ judgements about each risk of bias domain for each included RCT separately are available upon request.

**Effects of RPM on patients with diabetes and/or hypertension**

In this section, each of the subsections include a narrative summary of the findings for a primary outcome, as well as a presentation of the results by means of forest plots. In addition, we give results of the GRADE assessment (our evaluation of the certainty of the evidence) for each of the eight primary outcomes. Table 3 provides an overview of the conducted GRADE assessments.

**Table 3 Summary of findings table: RPM compared to usual care for chronic diseases**

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Effect</th>
<th>No of participants (studies)</th>
<th>Certainty of the evidence (GRADE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HbA1c [%]</td>
<td>Most studies showed a reduction, from 0.23% lower to 1.08% lower. However, only four exceeded the suggested minimum clinically importance difference of ≥0.5% 1. RPM probably slightly reduces Hba1C.</td>
<td>2235 (10 RCTs)</td>
<td>⬤⬤⬤◯ MODERATE a</td>
</tr>
<tr>
<td>Systolic blood pressure [mmHg]</td>
<td>No studies showed a statistically significant effect, but tended to benefit RPM. RPM probably leads to a slight reduction.</td>
<td>1407 (7 RCTs)</td>
<td>⬤⬤⬤◯ MODERATE a</td>
</tr>
<tr>
<td>Diastolic blood pressure [mmHg]</td>
<td>No studies showed an effect. RPM may make no difference.</td>
<td>1207 (6 RCTs)</td>
<td>⬤⬤⬤◯ LOW b,c</td>
</tr>
<tr>
<td>Total cholesterol</td>
<td>No studies showed an effect. RPM may make no difference</td>
<td>664 (3 RCTs)</td>
<td>⬤⬤⬤◯ LOW c,d</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Effect</th>
<th>No of participants (studies)</th>
<th>Certainty of the evidence (GRADE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients with hospitalizations or ER visits, all-cause</td>
<td>No effect. RPM may make no difference.</td>
<td>249 (1 RCT)</td>
<td>LOW e,f</td>
</tr>
<tr>
<td>Quality of life (SF-12/SF-36), mental health component</td>
<td>Two studies showed no effect, and one showed a small benefit to usual care patients. RPM may make no difference.</td>
<td>698 (3 RCTs)</td>
<td>LOW d,g</td>
</tr>
<tr>
<td>Quality of life (SF-12/SF-36), physical health component</td>
<td>Usual care reported higher scores, with a small effect size. RPM probably harms this outcome.</td>
<td>698 (3 RCTs)</td>
<td>MODERATE g</td>
</tr>
<tr>
<td>Hospital Anxiety and Depression Scale total score</td>
<td>No effect. RPM may make no difference.</td>
<td>257 (1 RCT)</td>
<td>LOW f,h</td>
</tr>
</tbody>
</table>

a. In four RCTs, patients differed significantly from the Norwegian patient population (e.g. American war veterans, urban poor, only men)
b. In two RCTs, patients differed significantly from the Norwegian patient population (e.g. American war veterans, urban poor, only men)
c. Effect estimates favor both RPM and usual care.
d. Wide confidence intervals, with studies showing both a moderately negative effect and a moderately positive effect
e. Likely bias related to study funding
f. One study
g. Performance bias, detection bias, attrition bias, and other bias.
h. Wide confidence interval and small number of participants.
<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Effect</th>
<th>№ of participants (studies)</th>
<th>Certainty of the evidence (GRADE)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**GRADE Working Group grades of evidence**

**High certainty:** We are very confident that the true effect lies close to that of the estimate of the effect.

**Moderate certainty:** We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.

**Low certainty:** Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect.

**Very low certainty:** We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect.

Figure 3 displays forest plots for each primary outcome. Forest plots omit group sizes and raw data values, such as mean values for RPM and usual care groups, for ease of reading. These numbers are available from the authors upon request.

Among the ten RCTs that reported HbA1c, there appeared to be a small pooled reduction among the RPM group by the end of interventions (Figure 3). With three exceptions that reported either no difference or a slight benefit to usual care patients ([17, 19, 30]), the RPM patients in the RCTs reported an average of 0.23-1.08% lower HbA1c scores than the usual care groups. However, only in four studies was this reduction above the suggested threshold for clinical meaningfulness, ≥0.5% (see [34-36]). Moreover, due to considerable heterogeneity among the studies and uncertainty about the applicability of the findings, we have low confidence in this finding.

*Figure 3 Forest plots of primary outcomes*
### Outcome 4: Cholesterol

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Mean Difference</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IV, Random, 95% CI</td>
<td>IV, Random, 95% CI</td>
</tr>
<tr>
<td>Stone 2010</td>
<td>-10.90 [-23.93, 2.13]</td>
<td></td>
</tr>
<tr>
<td>Wild 2016</td>
<td>-7.73 [-17.44, 1.98]</td>
<td></td>
</tr>
<tr>
<td>Nicolucci 2015</td>
<td>5.00 [-4.42, 14.42]</td>
<td></td>
</tr>
</tbody>
</table>

### Outcome 5: Patients with hospitalizations or emergency stays

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Risk Ratio</th>
<th>Risk Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IV, Random, 95% CI</td>
<td>IV, Random, 95% CI</td>
</tr>
<tr>
<td>Nicolucci 2015</td>
<td>0.93 [0.44, 1.97]</td>
<td></td>
</tr>
</tbody>
</table>

### Outcome 6: Health-related quality of life (SF-12/36), mental component summary

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Std. Mean Difference</th>
<th>Std. Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IV, Random, 95% CI</td>
<td>IV, Random, 95% CI</td>
</tr>
<tr>
<td>Dario 2017</td>
<td>-0.04 [-0.31, 0.23]</td>
<td></td>
</tr>
<tr>
<td>Schillinger 2009</td>
<td>0.11 [-0.17, 0.38]</td>
<td></td>
</tr>
<tr>
<td>Nicolucci 2015</td>
<td>0.31 [0.06, 0.56]</td>
<td></td>
</tr>
</tbody>
</table>

### Outcome 7: Health-related quality of life (SF-12/36), physical component summary

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Std. Mean Difference</th>
<th>Std. Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IV, Random, 95% CI</td>
<td>IV, Random, 95% CI</td>
</tr>
<tr>
<td>Schillinger 2009</td>
<td>0.12 [-0.16, 0.39]</td>
<td></td>
</tr>
<tr>
<td>Nicolucci 2015</td>
<td>0.16 [-0.09, 0.41]</td>
<td></td>
</tr>
<tr>
<td>Dario 2017</td>
<td>0.28 [0.00, 0.55]</td>
<td></td>
</tr>
</tbody>
</table>
Systolic blood pressure was measured by seven studies, with no differences between the RPM and usual care patients overall. Studies reported mean values both above and below NICE’s recommended target of 135/85 [37]. Due to the consistency of each RCTs’ estimates, it is likely that the true effect of RPM on systolic blood pressure is close to the estimates we see here, namely, a small reduction.

Diastolic blood pressure was measured by six studies, again with no mean differences between the RPM and usual care patients. In all studies, mean diastolic values were below NICE’s recommended target of 135/85 [37]. We are not confident that the true effect of RPM on diastolic blood pressure is close to the effect estimates we see here, both because of moderate heterogeneity and effect estimates spread widely and in directions that favor both RPM and usual care.

Total cholesterol was measured by three studies. None of the studies reported a mean difference between the RPM and usual care patients. We have low confidence in this finding, given large heterogeneity and individual effect estimates that show not only opposite findings, but findings with almost no overlap with one another.

One study reported the amount of patients who had a hospitalization or an emergency room visit during the course of the intervention, with no difference between the two groups. As this finding drew from only one RCT, which was also judged to have a high risk of bias due to commercial funding, we are not confident that the true effect of RPM on this outcome is close to the estimate presented here.

Health-related quality of life (HRQOL) was one of the few patient-reported outcomes. The standardized mean difference of HRQOL in the mental component summary of two HRQOL instruments was no different for the RPM or usual care patients in three studies. We have low certainty in this finding, and RPM may in fact have an effect on this component, potentially negatively. In the physical component summary of the same instruments, the usual care group reported slightly higher mean HRQOL than the RPM group. In this component, the three studies’ effect studies agreed.
more with one another, but our certainty was still downgraded due to potential biases. We are moderately certain that RPM has a small negative impact on physical HRQOL.

One study reported the total score from the Hospital Anxiety and Depression Scale (HADS), with no difference between the RPM and usual care patients.

Only one of 23 secondary outcomes showed an effect of RPM, and this effect was negative: the usual care group weighed 5.40 kg less at the end of the intervention in Wild et al. [33], with a range of 0.42 kg less to 10.38 kg less. Twenty-two of 23 secondary outcomes, reported by one to two RCTs each, showed no effect of RPM. These included achievement of normal blood pressure, achievement of target HbA1c, daytime ambulatory systolic blood pressure, daytime ambulatory diastolic blood pressure, body mass index, overall health-related quality of life, diabetes knowledge, HADS anxiety subscale, HADS depression subscale, contact with general practitioner, practice nurse visits, primary care physician encounters with procedures, amount of emergency department visits, all-cause hospitalization, all-cause emergency hospitalization, bed days for all-cause hospitalization, bed days for diabetes-related hospitalized patients, patients who visited a specialist, amount of outpatient visits, patients with home visits, start-up, ongoing, and total costs, and costs per quality-adjusted life year gained. Forest plots of secondary outcomes are available from the authors upon request.

**Discussion**

This overview of systematic reviews sought to assess the effectiveness of a specific type of remote patient monitoring (RPM) on chronically ill patients’ clinical outcomes and healthcare utilization. Only eleven RCTs captured by four high-quality reviews. We summarized results from 31 outcomes (8 primary and 23 secondary), and RPM had two positive effects (slight reductions in HbA1c and systolic blood pressure) compared to standard treatment, and two negative effects (reduction in the physical health component of health-related quality of life and increase in weight). RPM appeared to have no effect on the remaining five primary outcomes (diastolic blood pressure, cholesterol, the mental health component of health-related quality of life, the number of patients with a hospitalization or emergency room visit, and Hospital Anxiety and Depression Scale scores), and no effect on 22 of 23 secondary outcomes.

We assessed 155 systematic reviews and their approximately 3,500 RCTs, and only four high-quality reviews met our criteria. Of these four reviews’ 176 included RCTs, only eleven reported on interventions that met our definition of RPM, and only with patients with diabetes and/or hypertension. This indicates that the type of RPM of interest to the Norwegian Directorate of Health is
not commonly implemented. However, our findings overall comported with previous reviews that have used broader definitions of RPM, such as those utilizing internet-based technologies, fully automated programs without provider input, and interventions organized in specialist health services.

Reviews with the broadest definitions of RPM have reported clinically meaningful reductions in HbA1c [16], including in primary care settings [38], although both studies’ meta-analyses contained substantial amounts of unexplained heterogeneity. Reviews that have parsed RPM further into component types have found no significant effect on HbA1c (either statistically, clinically, or both) of automated telephone messages [27], text messages [28], website-only [39], and mobile phone-only [39]. Combined website and mobile strategies were found to be effective, -0.77% [39], as was web-based RPM when including personal digital assistants, tablet, computer, and smartphone interventions [28]. The delivery method or components of RPM may moderate the effect it has on HbA1c. RPM using combination mobile-and-website components and RPM using online components were more effective than other single-component technologies, such as text-messaging, automated telephone, website, or mobile technologies. Our included RCTs might have been collectively less effective because they tended to utilize single-component technologies. Our finding of little or no effect of RPM on blood pressure of patients with diabetes and/or hypertension contributes to a mixed body of evidence [8, 27, 40-43].

Given that RPM defined more broadly has been reported to have positive effects on HbA1c and, in some cases, blood pressure among people with diabetes and/or hypertension, our definition of RPM appears to have captured the types that are least effective. If RPM is to be implemented among post-covid patients or other chronic disease patients, we cannot conclude that it is more (or less) beneficial than standard treatment.

Patient groups should be involved in RPM implementation and evaluation, to maximize potential for modification and ultimately efficacy. Patients may have preferences as to the frequency of feedback from providers, the content of such feedback, and even the method of contact. Pekmezaris et al. provide one example of a participatory approach to designing a RPM program for heart failure patients [44], while Ware et al. describe suggestions made by patients for program modification after conclusion [45]. Further impetus to involve patients in the development of effective RPM strategies are the consistent reports of racial, ethnic, and socioeconomic equalities in other types of monitoring and telemedicine programs [46, 47]. RPM strategies must be designed, or re-designed, not only with cost effectiveness, scalability, and health outcomes in mind, but also with health equity as an equally important requirement.
This review was conducted pre-pandemic. How relevant are our findings now, in the context of COVID-19? We are aware of one recent rapid review of “remote home monitoring” organized through primary or specialist care, and provided mainly to patients with suspected COVID-19 infection before being admitted. The majority of the included 27 studies may have met our definition of RPM, as they used self-monitoring on paper/online forms or wearable sensors. However, methodological quality of these studies was poor, and outcomes could not be summarized; we therefore lack knowledge of the impact of monitoring strategies. In another recent rapid review of barriers to using “remote monitoring technologies” with COVID-19 patients, equity-related barriers were the most common. These included financial barriers and lack of access to technology, membership in a patient group with particular needs (including chronic diseases), and low health literacy. Wilhite et al.’s recent survey of health care providers reports much of the same barriers to remote monitoring during COVID-19, particularly patients’ lack of access based on socioeconomic status, and patients’ and providers’ low technology skills. To summarize, we do not yet know the most effective way to remotely manage COVID-19 patients, and current strategies have not become magically more equitable than pre-pandemic remote monitoring. Pre-pandemic lessons about health technology inequality must inform current pandemic responses.

Strengths and weaknesses

A strength of this overview is the definition of RPM developed in collaboration with the commissioner/policy maker, and an exhaustive search strategy that allowed us to capture interventions that were not called RPM but nevertheless met our definition. While the specificity of the definition required screening of approximately 3,500 RCTs included in 155 systematic reviews, a time-consuming practice atypical of an overview of systematic reviews, it has also ensured that the interventions summarized in this overview are relevant to Directorate of Health. Even working within this specific definition, the interventions involved a variety of actors, data transmission methods, data evaluation methods, and response options. Each of these can be used as possible design options in moving forward with national RPM recommendations. We were analytically limited by the methodological choices of both the included systematic reviews and RCTs, such as two studies’ failure to report standard deviations, and two studies reporting outcomes adjusted for baseline values while the rest reported unadjusted values. As we did not calculate summary effect estimates, these situations are not particularly problematic, but is worth keeping in mind when viewing forest plots.
Knowledge gaps

It is unsurprising that the impaired vision/hearing systematic review was an empty review, as this was an exclusion criteria for many of our identified RCTs – despite the fact that impaired vision and hearing are conditions that will only increase with age, along with other chronic diseases; see for example Fisher et al. [49]. While physical multi-morbidities were reported by many RCTs, most excluded based on psychiatric morbidity. Excluding patients on the basis of psychiatric morbidities is, unfortunately, standard practice in clinical trials; investigators often assume that these potential patients will struggle more with treatment adherence than other patients. Assuming that patients with mental health problems will not comply may be selling them short [50, 51]. Excluding patients who use any type of psychotropic medicine, or those with anxiety, depression, alcohol, and/or substance problems, belies the prevalence of these conditions among people with diabetes and/or hypertension [52-55]. There is both potential and need to tailor RPM to people with impaired vision/hearing and with comorbid mental health problems, particularly if they have other conditions that limit mobility and utilization of in-person health services. RPM technologies should be developed following basic universal design principles to be suitable for people with disabilities, and we encourage the inclusion of people with disabilities and comorbid mental health problems in future trials of RPM for chronic physical conditions.

Conclusion

In this comprehensive overview of four systematic reviews, we aimed to assess the effectiveness of a specific type of remote patient monitoring (RPM) on clinical and healthcare utilization outcomes for chronic disease patients. We found a slight benefit of RPM on HbA1c, and a small negative effect of RPM on one type of health-related quality of life, which have both been reported in previous reviews utilizing different or broader definitions of RPM. These somewhat disappointing results may be because RPM facilitates data transmission, analysis, and feedback, but does not necessarily assist patients in making or sustaining the medication, diet, or physical activity change that are often necessary for these conditions. RPM could be seen as a bridge to necessary further support, but not superior by itself to usual care. More complex RPM interventions may be required to support such complicated behavioral change, such as interventions combining multiple

Figure 4 A proposed research and clinical agenda for remote patient monitoring

- Include people with impaired vision/hearing in future trials
- Include people with psychiatric diagnoses in future trials
- Develop RPM technologies to follow universal design principles, to be accessible for people with disabilities
- Co-design RPM technologies with target patient groups
- Explore and address equity-related barriers to utilization, such as to socioeconomic status, location, age, race/ethnicity, and technology literacy
components, or perhaps involving specialists from the beginning. There is a clear need to capitalize on RPM’s innovative capacity to serve people with hearing, vision, psychiatric, and cognitive difficulties.

Extensive experience has been gained during the COVID-19 pandemic about RPM and other types of remote care. While COVID-19-specific RPM outcomes have not yet been synthesized [2], we hope that best practices and lessons learned during the pandemic will be carried forward to provide high-quality RPM for post-covid patients as well as other chronic disease groups. More complex interventions are on the horizon [56], and the COVID-19 pandemic has catapulted forward the use of machine learning and artificial intelligence [57, 58]. Enhanced RPM strategies can collect and analyze massive amounts of real-time data, genomic information, and other risk factors, and they have the potential to increase accuracy and speed of clinical decision-making and follow-up. Policymakers must keep in mind equity is as important as efficacy. The pandemic has revealed nothing if not the unacceptable global and national disparities in healthcare access and health outcomes. RPM has the potential to extend care to people who fall through the cracks of traditional services – as well as the potential to further privilege those with financial and technological resources and literacy. Involving target groups and underserved subgroups in RPM design and implementation may be the key to seeing significant benefits for people with chronic disease.

Author contributions

AEM was the project leader, screened abstracts, assessed full-texts, conducted quality assessment, extracted data and analyzed, participated in GRADE assessments, and drafted the manuscript. SSO assessed full-texts, checked data extraction, conducted quality assessment, and participated in GRADE assessments. TJB and PSJJ assessed full-texts, conducted quality assessments, and participated in GRADE assessments. RCB was the principal investigator, drafted the study protocol, screened abstracts, and resolved conflicts relating to study inclusion. All authors contributed to the manuscript and read and approved the final version.

Author disclosures

AEM: No conflicts of interest.

TJB: No conflicts of interest.

PSJJ No conflicts of interest.
SSO: No conflicts of interest.
RCB: No conflicts of interest.

Data availability
Data is available by the authors upon request.

Funding statement
The white paper published on our institution’s website was commissioned by the Norwegian Directorate of Health.

References
15. Veritas Health Innovation, Covidence systematic review software [computer program]. Melbourne, Australia.


