Technical Brief Disposition of Comments Report

Research Review Title: Automated-Entry Patient-Generated Health Data for Chronic Conditions: The Evidence on Health Outcomes

Draft report available for public comment from June 26, 2020 to July 31, 2020

Research Review Citation: Treadwell JR, Reston JT, Rouse B, Fontanarosa J, Patel N, Mull NK. Automated-Entry Patient-Generated Health Data for Chronic Conditions: The Evidence on Health Outcomes. Technical Brief No. 38 (Prepared by the ECRI-Penn Evidence-based Practice Center under Contract No. 290-2015-00005-I.) AHRQ Publication No. 21-EHC012. Rockville, MD: Agency for Healthcare Research and Quality. March 2021. Posted final reports are located on the Effective Health Care Program search page. DOI: 10.23970/AHRQEPCTB38.

Comments to Draft Report

The Effective Health Care (EHC) Program encourages the public to participate in the development of its research projects. Each draft report is posted to the EHC Program Web site or AHRQ Web site for public comment for a 3- to 4-week period. Comments can be submitted via the Web site, mail, or email. At the conclusion of the public comment period, authors use the commentators' comments to revise the draft report.

Comments on draft reports and the authors' responses to the comments are posted for public viewing on the Web site approximately 3 months after the final report is published. Comments are not edited for spelling, grammar, or other content errors. Each comment is listed with the name and affiliation of the commentator, if this information is provided. Commentators are not required to provide their names or affiliations in order to submit suggestions or comments.

This document includes the responses by the authors of the report to comments that were submitted for this draft report. The responses to comments in this disposition report are those of the authors, who are responsible for its contents, and do not necessarily represent the views of the Agency for Healthcare Research and Quality.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #2	Introduction	In addition to disparities in access to the devices themselves (page 1, line 50), patients in rural and underserved communities may not have sufficient internet access or other resources to utilize the device as directed or in the manner studied in a RCT.	We noted in the Discussion the issues about lack of access to the internet or smart phones as one of the reasons rural populations are underrepresented in existing studies. We have added a similar comment to the Background.
Peer Reviewer #2	Introduction	Clinicians and healthcare systems may not be equipped to process or store these data in a secure manner. A further issue is that there are no billing codes for reviewing PGHD data.	We agree and we have added this point in the background section.
Peer Reviewer #2	Introduction	Guiding Questions are reasonable and appropriate to the report. Notably, though population-specific data is mentioned in the guiding questions it is not reported for each chronic disease. Rather it is summarized at the end of the report. These data should be presented for each section to provide essential context for study findings.	For each clinical condition, we have added several details to the main body of the report, concerning the characteristics of the enrolled patients: age, gender, baseline disease severity, rural populations, and technological access/expertise.

Commentator & Affiliation	Section	Comment	Response
Key Informant #1	Introduction	This technical brief raises the important issue of widespread marketing of automated PGHD technologies which make claims that are not validated by the FDA. These technologies lie in the grey area between regulated devices and recreational technologies. In light of the fact that FDA regulation would hinder innovation in this area enormously, the AHRQ can perform an important service by evaluating the effectiveness and safety of these technologies- albeit based on a wide variety of research studies with varying quality.	Thank you for your comment.
Key Informant #1	Introduction	The background discussion did not address one issue which, if there is actual literature about this, it would be helpful to include. What are consumer's perceptions about these technologies? Although it seems self-evident- to what degree do consumers expect these measurements to be valid? How much money has been invested in them on a national market level? Is there reason for concern that a significant amount of consumer energy and time is wasted in engaging in these technologies?	We summarized a recent review (Vo, 2019) to address your comment. Thank you.

Peer Reviewer #4	Introduction	The background of the study is well described. I am, however, surprised that two important issues are not addressed in the study. 1. The first is the issue of data security and privacy. Many of the available apps collect patient data and some send these data to healthcare professionals. Many providers do not disclose what they do with the data collected; where is it stored, what do they use it for, how is privacy secured, etc. This is a major concern. I understand the current review was focused on effectiveness, but, given the importance of the topic, I would expect data security and privacy would have been included or at least mentioned in the report. 2. The second issue is the theoretical foundation (or lack thereof) of the available interventions. Self-management and self-monitoring as an important element of it, is aimed at supporting patients to change their behavior in a positive health promoting way. What behavior is the most important depends on the disease, the condition and the habits of the patient. Changing behavior is a very complicated task and in the literature many theoretical models have been described that help to explain the relevant factors and guide the development of interventions. A problem in the field of study this report deals with is that most apps lack a solid theoretical foundation. This leads to an important question that can be asked for many apps: why on earth would this have an effect? A simple analogy: when we drive a car we can continuously see our speed, but this does not prevent most of us from driving too fast; only in combination with other measures like fines for speeding, clear traffic signs, warning systems, and - as a basis - the requirement of having a driving license based on practical skills and theoretical knowledge about the traffic rules etc., there is a chance that most people	We agree and have added info about data and security concerns in the Background section. We have also added in the background section, the connection of how PGHD from both the patient and clinical perspective can improve health outcomes.

Commentator & Affiliation	Section	Comment	Response
		don't drive too fast. When applied to the topic of the current report I think it is important to understand the role of self-monitoring as part of a complexity of factors and other elements that affect health behavior. It is hard to imagine how self-monitoring alone would lead to positive health outcomes; only when patient have knowledge, are stimulated/supported, have a certain level of motivation, self-monitoring has any chance of being effective. That is why I am not surprised that the results of the review show such weak or no evidence of effectiveness. It is in my view a weakness of the current study that this theoretical background has not been looked at. The authors state that research is needed to study the isolated effect of self-monitoring. I would question that and say that research is needed to study how self-monitoring can best be embedded in effective behavior change interventions aimed at supporting self-management of chronic diseases.	
Peer Reviewer #4	Introduction	The guiding questions are relevant and important. As mentioned above I missed a question about data security and privacy, as well as a question about the theoretical background of the interventions included, but apparently the experts interviewed at the start of the study have not mentioned these issues.	Thank you for your comment.
Key Informant #2	Introduction	The background provides a good summary of the rationale for technology growth to improve management of chronic conditions. It might be helpful to give a sense of the consumer marketplace for mHealth devices that capture data.	We have added in the background the current type of technologies that exist and forecast numbers.

Commentator & Affiliation	Section	Comment	Response
Key Informant #2	Introduction	2. Page 1, 2nd paragraph: The reviewer suggests clarifying up front definitions - consumer health technologies vs. PGHD technologies. Which consumer health technologies are devoid of data collection from the patient? Perhaps provide the definition of consumer health technologies and clarify that this technical brief focuses only on the subset that captures PGHD.	We have added the PGHD definition of the Office of the National Coordinator for Health Information Technology to this section. We also included our definition of "consumer" devices and we now clarify that we only focus on the subset that captures automated-entry PGHD.
Key Informant #2	Introduction	3. While discussion exists regarding FDA approval no citations are provided. It might be worth reviewing and including some mention of the FDA's effort to review digital health technologies. https://www.fda.gov/medical-devices/digital-health/digital-health-software-precertification-pre-cert-program	We have added citations to this discussion. We have also added discussion of FDA's efforts to review digital health technologies on page 1 of the Background section.
Key Informant #2	Introduction	4. Page 1 line 17: consider providing a definition of PGHD.	We have added the PGHD definition of the Office of the National Coordinator for Health Information Technology to this section. We also included our definition of "consumer" devices and we now clarify that we only focus on the subset that captures automated-entry PGHD.

Commentator & Affiliation	Section	Comment	Response
Key Informant #2	Introduction	5. Page 1 line 26: the first reference of PGHD refers to it in the singular (is) but data in this context would be plural.	We reworded this to "The marketplace for PGHD has grown rapidly…"
Key Informant #2	Introduction	6. Page 1 line 26 references PGHD as a rapidly growing field. This reviewer is certain PGHD is a field rather digital health tech development is a marketplace. There certainly is interest in studying how PGHD advances healthcare but that would be more in the space of informatics and how people interface with data and technology.	We agree and have changed "marketplace" to field.
Key Informant #2	Introduction	7. Page 1 line 52: The authors note the concerns around accuracy of data capture which is an important point. Another consideration is the lack of science/evidence behind how to present data in a manner that directs/informs decision-making. For example, if you show a directional trend it doesn't mean that meaningful change has occurred (even if the visual change looks impressive). Thus a person might believe that improvement is occurring when in fact it is not.	We agree that a directional trend does not imply meaningful change. This is generally why we focused on health outcomes. And when we needed to look at surrogates, we gave heavy consideration to the minimal important difference.
Key Informant #2	Introduction	Overall the guiding questions seem adequate.	Thank you for your comment.
Key Informant #3	Introduction	The guiding questions were good and on point. However, not all parts of the guiding questions were addressed in the findings (see comment for findings).	See our response to your pertinent comment below
Key Informant #3	Introduction	The background was brief, but adequately described the clinical problem and the need for the brief.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Key Informant #4	Introduction	Some of the guiding questions conflate outcome assessments and therapeutic interventions, such as question #1. As operationalized throughout the report, some of the PGHD were used as an intervention (e.g.. WW app or the guided weight loss in the diabetes example) and as a tracker of health outcomes. This overlap makes interpretation of the response to question #1 challenging.	Guiding Question #1 is intended to orient the reader to numerous facets of the evidence, and the other Guiding Questions delve into more specifics. For example, Guiding Question #3 involves effectiveness.
Key Informant #4	Introduction	It might be helpful to reference FDA general wellness guidance to clarify what is considered a medical device that would have undergone FDA review from those that do not or are classified as a medical device under enforcement discretion.	Thank you. We have added this reference to the Background section (page 1).
Peer Reviewer #2	Methods	Page 3 (lines 41-42) – the decision to restrict studies to consumer purchasable devices is understandable but may not capture emerging technologies such as Google watch, which are increasingly used for research purposes may provide important data for this review.	We did not limit the scope to wearables. Regarding the Google Watch, this is an interesting point, but we did not identify any studies that evaluated this device. However, this may be a good target for future research studies in this area.
Peer Reviewer #2	Methods	I highly support the inclusion of surrogate outcomes (e.g., health trend data) in this review.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #2	Methods	In general, the definitions for health vs “non-health” outcomes is a concern across all sections of the report and likely accounts for the low percentage of studies reporting “outcomes” for each chronic disease (e.g., pg 34, line 11 “only 21 of the 50 HTN studies report on health outcomes..”. For example, on page 3 line 44 – I am not sure I would define BP and A1c as non-health outcomes. In some instances, those are important health outcomes. Primary prevention interventions for cardiovascular disease may use those markers as objective clinical outcomes whereas interventions for secondary prevention tend to focus on preventing future hospitalization, morbidity and mortality in those with established disease. For clinicians in primary care vs specialty care settings, this distinction would provide clinically useful information for some of the chronic health conditions included in this report (diabetes, hypertension, etc).	We included HbA1c as a surrogate outcome in studies of diabetes prevention, and BP as a surrogate outcome in studies of hypertension. They are surrogate markers for health outcomes, but are not health outcomes on their own.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #2	Methods	Further, sub-categorization of objective (hospitalization) vs subjective health outcomes (depression scores, qol) should be considered.	We disagree with the suggestion to delineate objective and subjective outcomes. The report already has two categorizations of outcomes (surrogate/health, usability/effectiveness/harm, and different types of categories within those categories) and we do not think the objective/subjective distinction would add value. Please note that when we do discuss health outcomes, we provide the exact outcome (such as hospitalization or QOL) and so having a new category of objective/subjective seems unnecessary.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #2	Methods	It may be worth categorizing outcomes as objective health outcomes and subjective health outcomes for primary prevention and secondary prevention groups, respectively.	We disagree with the suggestion to delineate objective and subjective outcomes. The report already has two categorizations of outcomes (surrogate/health, usability/effectiveness/harm, and different types of categories within those categories) and we do not think the objective/subjective distinction would add value.
Key Informant #1	Methods	The investigation did not significantly alter the guiding questions and the presentation of each chronic illness followed by these guiding questions simplified the presentation of results. Some assumptions about how health outcomes were defined, and the concept of "isolation of the technology's effect" need further justification, but are within the scope of the guiding questions as they are presented. Specific examples of where such justification is needed will be addressed in the review of the Methods section.	Thank you for your comment.
Key Informant #1	Methods	The Methods section is well described and the methods are highly rigorous. They are presented in a way that this review could be replicated easily in five years using the same methods to yield comparative results. Some of the inclusion and exclusion criteria require further elaboration.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Key Informant #1	Methods	1) What do the authors mean when they exclude technologies that rely on manual input? If an asthma application requests users complete a subjective breathlessness measure or if quality of life is assessed by user self-report – are these not included? If so, why? How else would these be measured? In short, it is not clear what this exclusion criterion might be referring to, and some examples may help the reader to understand this better. Given that quality of life, was a health outcome, one may assume they are allowing for consumer completion of measures but it is not clear.	If a study used a PGHD intervention that met inclusion criteria, we included any relevant outcomes reported in the study, regardless of whether the outcomes were captured automatically. We have now added clarifying text to the Methods section of the report.

Commentator & Affiliation	Section	Comment	Response
Key Informant #1	Methods	2) The outcomes were tiered to an extent with mortality, survival, Emergency Room visits, admissions etc etc designated as health outcomes and illness related variables as surrogate outcomes. Please justify this further as surrogate outcomes may be the specific outcome consumers are trying to target. In short, consumers would not view these as "surrogate". This was most evident in hypertension where blood pressure is listed as a surrogate outcome. Moreover, the incidence of some of the health outcomes is likely low for the target duration for many of these trials so it sets a high expectation of effectiveness on the primary outcomes. It seems that the health outcomes were transdiagnostic and could apply to all, while the surrogate outcomes were specific to the disease. Perhaps this classification would be more accurate. This is really just an issue of nomenclature, but it does raise the possibility that someone who briefly skims this review, concludes that consumer health technologies did not overwhelmingly impact health outcomes as defined in this review, yet they did appear to have some impact on what consumers health goals such as systolic blood pressure in hypertension.	We have clarified in the methods section our distinction between surrogate outcomes and health outcomes.
Key Informant #1	Methods	3) When the inclusion/exclusion states that any comparator is acceptable- does this include no intervention comparator? If so, that should be stated.	Yes it does; we edited the PICOTS table to accurately reflect our requirements about comparisons.

Commentator & Affiliation	Section	Comment	Response
Key Informant #1	Methods	The risk of bias evaluation strengthens the rigor of the analysis and is well defined. Is there any concern about financial interest in some of the studies conducted and, if so, was that considered as part of bias risk?	We did not explicitly include financial interest as an item for our risk of bias assessment. We suspect that any risk of bias issues arising from this would have already been captured by the risk of bias items that we did use.
Key Informant #1	Methods	The report lists several discussions with Key Informants and the resulting discussions. I agree with the report investigators choices in response to each of these issues such as not excluding telehealth as a search term.	Thank you for your comment.
Peer Reviewer #4	Methods	The methodology of the study is very strong. The study has been executed with great rigor and thoroughness. The data are very well presented and discussed. This work is impressive in terms of amount and quality.	Thank you for your comment.
Key Informant #2	Methods	1. Engagement with Key Informants is described including how feedback shaped the review.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Key Informant #2	Methods	2. It might be helpful in the context of PICOTS to define harms. Were these harms related to clinical outcomes only? It appears unintended consequences or burden to patients (e.g., distress from seeing no improvement) was not captured. This may simply not be part of literature covered and may be valuable for considering in future work (especially as it relates to attrition of use).	The literature did not report these types of indirect harms. We have now noted in the PICOTS table that we included any harms that studies reported to be related to PGHD interventions. We have included a more general statement on the limited assessment of harms systematically in the current evidence base, and we now specifically mention the possibility of patient distress due to no improvement
Key Informant #2	Methods	3. For the assessment of risk of bias, please define the composite for categorizing into low, moderate, high risk. Was it a yes/no for each with all weighted equally or did some parameters hold heavier weight when determining bias?	We have added information clarifying our risk of bias assessment process.
Key Informant #2	Methods	4. For the evaluation of Economic Evaluations (page 19) how was risk of bias determined (or what parameters were assessed in making this judgement as a modified version of the CHEC is referenced.	The specific items appear in appendix tables C-27, C-34, and C-60, and the main text now notes this.

Commentator & Affiliation	Section	Comment	Response
Key Informant #2	Methods	5. The process (and outcomes) of device similarity is not clear. In particular, how is the reader meant to interpret this information (or why is device similarity important for decision-makers?) What criteria was used to determine similarity? These details seem important for thinking about accuracy from model to model, how data is submitted, transmitted, displayed, etc. but as defined, it is not clear.	We added clarification for why readers might be interested in device similarity. For example, if a device is very dissimilar to any products in the marketplace, then the trial results are less relevant to decision makers. This is because, even if the trial showed a device benefit, the decision maker could utilize that benefit by buying a similar device on the marketplace. The methods section clarifies the criteria for judging device similarity.
Key Informant #2	Methods	6. As noted earlier, the goals and methods of the usability testing for the WW app is not clear. Why was this application selected? What were the methods and processes used for this?	We have added text to the methods section about why we conducted this usability analysis, how we chose that particular app, and how it fits in to the overall picture.

Commentator & Affiliation	Section	Comment	Response
Key Informant #2	Methods	7. Similar to the point above, please provide additional rationale and details on the methods for conducting the device evaluations conducted by the engineers. Who were the engineers? How many participated? Why were the particular devices selected (or others excluded?) What protocol was followed? If not listed in methods please consider as an appendix to the report.	We have added additional information regarding the protocols used in the sections describing the device evaluations.
Key Informant #3	Methods	Page 7 - It's not clear why the usability assessment of the Weight Watchers (WW) app was included in this technical brief. I believe it takes away from the true focus of the brief. The authors did not explain the need to include it nor the reason for only focusing on the Weight Watchers app. There are other weight loss programs available too. Why were they not included? It seems more like a product promotion for Weight Watchers. I recommend removing this assessment from the technical brief.	We have added text to the methods section about why we conducted this usability analysis, how we chose that particular app, and how it fits in to the overall picture.
Key Informant #3	Methods	Page 5 -- Was there a beginning date for the search strategy? (lines 5--8)	Each search was from inception of the databases, and we have now added this to the methods section
Key Informant #4	Methods	In grading the risk of bias, the descriptors of low, high or unclear were listed for each item and then they were combined into low moderate or high risk of bias. It is not clear what criteria was provided to guide how these labels were applied. it would be helpful to reference an appendix or guide that informed the classification	We have added information clarifying our risk of bias assessment process.

Commentator & Affiliation	Section	Comment	Response
Key Informant #4	Methods	Moreover, on page 6 the terms as well as the usability evaluation, it would be useful to understand what influence the grading of "moderate" or "low"	We have added information clarifying our risk of bias assessment process.
Peer Reviewer #1	Results	1. It is not clear why the usability assessment of the Weight Watchers app was conducted. It seems to be a bit out of place with the rest of the report since it is the only app that was discussed.	We have added text to the methods section about why we conducted this usability analysis, how we chose that particular app, and how it fits in to the overall picture. We added that our usability analyses was a pilot study of how additional data might complement the standard research literature, and we have cited a relevant paper by Bates et al. (https://jamanetwork.com/journals/jama/article-abstract/2707668)
Peer Reviewer #1	Results	e. Findings: 1. One thing that I found a bit confusing was that there is a mention in the appendix to a study being conducted for “Mobile Physical Activity for Type 1 Diabetes” (page C-1). This doesn't seem to match any of the 11 conditions stated.	This was an error on our part; we have deleted the study

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	Results	2. Table C-3 in the appendix has conditions listed including “motivational interviewing” (page C-3), “obese” (page C-3), “Lung transplantation” (page C-3) that seem out of place, alongside the other conditions.	Those terms had been previously included in the clinicaltrials.gov table of trials that actually are not relevant to this report, so we deleted them
Peer Reviewer #1	Results	3. Page 30 line 17. The grammar needs correcting here. “Of the four diabetes-prevention-related records in clinicaltrials.gov, none made PGHD-related comparisons and stated that they were collecting data on a health outcomes.”	That count has been edited and the sentence removed, due to some of the previously listed studies not being relevant to this report
Peer Reviewer #1	Results	4. Page 29 line 3. This paragraph seems to be related to “preventing diabetes” but it’s in the “sleep apnea” section.	We corrected this error; thank you
Peer Reviewer #1	Results	5. Page 35 line 21. There are a number of instances where the paragraphs/spacing seems off. This is an example.	We corrected this error; thank you
Peer Reviewer #1	Results	6. page 83, line 7. This bullet points and paragraph/spacing seem to be off here.	We corrected this error; thank you
Peer Reviewer #2	Results	page 11, line 49 – there may be a typo concerning the number of studies not yet recruiting.	These numbers have been updated, thank you
Peer Reviewer #2	Results	In the section on PGHD for the prevention of diabetes, were there data on secondary prevention interventions? The studies reported in the technical brief appear to be focused on primary prevention – if data are available concerning prevention in those with pre-diabetes or improved management of those diagnosed with diabetes, those data would enhance the utility of this report.	This report only involved primary prevention; we have now clarified this in the PICOTS table

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #2	Results	Pg.35 line 29 – rather than a “lack of user support”, consider using the phrase “inadequate technical support.”	We made this change; thank you.
Peer Reviewer #2	Results	Page 38 (line 18), the definition of BP control changes in 2017 to 130/80 mm Hg rather than 140/90 – based on new ACC and American Heart Association (AHA) guidelines for the detection, prevention, management and treatment of high blood pressure. It would be worth while to note if any current/completed studies reported data based on the new guidelines for BP management.	All studies which examined isolated effects and reported on BP control were completed before the publication of the 2017 guidelines. We added more information on how BP control was defined across these studies.
Peer Reviewer #2	Results	In the section on heart failure outcomes (page 53), stating the NYHA class along with the outcome would enhance interpretation of these data as a patient with NYHA class 1 HF differs significantly from an individual with class 4 HF. Additionally, if there are data on women, HFpEF vs. HFrEF (preserved vs reduced ejection fraction), and young people that would be important to report.	We have now added information on NYHA class to a new section called "Additional Patient Characteristics". Although not stated in all studies, the four studies that reported NYHA class enrolled a majority of patients in NYHA Class II or III.
Peer Reviewer #2	Results	In the section on cardiac arrhythmias (page 55), were data from current/completed wrist-worn device trials excluded?	We excluded studies that did not meet our stated inclusion criteria, otherwise any relevant study on wrist-worn devices would have been included.

Commentator & Affiliation	Section	Comment	Response
Key Informant #1	Results	The findings are presented well, supported by specific data extracted from research studies and organized according to the Guiding Questions. As stated earlier, navigation of the document would be much easier if all subheadings included the chronic disease it was referring to.	Thank you for your comment.
Key Informant #1	Results	Each of the chronic disease categories included a thorough description of the intervention characteristics including interoperability, usability, and fidelity/validity. Overall, these reviews support the categorization for impact on isolated health outcome presented on Page x.	Thank you for your comment.
Key Informant #1	Results	The Evidence Map is sound for each of the conditions presented and the writing is clear except where noted below.	Thank you for your comment.
Key Informant #1	Results	The inclusion of the adherence data is extremely important as most studies report remarkably low continuous usage for the outcomes they report. This issue of adherence should be highlighted more in the discussion as it may be one of the greatest barriers to any impact of these devices.	We have added specific content to the summary section noting these issues of generalizability given the high variability of adherence in the evidence we reviewed. We now note adherence as a ripe area for future research in the “Next Steps” section.

Commentator & Affiliation	Section	Comment	Response
Key Informant #1	Results	The investigations insistence on identifying the specific effect of the application in multicomponent trials is rigorous and helpful. However, their specification of additional interventions may be overly stringent. This makes sense if one wants to tease out the addition of a coach or other contact intervention. However, for Obesity, the investigators do not include the addition of physical activity goals (e.g. 10,000 steps a day) in obesity trials yet these types of goals are often part of the technology interface. Moreover, one of the overarching aims of wearable technologies is improving patient self-management of chronic illness, and motivation enhancers (such as goal setting) are inherent in the overall intervention. At the very least, this needs more justification because it appears to be throwing the baby out with the bathwater. Motivational components are a key value added for these technologies.	We felt that activity goals is an intervention that could live on its own, therefore we did count it as an additional component of a multicomponent intervention

Commentator & Affiliation	Section	Comment	Response
Key Informant #1	Results	The review of the PGHD for hypertension makes an enormous contribution to the literature. However, as mentioned earlier, it is hard to understand why blood pressure is a surrogate health outcome rather than an actual health outcome. Also, for the discussion of reviews by device engineers, it seems all technologies fall short on validity criteria that FDA approved devices are required to meet. However, all but one were rated as safe. These technical and engineering experts probably believe there is a range of inaccuracy that is allowable, however, it if it falls below FDA or CDC criteria for validity- does that not pose a safety risk? The definition of how safety was determined needs more explanation so that such discrepancies are better understood.	We have clarified in the methods section our distinction between surrogate outcomes and health outcomes. We have added more information regarding the methodology the device engineers used to evaluate the accuracy and safety of the BP monitors. Note that these monitors are considered consumer-marketed devices and not intended for use in medical applications, and therefore are not required to meet the same standards as FDA-reviewed medical devices.
Key Informant #1	Results	There are many interesting findings reported throughout the review such as subgroup differences favoring positive impact in minorities for hypertension technologies, and for pediatric patients using technologies addressing asthma.	Thank you for your comment.
Key Informant #1	Results	Finally, all of the visual representations of the review results are excellent and ease the readers task of integrating a complex list of information. Specifically, Figure 6 on page 82 summarizes a large, complex review of this vast literature.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #4	Results	The findings are well presented. As discussed under background above, the results are not surprising. It would have been more informative if the authors had also discussed and presented the theoretical background of the interventions (as far as available; many studies do not report this) included.	We have added in the background more information on the potential benefits of PGHD for both the patient and providers.
Public Reviewer #1 (René Quashie, Consumer Technology Association)	Results	A. Despite positive findings, the report assigns “Possible positive effect(s)” to only four therapeutic areas. Such an assessment undervalues the positive impacts observed in the studies and noted in the report itself.	As our Methods section clarifies, we used a systematic process for deciding to use the word "possible" when describing a "positive effect". Note that the revision does not use surrogate outcomes in an attempt to make statements about health outcomes, but instead it addresses those two outcome categories separately

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1 (René Quashie, Consumer Technology Association)	Results	1. In hypertension, only 7 of the 30 studies found that diastolic blood pressure (DBP) differences in studies of isolated effects of devices presence/absence found “favors no PGHD.” In addition, of the 23 studies finding “Favors PGHD”, 11 found they did so more than 2mm Hg—while no studies had a finding of not favoring PGHD by 2 mm Hg or more. Figures 5 and 6 graphically illustrate the bunching of the studies to the “left” (positive PGHD) side of the chart and table, respectively.	While many of the treatment differences favored PGHD over usual care for SBP and DBP in the studies that isolated the effects, for both outcomes, multiple mean effects were not clinically important, favored neither treatment, or favored usual care. Therefore, there is too much inconsistency to categorize it as a "likely positive effects.” Instead we deemed it a “possible positive effect”.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1 (René Quashie, Consumer Technology Association)	Results	2. For coronary artery disease, all results were in favor of PGHD regarding mortality, yet the draft report only assigned the “Possible positive effect” label. (see Table 7).	Because those results were from a single study we did not know if they are reproducible, which is why we chose "possible positive effect." In our update search we identified one new study that found no difference in mortality between intervention and control groups, but we are still leaving the judgment as "possible positive effect" because there are positive trends across several outcomes.
Public Reviewer #1 (René Quashie, Consumer Technology Association)	Results	3. In the heart failure analysis, all the moderate risk of bias studies found in favor of PGHD in terms of quality of life--and of the other risk studies, most favored PGHD (Table 8).	True, but for studies that isolated the effect of PGHD only one moderate-quality study found a difference in quality of life, so we feel that the "possible positive effect" rating is valid.
Public Reviewer #1 (René Quashie, Consumer Technology Association)	Results	4. For asthma, most studies favored PGHD (Table 12).	There was one moderate risk-of-bias asthma study included in the report, so we a "possible positive" rating is valid.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1 (René Quashie, Consumer Technology Association)	Results	B. For the subcategories classified for all 11 chronic conditions, only 2 had instances of unfavorable PGHD findings, while for the others “unclear” and “Likely no effect” overall assessments, were either neutral or favorable PGHD findings.	For the 11 conditions, for health outcomes, we stated "possible positive effect" for 3, and "unclear" for the other 8. There were no situations where were suggested that PGHD effects are harmful. We did have 1 case where we judged possible positive effect for surrogate outcomes, 1 case where we judged likely positive effect on surrogate outcomes, and 1 case where we judged likely no effect for surrogates.
Public Reviewer #1 (René Quashie, Consumer Technology Association)	Results	1. Obesity (likely no effect)—most studies in all the measurements were favorable regarding PGHD. Table 3.	We stated for obesity that there is "likely no effect" on surrogate based on data shown in Figure 2 and Figure 3, that indicate that all of the point estimates were less than the minimal important difference. Thus the evidence has shown, rather consistently, no strong effect on either BMI or weight.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1 (René Quashie, Consumer Technology Association)	Results	2. Diabetes prevention—all results for the surrogate outcomes portion were positive regarding PGHD. Table 4.	Only one of the three included studies used an isolated-effect design. It did not report any health outcomes. This trial did find a benefit of PGHD on a surrogate outcome, but the trial was at high risk of bias. Taken together, we believe the evidence on diabetes prevention is unclear.
Public Reviewer #1 (René Quashie, Consumer Technology Association)	Results	3. Arrhythmia—one negative PGHD designation—offset by favorable results in the surrogate outcomes category. Table 9.	Yes, because the negative finding in the health outcome conflicted with the positive finding for the surrogate outcome we chose the "unclear effect" rating.
Public Reviewer #1 (René Quashie, Consumer Technology Association)	Results	4. Stroke—most results favorable to PGHD in health outcomes, for 2 of the 3 risk categories. Table 10.	The stroke trial did not use an isolated-effect design, making it impossible to tell whether observed differences were due to the PGHD device or to other interventions that patients received. We did include its data in the appendix so that readers can view that data.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1 (René Quashie, Consumer Technology Association)	Results	5. Parkinson's—one slight negative in quality of life in health outcomes. Table 11.	There were no included studies of Parkinson's disease. Table 11 involved COPD.
Public Reviewer #1 (René Quashie, Consumer Technology Association)	Results	2. Of the 111 unique studies, 93 (84.5%) were in the areas of obesity and hypertension— leaving only 18 (16.2%) unique studies in the other 9 chronic conditions combined. As discussed above, even in the two chronic conditions having a relatively high number of studies, AHRQ appears to be conservative in assessing the positive effects of PGHD.	A large number of studies does not imply that there are positive effects.
Public Reviewer #1 (René Quashie, Consumer Technology Association)	Results	3. The presence of studies from foreign countries complicates matters because other countries have different models of care, health delivery and reimbursement systems, making it difficult to accurately compare results without understanding the influence of these other factors. For obesity, 16 of the 23 studies were conducted overseas, and for hypertension, 26 of the 28 were conducted in foreign countries.	We agree that across-country differences in health systems complicates interpretation of these data. The evidence tables included each study's country, and a comparison among countries may be of interest in the future, but it was outside the scope of the current project.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1 (René Quashie, Consumer Technology Association)	Results	A. In addition, there were 116 unique devices, of which 26 were labeled as “unknown” in terms of whether they were “similar to current devices.” It does not appear that the report measured the health outcomes of RPM utilization based on the potential differing effectiveness of RPM devices.	Our inclusion criteria did allow for comparison of different PGHD devices, so any RCTS that compared different PGHD devices and met other inclusion criteria were included.
Public Reviewer #1 (René Quashie, Consumer Technology Association)	Results	4. Nothing in the report undercuts the value of automated PGHD for chronic disease monitoring or management. The clear benefit of automation is that it enables studies of the effectiveness of specific devices to be grounded in accurate real-world data. Additionally, consumer PGHD and all connected devices are creating vast repositories of data that are useful for continuing research. Given the issue of reproducibility in peer reviewed research it can be expected that the transparency of real world-data will provide healthcare benefits.3	Thank you for your comment.
Public Reviewer #1 (René Quashie, Consumer Technology Association)	Results	5. 46 of the 111 studies focused on obesity and diabetes prevention, which are notoriously difficult to manage. It is unreasonable to expect that the use of a pedometer or accelerometer alone—or even in conjunction with a scale—would reliably deliver improvements to a broad number of individuals.	We agree that these conditions are difficult to manage.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1 (René Quashie, Consumer Technology Association)	Results	6. Very few of the studies reported on health outcomes. This lack of crucial information is a common problem in medical device and therapeutic interventions where pre-market validation is typically based on markers—and not outcomes. Moreover, most medical products already on the market have never been subjected to long-term health outcomes reviews.	We agree that it's rare to measure health outcomes, and the primary purpose of this project was to document where they have been measured and shown to improve. We did extend consideration to surrogate outcomes when the evidence on health outcomes was unclear
Key Informant #2	Results	1. This reviewer appreciates the authors breaking the findings out by condition studied as it seems a logical way to organize the information.	Thank you for your comment.
Key Informant #2	Results	2. Page 11 line 49: 20206 is referenced after the date and I believe this is an error.	We corrected this error; thank you

Commentator & Affiliation	Section	Comment	Response
Key Informant #2	Results	3. Figure 1: Similar to the comments in methods, it isn't clear how this information is to be used. Given that the results are presented by condition and many of the devices are cross cutting, it would be interesting, perhaps, to see how often unique devices cross across different conditions.	Please see the revised table in the Overview section, which now shows the number of devices in each device category for each clinical condition. This allows the reader to see the preponderance of different device categories in some conditions but not others. For example, the hypertension trials used 41 different BP monitors and 4 different pedometers, whereas the obesity trials used 17 different pedometers and 5 different BP monitors.
Key Informant #2	Results	4. Page 12: adherence measurement is referenced. Is it possible to distinguish here (and for other sections) how adherence was defined? It would be interesting to see if there are set criteria for adherence to understand if there is a 'dose-response' that is considered.	Adherence measurement was specific to each clinical condition and how adherence was measured is included within Guiding Question 1 as a subsection.
Key Informant #2	Results	5. Page 13 line 6: minor detail that the abbreviation (AE) should follow adverse events.	We corrected this error; thank you

Commentator & Affiliation	Section	Comment	Response
Key Informant #2	Results	6. Page 13 line 23: Please clarify if patient usage is the same as adherence.	The two concepts are not identical, but clearly they are very related. "Adherence" is a research term that indicates whether patients did what trial investigators asked them to do. "Usage", in the context of this report, simply means how often patients used the device(s) they were given, independent of investigators' expectations. We focused on usage, which is the term we used here, since we felt it was more important.
Key Informant #2	Results	7. Page 13 results for accelerometers/pedometers references different ways of wearing the product – did the way in which the device was worn (i.e., wrist, clip-on) correlate with duration or longevity of use which may then influence the ability to achieve outcomes?	The project was not scoped to allow us to delve into these types of detailed questions

Commentator & Affiliation	Section	Comment	Response
Key Informant #2	Results	8. Page 14 Usability Testing of the WW Online app. As noted in the overall feedback, consider pulling this to a separate section.	We have added text to the methods section about why we conducted this usability analysis, how we chose that particular app, and how it fits in to the overall picture.
Key Informant #2	Results	9. Page 15 line 41: The first sentence of this paragraph is challenging to read and extract the intended message.	We agree that the sentence was trying to communicate too much; we now use bullet points for clarity.
Key Informant #2	Results	10. Figure 3: the image is not showing in the draft accessed.	We have fixed this error; thank you
Key Informant #2	Results	11. Page 21 (and relevant for other sections) please indicate if other harms (unintended consequences of tracking) noted by studies? For example, hyper-tracking that could lead to emotional distress?	We examined harms in guiding question #5 for each clinical condition
Key Informant #2	Results	12. Page 22 line 26 (Inclusion/Exclusion Criteria): Why is race called out in the US study but not in the other two? This gives the impression that the other 2 studies represented more diverse populations but isn't reflected anywhere. Please confirm/clarify.	We removed the race information from that sentence, because it was not germane to the point being made
Key Informant #2	Results	13. Page 30 line 5: the line references diabetes but should reference sleep apnea.	We have fixed this error; thank you

Commentator & Affiliation	Section	Comment	Response
Key Informant #2	Results	14. Page 34 (Technical report on wearable and handheld BP device): This reviewer suggests placing this report in a section separate from the results of the literature review (similar recommendation for the WW Usability test). As the methods are different than the literature review it seems it would be best placed separately. Further, especially for BP monitors that are utilized in other areas of study, the results are cross cutting across different studies.	Thank you for your comment.
Key Informant #2	Results	15. Page 35 line 6 the statement 'This consumer-marketed device has not undergone FDA review for its BP-measuring technology, and it does not have the same accuracy as BP monitors regulated by FDA.' is not cited. Is the from packaging or the conclusion of the authors?	This is an ECRI statement, not from the packaging. We have clarified this, and added citations to the reports produced by the device engineers.
Key Informant #2	Results	16. Page 35 line 13 results indicate phrases such as less accurate. How was this defined? Is there any quantified results from the assessment/evaluation carried out or thresholds use to draw conclusions?	We clarified at the beginning of the section how accuracy was defined.
Key Informant #2	Results	17. Page 35 line 26 indicates cost - What factored into cost of ownership and why over a 3-year period?	We clarified at the beginning of the section how cost was determined.
Key Informant #2	Results	18. Page 35 line 37 indicates less accuracy. Was this determined to be statistically less accurate (or to what extent?) Were measures consistent over time? There might be value in consistency as one can calibrated it against other devices.	We clarified at the beginning of the section how accuracy was defined.

Commentator & Affiliation	Section	Comment	Response
Key Informant #2	Results	19. Page 35 line 52: Is $70/this just for the device or does use of the app come with a cost? Same comment above as clarification for what factored into cost.	We clarified at the beginning of the section how cost was determined.
Key Informant #2	Results	20. Page 36 line 11: Details are provided for AAMI SP10 criteria. If this is the criteria for which the devices were evaluated again, please provide this detail in the methods (or perhaps earlier in the section). This level of detail about how devices are evaluated is very helpful for the reader.	We now discuss this criteria at the beginning of the section.
Key Informant #2	Results	21. Page 37 line 38: I believe 'social foundation' intended to be 'social functioning'.	Thank you. We have corrected this error.
Key Informant #2	Results	22. Page 53 line 41: I am confused by this statement, if the result was not statistically significant, then direction of effect should not be reported (is simply was not statistically significant under the parameters tested.	We have deleted the statement on direction of effect.
Key Informant #2	Results	23. Page 57 line 27: Per prior comments, consider placing the device evaluation outside of the literature review. In addition, please consider adding more detail about the methods followed for the evaluation including who was involved in the design, collection, and analysis to data.	We have added more information to the device evaluation sections, but we feel these sections still belong in the main text.

Commentator & Affiliation	Section	Comment	Response
Key Informant #2	Results	24. Page 57 line 40: Interoperability is noted as good due to being able to generate and send a PDF report. What is the criteria for good? It might be worth noting that while a PDF can be imported into an EHR, it is not easily found once there. So this form of interoperability is relatively basic (i.e. not much better than a scanned document) versus placing discrete data into a record such that it can be found easily in workflow as well as tracked and viewed over time. Defining interoperability and basis for rating scales is important.	We have now defined interoperability in the text. Being able to send information that can be added to EHR was considered the minimum criteria to get a Good rating for interoperability. Each clinical condition now has a paragraph about interoperability.
Key Informant #2	Results	25. Page 57 line 51: A 2-year time frame is used here but 3-year time frame is used in the evaluation of the BP monitors. Please indicate why this timeframe was selected.	The time span was based on the expected battery life of the Kardiaband (2 years).
Key Informant #2	Results	26. Page 58 line 6: 'Workflow was rated fair because user entered symptoms are not included in the report that users can email…' Please clarify if this is the definition of interoperability or is this a separate feature?	We have now clarified the definition of interoperability in the text
Key Informant #2	Results	27. Page 73 line 51: Confirm the use of 6MWD and 6MWT are the same (or used interchangeably) with intention - 6 minute walk distance/6 minute walk test.	6MWT and 6MWD are the same test. We changed the text to 6MWT in all cases for clarity. Thank you.
Key Informant #2	Results	28. Page 74 line 39: The acronym CCC is used – is this the clinical call center referenced later in the paragraph? Consider simply spelling out to reduce the use of acronyms since not frequently used other places in the report.	CCC is clinical call center. We made the correction. Thank you.
Key Informant #2	Results	29. Page 75 line 38: Minor note that the Euro symbol is needed in place of the 'E'.	We made the correction. Thank you.

Commentator & Affiliation	Section	Comment	Response
Key Informant #3	Results	The findings section was mostly easy to follow and the summary tables towards the end of the guiding questions were helpful. However, the authors did not really address all of the parts for the guiding questions. Specifically, guiding question #3 asks, in part, "Does this vary across different patient populations, different settings, or other modifiers of effectiveness?" However, the findings did not seem to address this. A clearly written sentence or paragraph on whether or not studies addressed patient population differences, for example, would have been helpful.	For each clinical condition, we have added several details to the main body of the report, concerning the characteristics of the enrolled patients: age, gender, baseline disease severity, rural populations, and technological access/expertise.
Key Informant #4	Results	It would be helpful to format the text to more easily see each study and their associated risk of bias. In addition, a subheader to show the engineering evaluation would be helpful to guide the reader through the manuscript.	Each study's risk of bias is shown in the appendix tables, along with the individual risk of bias items for each study. There is one table per clinical condition. The similarity evaluations by device engineers also appear in the appendix. Because these tables are so long, we did not feel it important to include them in the main report.

Commentator & Affiliation	Section	Comment	Response
Key Informant #4	Results	QOL was used throughout the manuscript. This term is very generic and encompasses more than health-related QoL. For measures that it is unclear what concept was actually measured, it may be better to say health-related questionnaire. In addition, if a specific questionnaire was listed it would be better to describe the scores based on the questionnaire and not say QoL.	We agree that the term is generic, and encompasses many questionnaires. We have included the details about each questionnaire in the tables.
Key Informant #4	Results	For the obesity response to guiding question #1, it seems as if the WW app usability work was out of place and seemed to be a minor reflection of the PGHD used for this condition. In the text it was acknowledged that the majority were scales and accelerometers. In addition, the WW app was not solely an outcome assessment since it was also an intervention. I would recommend removing it from the report and perhaps providing qualifying statement addressing why.	We have added text to the methods section about why we conducted this usability analysis, how we chose that particular app, and how it fits in to the overall picture.
Peer Reviewer #2	Discussion and Conclusions	Information reported concerning the effects of PGHD on health outcomes in diverse populations, particularly women, older and younger adults, rural populations and racial/ethnic minorities is summarized on page 82. Providing summary population data for each section/chronic disease would greatly enhance the interpretation of the findings included in this report.	For each clinical condition, we have added several details to the main body of the report, concerning the characteristics of the enrolled patients: age, gender, baseline disease severity, rural populations, and technological access/expertise.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #2	Discussion and Conclusions	The report briefly mentions critical issues concerning how patients and caregivers can share PGHD with clinical providers and challenges with integration into electronic health records (page 83, line 18). The potential value of device-generated PDFs is also mentioned in several areas of the report but I would caution the language used to characterize that as an advantage. For example, on page 36 (line 17), the report notes that PDFs can be sent to providers which characterized as a potential advantage of this technology. However, most patients do not have an email address for their providers (there may be a general clinic email address provided to patients but rarely is an individual provider's email provided) so that is not feasible in many cases. Patient portals (my chart) are increasingly used but again, they often do not allow patients to upload external PDFs to this system for cybersecurity purposes. Thus, patients would have to print individual PDFs and bring them to a future clinic visit or mail them to the clinic to be scanned in by clinic staff and reviewed by a provider who cannot submit a bill for that time. CPT codes are not currently available for the review of PGHD which is a major barrier to provider adoption. Page 57 (line 41) also refers to this as an advantage of Kardia mobile. If patients do have an email address for a provider, each PDF has to be sent in a separate email. In our cardiology clinic we have received 20 emails from a single patient in a single day. Thus, I would caution any references to this feature as an advantage given the current limitations.	We agree and we have added this point in the background section.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #2	Discussion and Conclusions	The potential for psychological harm should be added to this report. Anxiety is prevalent in chronic disease populations and excessive monitoring with a wearable device could lead to excess healthcare utilization and quality of life impairment. Our group recently had a manuscript accepted for publication documenting a link between health anxiety and smartwatches in patients with atrial fibrillation. Including a statement that recognizes the psychological impact of these technologies should be considered in this section.	We have included a more general statement on the limited assessment of harms systematically in the current evidence base
Key Informant #1	Discussion and Conclusions	The results are well reviewed in the summary and the discussion of participant representativeness a good one. Noting the lack of rural and women participants is a strength as rural patients may benefit more due to limited access to healthcare and women, in general, are underrepresented in clinical trials. The limitations of the review, in terms of an everchanging field of available technologies is a good one.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Key Informant #1	Discussion and Conclusions	Though no studies examining Apple products met criteria for inclusion in this study – some of the usability and engineer review of these products would have been informative. Perhaps that could be included in the Next Steps. This review emphasized the isolated effects of these technologies, but it may be more helpful to the field to examine multicomponent interventions, providing they do modular randomized controlled trial designs to tease apart specific effects. In addition, the problem of long-term adherence needs to be addressed.	We note that device engineers evaluated the AliveCor KardiaMobile and KardiaBand Smartphone-enabled ECG monitors; both use the Kardia app, which can be used on a smartphone or Apple Watch. We have also mentioned in Next Steps that RCTs are needed to evaluate the Apple Watch's Kardia apps for arrhythmia detection to determine if they have any impact on health outcomes.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #4	Discussion and Conclusions	The summary of the results in the report is adequate, but the discussion of the implication could be a bit more extensive and wider. The authors focus on effectiveness, which is understandable as that was the task, but the results raise several important issues that need to addressed to further develop this field as a part of the way we deal with chronic diseases in our healthcare systems. The report shows how fragmented this field is, with very little solid evidence and very weak foundation of intervention. A discussion about how to further 'regulate' this field, how to support patients and healthcare professionals in choosing what to use and what not, and how to develop more solid interventions with a strong theoretical basis, etc., would be an important addition.	A discussion of how to regulate this huge field is beyond the scope of this report as well as outside of the review team's expertise. Patients and healthcare providers will make their choices based on numerous factors, only one of which is the supporting evidence. Evidence reviews like this can help patients and physicians make decisions, and also perhaps motivate funders and manufacturers to develop better evidence. As to how to develop interventions with a stronger theoretical basis, this lies in the purview of device manufacturers.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #4	Discussion and Conclusions	See remarks under f. I think there is a great need to for regulation of this field, that currently has a bit of a 'wild west' character. Effectiveness is of course a key element in any regulation solution, but reliability/accuracy, data safety, privacy, and theoretical foundation should also be included as key criteria. Another important next step is the development of some kind of decision support system for patients and professionals to choose the most appropriate technology.	We agree with this comment and feel that the first step is to find technology that is effective and then to recommend clinical decision support for both the patient and clinician in leveraging the effective technology.
Key Informant #2	Discussion and Conclusions	1. The summary and implications touch on important aspects covered with Table 13 highlighting the primary findings succinctly.	Thank you for your comment.
Key Informant #2	Discussion and Conclusions	2. Of note, on page 82 of the report, the summary presents new findings regarding implications of underrepresentation of females and rural populations. These findings were not specifically called out in the different sections (although summarized in the tables). Of note, sex-based assessments should link to rationale for why this variable is important as is noted with rural and the relationship to access to care.	For each clinical condition, we have added several details to the main body of the report, concerning the characteristics of the enrolled patients: age, gender, baseline disease severity, rural populations, and technological access/expertise.

Commentator & Affiliation	Section	Comment	Response
Key Informant #2	Discussion and Conclusions	3. What are the implications from the end-user study of the WW app or device evaluations? This is missing from the summary and implications sections.	We have added text to the methods section about why we conducted this usability analysis, how we chose that particular app, and how it fits in to the overall picture.
Key Informant #2	Discussion and Conclusions	4. Page 83 line 7: as noted previously, it is unclear what a decision-maker is intended to do with device similarity.	We added clarification for why readers might be interested in device similarity. For example, if a device is very dissimilar to any products in the marketplace, then the trial results are less relevant to decision makers. The methods section clarifies the criteria for judging device similarity.
Key Informant #2	Discussion and Conclusions	5. There is a lack of a conceptual framework that could help organize future research and policy. This is particularly important as there should be a clear hypothesis as to why technology is incorporated and how tracking is expected to result in changed outcomes.	We have added in the background more information on the potential benefits of PGHD for both the patient and providers.

Commentator & Affiliation	Section	Comment	Response
Key Informant #2	Discussion and Conclusions	1. Are there any next steps based on the end-user feedback conducted with the WW app or the device evaluations conducted with the BP monitors?	For obesity, we could list some next steps based on the usability testing, but those would not nearly as important as the next step that we did list for obesity, specifically the need to measure health outcomes such as quality of life. Therefore, to ensure the prominence of that main point, we did not add next steps based on the usability testing. For the BP monitor device engineer report, we did not feel any clear next steps were important enough to mention. In the Methods, we added that our usability analyses was a pilot study of how additional data might complement the standard research literature, and we have cited a relevant paper by Bates et al. (https://jamanetwork.com/journals/jama/article-abstract/2707668)

Commentator & Affiliation	Section	Comment	Response
Key Informant #2	Discussion and Conclusions	2. The use of mHealth creates a socio-technical system. This includes the way humans/tech/data interface. How should this be explored and reported in research as the fidelity of interventions not just for patients but how clinicians use and act on data is important to consider in evaluating outcomes.	We agree with this comment and have included in the Background section how PGHD can benefit both patients and clinicians to improve a patient's health.
Key Informant #2	Discussion and Conclusions	3. Is there any future research warranted regarding intermittent use of PGHD devices? Fatigue from monitoring may occur in chronic or progressive conditions. For example, short term use for newly diagnosed conditions (or post hospitalization) may be more successful in helping people obtain control of symptoms than longer term sustained health. Many studies discussed showed change in use over time. Perhaps continuous use of technology isn't effective, what about other strategies?	This is a valid point but we would need to define short term and long term use of these technologies and this may vary according to chronic condition. We have added the concept of intermittent use to our Next Steps section, in the context of adherence.
Key Informant #2	Discussion and Conclusions	4. The summary noted under-representation of females and rural populations. What next steps are appropriate and does this differ by condition?	We have noted this as a limitation in both the summary and the Next Steps future research section.
Key Informant #3	Discussion and Conclusions	Page 82. It may be helpful to indicate how many studies focused on rural. The summary currently indicates "few studies included rural populations"(lines 41-44). Is it possible to give the exact number?	We have now noted that 9 studies enrolled rural populations.

Commentator & Affiliation	Section	Comment	Response
Key Informant #3	Discussion and Conclusions	Since rural populations have higher rates of many of the conditions that were part of this technical brief, I recommend adding a paragraph highlighting the need for future research that includes geographically isolated populations. Please see CDC's MMWR Rural Health Series, and specifically papers published in January and February of 2017 (https://www.cdc.gov/mmwr/rural_health_series.html).	We have added a paragraph to the end of the next steps section.
Key Informant #4	Discussion and Conclusions	f. Summary and Implications: It might be important to know the duration of the effect and whether it attenuates over time.	We have added additional information in the summary section to draw attention to this limitation of the current evidence base.
Key Informant #4	Discussion and Conclusions	Correlations between other measures of how a patient feels and functions and PGHD may be interesting and could support the measures.	This is an interesting point, and may be a good target for future research in this area

Commentator & Affiliation	Section	Comment	Response
Key Informant #4	Discussion and Conclusions	Accelerometers are mentioned a lot but how they may be analyzed or used in the trials varies quite a bit. Harmonizing on how the data from these technologies might measure an important concept in different patient populations would be important to acknowledge	For the most part, studies used accelerometers to measure steps per day, therefore functionally, they were the same as pedometers. In theory, however, you are correct that they can be used for additional purposes. But given the huge overlap with pedometers in the trials we examined, we did not think the theoretical difference was important enough to discuss.
Key Informant #4	Discussion and Conclusions	This section would benefit from some editorial improvements to ensure it connects the ideas and sounds like one voice.	We have edited that section to improve the consistency of voice
Peer Reviewer #1	General	2. It is not clear to me why the report focuses specifically on the 11 conditions. For instance, why focus on type 2 diabetes prevention and not type 2 diabetes?	A previous AHRQ report had focused on diabetes itself, so we chose to focus our efforts on diabetes prevention.
Peer Reviewer #1	General	3. There are a number of cases in the report where diabetes is referred to, but diabetes is not one of the conditions included in the report. Diabetes prevention is included, but not diabetes. For example, on page 16 line 33, “HbA1c for diabetes” is stated. Maybe changing these to state “diabetes prevention” would help.	We corrected this to diabetes prevention where necessary

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	General	4. I think it may be more clear to state that the “diabetes prevention” is “type 2 diabetes prevention”.	We have noted in the inclusion criteria that in this report, when we refer to "diabetes prevention", we mean type 2 diabetes prevention.
Peer Reviewer #1	General	c. Guiding Questions: 1. The issue stated in page 3/line 53, I see as of great importance. “Other applicability concerns raised by KIs involved the possibility that study participants have unrepresentatively high familiarity and comfort with technology (e.g., ready internet access or already have a cell phone) and are more likely to be from urban (not rural) areas.” I see it was discussed in the sections on Coronary Artery Disease and Cardiac Arrhythmias. But from what I can tell, I wasn’t able to find it in the other sections. Apparently there was substantial in-person support for participants using the devices, which seems like an important factor to discuss in implementing the findings of this research into real world settings. Besides the focus on rural populations, should other populations possibly be considered? For example, there may be reason to focus on individuals with family incomes below the poverty of their communities?	For each clinical condition, we have added several details to the main body of the report, concerning the characteristics of the enrolled patients: age, gender, baseline disease severity, rural populations, and technological access/expertise.
Peer Reviewer #2	General	Overall, the report is well-written, and the topic will be of great interest to researchers, providers, payers, patients and healthcare leaders. The authors should be commended for their important work in this area.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #2	General	Page 1 (line 19). Mentions mobile apps and wearables. There are a growing number of biosensors and implanted sensor technologies that also record ambulatory health trend data. For example, pacemakers and implanted cardiac devices measure health outcomes and health trend data described in this report. I am curious as to why the report was limited to only consumer wearables?	We did not limit the scope to mobile apps and wearables (for example, blood pressure monitors are not truly "wearable" as there are used only to measure BP at a given time not worn continuously). The scope of all PGHD was considered too large to be covered in a Technical Brief. Given the proliferation of consumer wearables, AHRQ and the key informants decided this was the area of greatest interest to focus on in this report. The EPC program invites nominations for new topics (https://effectivehealthcare.ahrq.gov/get-involved/suggest-topic) and if the topic of biosensors and implanted sensor technology is nominated , AHRQ will consider it for a new review.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #2	General	In addition to noting the potential annual cost of a device, it would be helpful to note whether such devices are typically covered by health insurance companies. Admittedly, that can vary between specific payers but even a general statement about typical payer coverage could provide additional context.	We have added what is known about insurance coverage on these technologies and why reimbursement in the future is warranted.
Peer Reviewer #3	General	This is a sound, well-done Technical Brief. I do not really have specific feedback to provide on the methodology, results, or conclusions. There is one thought for AHRQ to consider incorporating into its "disclaimer statement" at the beginning of the report - that AHRQ is simply providing an analysis of information in the literature and not either endorsing or recommending against a specific commercial product ... in order to minimize potential blowback/liabilities/legal concerns from the entities producing these devices. For example, AHRQ could write something to the effect of: "The information in this report is intended to help health care decision-makers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of healthcare services. This report is not intended to be an endorsement of or recommendation against use of a specific commercially available device or product. This report is also not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information (i.e., in the context of available resources and circumstances presented by individual patients)."	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Key Informant #1	General	This systematic and detailed review of validity and efficacy of automated-entry patient-generated health data (PGHD) for chronic conditions addresses an important need because the applications are not regulated by the Food and Drug Administration (FDA) yet they make many marketing claims that imply efficacy. Moreover, there are thousands of applications available, making it difficult for consumers to discern actual value. The investigators successfully identified, characterized, and summarized the evidence base for PGHD applications claiming benefits in 11 chronic diseases/clinical conditions: obesity, diabetes, prevention, sleep apnea, hypertension, coronary artery disease, heart failure, cardiac arrythmias or conduction abnormalities, stroke, Parkinson's disease, chronic obstructive pulmonary disorder, and asthma.	Thank you for your comment.
Key Informant #1	General	Though the reviews for all conditions are informative, the review of PGHD for hypertension in particular makes an enormous contribution to the literature. The studies were rigorously reviewed, which allows some trust in the possibly positive impact of these PGHD technologies. In contrast, little to no support found for effectiveness with obesity, one of the most common uses of these technologies, is surprising and also highly impactful.	Thank you for your comment.
Key Informant #1	General	The report makes excellent use of data visualization techniques to help the reader easily understand the results. These combined with the straightforward tables summarize well a highly complex review.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Key Informant #1	General	1) Although there are technologies that target specific diseases, many focus on specific components of diseases such as blood pressure, sleep quality, or activity level. It is not clear if the technologies they are reviewing state that they specifically address these conditions as a whole, or if they are focusing on specific self-management components of the illnesses.	The full intended effects of the technologies were not explicitly stated in the studies. In many cases, the intent is obvious, such as the monitoring of blood pressure to increase patients' awareness of factors affecting BP and encouragement to take medications.

Commentator & Affiliation	Section	Comment	Response
Key Informant #1	General	2) The investigators focus primarily on automated PGHD. This decision needs more justification/clarification as many technologies combine automated PGHD with patient reported outcomes collected within the application using easy to administer measures. Many chronic illnesses include patient reported outcomes in management and determination of efficacy. Are these aspects of the application be excluded in this review and why?	Technical Briefs have time and budget limitations that necessitate limitations on the scope of a topic as large as PGHD interventions for chronic diseases. AHRQ, ECRI and the Key Informants decided to limit to the subset of PGHD devices that would be of most interest to readers. These were devices that employed automated data collection but could have automated or manual transmission of collected data to the provider. However, we reported any patient-centered outcome that was reported in any study that met inclusion criteria, regardless of how it was collected. We have added this information to the Methods section.

Commentator & Affiliation	Section	Comment	Response
Key Informant #1	General	3) The designation of key and surrogate outcomes needs more justification. The term surrogate implies an indirect measure of disease, yet many of the surrogate outcomes were the primary target of the technology (eg blood pressure in hypertension).	We have clarified in the methods section our distinction between surrogate outcomes and health outcomes
Key Informant #1	General	4) One minor organizational issue: It would be of great help to the reader to include the target chronic disease in the subheadings for each section so that s/he does not have to find the beginning of each section and then search for the specific subsection when searching for specific information. (e.g. Isolated Effects on Health Outcomes: Diabetes). Though this may lead to a lot of repetition- it will allow for easier navigation between sections of the document.	We added these subheadings, to ease navigation
Key Informant #1	General	Finally, the additional components in this review, including the usability study of target applications including the Weight Watchers Applications, and the engineer evaluations of validity were invaluable and not only provide insight into the specific consumer health technologies- they illustrate the types of evaluations these applications need to undergo when FDA regulation will not be applied.	Thank you for your comment.
Key Informant #1	General	Minor Editorial Issue: Page 7 line 40-41 there appears to be a language usage issue. There is a paragraph on page 30 that refers to Diabetes- should this be in the section about Apnea?	We corrected this error; thank you

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #4	General	This report provides an extensive and very complete overview of all the research that has been published or is underway about the effects of using consumer devices for self-collecting health data in patients with chronic diseases. This is an extremely important report because there are many such devices with connected apps available, most of which have never been evaluated properly and some of which may be even problematic in terms of accuracy, safety, and data security. This situation may be described as a kind of 'wild west'; there is a clear need to regulate this field and, most importantly, to offer some guidance to patients and healthcare professionals regarding what to use and what not.	Thank you for your comment.
Peer Reviewer #4	General	The authors have done an impressive amount of work, with great rigor and thoroughness. The result is a report that captures almost certainly all relevant research in this field. This is a great achievement.	Thank you for your comment.
Public Reviewer #1 (René Quashie, Consumer Technology Association)	General	While AHRQ's report provides valuable information for discussion simply by citing 111 studies, CTA believes that it does not adequately differentiate or reflect the importance of PGHD and service models such as remote patient monitoring (RPM).	There are likely many additional issues about PGHD technologies that our scope did not encompass. This report may be a first effort at summarizing the field as pertains to direct health outcomes such as mortality and quality of life data, and future reviews may shed more light on the underlying issues.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1 (René Quashie, Consumer Technology Association)	General	1. The draft report identifies examples of positive results, with respect to coronary artery disease (CAD), heart failure (HF), hypertension and asthma, through RPM utilization but does not adequately explore them.	A full exploration of the mechanisms for positive results was outside the scope of this report.
Public Reviewer #1 (René Quashie, Consumer Technology Association)	General	7. AHRQ must better recognize PGHD automatically captured and transmitted by a medical device or signal acquisition system and transmitted without a need for the patient to manually input device readings or other data. The inclusion criteria excluded medical devices with automated medical device data transmission to healthcare staff, systems, etc. The automation of medical device data from devices to healthcare professionals is an important, if not fundamental element, to managing chronic conditions as evidenced by Current Procedural Terminology (CPT®) coding, and CMS coverage and payment of chronic care remote physiologic monitoring, treatment management services.4 * * * * *	We did not exclude devices that automatically transmit data to healthcare staff. We did require that data be captured automatically, rather than manually entered as in questionnaires.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #2 (Brian Scarpelli, Connected Health Initiative)	General	Notably, CHI is the topic nominator for this Technical Brief, which was originally proposed to address the use of remote patient monitoring (RPM) devices in addressing chronic conditions, which would serve as a much-needed resource for policymakers considering much-needed policy changes throughout the federal government. While the scope of this Technical Brief was later altered by AHRQ to limit the scope of technologies to be evaluated to be non-prescribed consumer wearables, this AHRQ Technical Brief is nonetheless poised to provide much-needed assistance to a range of policymakers seeking to understand the role of digital health tools that bring new patient-generated health data (PGHD) into the care continuum. Such technologies are essential elements of advanced healthcare systems. Policymakers seek to understand the efficacy of digital health products as well as their cost-effectiveness, particularly as they contemplate changes to the American healthcare system in light of ongoing crises, including the ongoing COVID-19 pandemic. While the draft Technical Brief recognizes many examples of the positive impact of automated-entry consumer devices that collect and transmit PGHD, the draft Technical Brief does not accurately portray the benefits of digital health devices in addressing chronic conditions, and, unless corrected, will misinform the policymakers this Technical Brief it is intended to inform	We have taken steps to note limitations of the scope of our report and how future research may fill in gaps in the research. We have utilized objective evidence synthesis methods to draw conclusions about the current state of evidence given the scope of the report. Also note that the scope was changed from RPM through discussions with the nominator because of the findings from a previous AHRQ technical brief that there are already many systematic reviews on RPM which finds that RPM produces positive outcomes, and a new technical brief on RPM would have been duplicative. https://effectivehealthcare.ahrq.gov/products/telehealth/technical-brief

Public Reviewer #2 (Brian Scarpelli, Connected Health Initiative)	General	While the studies included in the draft Technical Brief clearly demonstrate the value of automated-entry consumer devices that collect and transmit PGHD in addressing further chronic conditions, the Brief fails to acknowledge such tools' role in the positive outcomes they produce. For example, the draft Technical Brief's conclusions for numerous chronic disease categories, including for obesity, diabetes, arrhythmia, stroke, and Parkinson's Disease, unfairly discount the strong evidence demonstrating favorable outcomes included in Tables 3, 4, 9, 10, and 11, and attached "unclear" or "likely no effect" labels to these chronic conditions. At a minimum, the Technical Brief should plainly share with readers (in both the abstract and the Technical Brief itself) the positive evidence in the studies used was positive, but that the Technical Brief's authors could not differentiate results of interventions between components of the PGHD system, which does not reflect negatively on the automated-entry consumer devices efficacy. In this sense, the Technical Brief does not appear to be fully responsive to the scoping questions put forward by AHRQ – for example, the integration of PGHD into the electronic health record or the use of data to inform treatment/care plans is raised in Question 2, but is not considered in the Technical Brief.	We have highlighted the potential effect of PGHD devices on both health outcomes and surrogate outcomes. As standard for synthesizing studies, we have analyzed risk of bias that may lead to uncertainty in the conclusion that can be drawn from the evidence base. We did identify and include studies that detailed changes in treatment plans as a result of PGHD device data often in a protocolled fashion and assessed biases for these studies in an objective manner. Risk of bias was at times high and prohibitively high in studies that could not isolate the effect of the PGHD intervention and otherwise would lead to potentially misleading conclusions on the effectiveness of such interventions. The revised report contains, for each clinical condition, a section on

Commentator & Affiliation	Section	Comment	Response
			interoperability, which encompasses the notion of getting data into the EHR.
Public Reviewer #2 (Brian Scarpelli, Connected Health Initiative)	General	As drafted, the Technical Brief's scope is not sufficiently clear. The scope of the Technical Brief is for non-prescribed consumer wearables, yet the Purpose does not clearly state this scope. Further, the Technical Brief should note that its use of symptoms management in place of outcomes is stricter than the standards used by the U.S. Food and Drug Administration (FDA) to evaluate regulated automated-entry consumer devices that collect and transmit PGHD. Further, the Technical Brief should clearly explain that its evaluation of chronic conditions chose to exclude evaluation criteria the FDA relies on, including surrogate outcomes" or "non-health outcomes." This context is critical for policymakers or any other stakeholders reading the Technical Brief's conclusions.	The scope included not only wearables, but also mobile apps and any automated-data collection devices. We have now clarified this in Table 1 in the Methods section. We also included surrogate outcomes, as already described in the Methods. The FDA uses a fairly complex set of standards to evaluate devices, and mimicking those standards is not in AHRQ's purview.
Public Reviewer #2 (Brian Scarpelli, Connected Health Initiative)	General	The value of automated-entry consumer devices that collect and transmit PGHD in addressing chronic conditions is, like many other interventions, more difficult to manage in the short-term. The Technical Brief should acknowledge this shortcoming due to its timeline and structure.	We agree that short-term follow-up is a major limitation of the PGHD literature. We have added text to the first paragraph under Next Steps, to emphasize that RCTs with long-term follow-up are needed.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #2 (Brian Scarpelli, Connected Health Initiative)	General	The Technical Brief's scope in evaluating interventions attempts to narrowly restrict consumer wearables by ignoring the larger effects that such wearables have as part of a PGHD system that is responsible for an intervention. "Automated-entry consumer devices that collect and transmit PGHD" are integral components of a singular interventions. This approach leads to flawed results that ignore the contributions of automated-entry consumer devices that collect and transmit PGHD in addressing chronic conditions. For example, the expectation that pedometers or accelerometers would individually impact a person's health, may be unrealistic. Pedometers or accelerometers as part of a PGHD system, that includes other modalities of communication coupled with a treatment regimen may help a person realize those goals.	We did not restrict the scope to wearables. We also did not exclude multicomponent interventions if they included a PGHD device that met inclusion criteria, we just noted that these studies usually did not evaluate the isolated effect of the PGHD device. Other modalities alone (i.e., without a device) may also be effective, and ideally more studies would be designed to measure the impact of the PGHD in the system vs. the other components of the system.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #2 (Brian Scarpelli, Connected Health Initiative)	General	The draft Technical Brief notes that there is an expectation that use of automated-entry consumer devices that collect and transmit PGHD are expected to have “possible positive effect(s)” for four of the 11 chronic conditions raised. Evidence summarized for these four chronic conditions (hypertension, coronary artery disease, heart failure, and asthma) overwhelmingly supports a strong recommendation of positive effects. However, the draft Technical Brief only proposes to attribute “possible” positive effects. The Technical Brief should be revised to accurately reflect the use of automated-entry consumer devices that collect and transmit PGHD for hypertension, coronary artery disease, heart failure, and asthma with a “positive effect” label.	As our Methods section clarifies, we used a systematic process for deciding to use the word "possible" when describing a "positive effect".

Public Reviewer #2 (Brian Scarpelli, Connected Health Initiative)	General	A key context is missing from the evidence review that led to the conclusion there is little to no effect for addressing automated-entry consumer devices that collect and transmit PGHD in addressing chronic conditions. Many other interventions show limited or no impact (e.g., counseling or beta blockers for hypertension). This missing context should be addressed in the Technical Brief.	We agree, many other interventions do indeed show limited or no impact but are still used for healthcare. Guidelines, decision makers, and policy makers attempt to bridge the evidence to action gap, but ultimately those issues and providing context for that are outside the scope of this report. There are exceptions to endorsing interventions that show little or no effect in clinical trials, however those issues again are outside of the scope of this report in which we do objectively do not endorse (nor not endorse) any particular intervention nor are we drafting recommendations for policy makers, patients, or healthcare professionals. The example of beta blockers for hypertension is an excellent one in which guideline developers have gone to great

Commentator & Affiliation	Section	Comment	Response
			lengths to recommend against the use of beta blockers as first line therapy for hypertension for the reasons you've highlighted as they have limited comparative effectiveness.
Public Reviewer #3 (Ben Rosner, University of California, San Francisco)	General	Regarding the results classification for isolated effects on health outcomes, there is the potential that PGHD devices can cause harm, and this does not seem to be captured in the 4 classification categories. By way of example, if a patient with congestive heart failure is using a scale to track weight, and the clinician is receiving those weight reports and either underprescribes or overprescribes additional diuretic based on these data, this could result in inappropriate hospitalization (or failure to hospitalize) for CHF exacerbation. Similarly, a sphygmomanometer with inaccurate results could lead a prescriber to inappropriately escalate or not escalate anti-hypertensives. Some of these outcomes may be difficult to track, while others such as hospitalization related to the primary condition might be easier. Nevertheless, a category for potential harm might be worth considering.	We reported any relevant health outcomes, but we have now added a category for potential harms related to PGHD interventions in Table 1.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #4 (Rob Havasy, Personal Connected Health Alliance)	General	Although there is great value in evaluating the state of evidence and impact of interventions, therapies, and services on outcomes, particularly, as noted by the draft technical brief, we are concerned about the Technical Brief, specifically about its focus, assumptions, and conclusions. PCHAlliance, through input from its members, provides substantive feedback on the Draft Technical Brief with the goal of ensuring the Final Technical Brief will provide meaningful information to consumers, health researchers, and clinicians and lead to care improvement.	The methods we used were explained in the protocol as well as the Methods section of the report. We feel the report does provide meaningful information, as we found areas where statements could, as well as other areas where more research is needed.
Public Reviewer #4 (Rob Havasy, Personal Connected Health Alliance)	General	The Draft Technical Brief notes that little underlying research isolates the effects of a connected device alone on health outcomes, and defines health outcomes as mortality, survival, ER visits, hospital admissions, disease severity, disease progression, and quality of life. While disease indicators, such as blood pressure, weight change, or blood sugar levels, seem to have been classified as surrogate outcomes.	Yes, disease indicators are surrogate outcomes. The primary purpose of the report was to examine health outcomes. We did include surrogate outcomes, and if the evidence on health outcomes was unclear, we then examined surrogate outcomes.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #4 (Rob Havasy, Personal Connected Health Alliance)	General	We believe this approach to the evidence review misunderstands the potential for connected devices to impact an individual's health. Connected digital technologies' most powerful effect is to help provide feedback to individuals and help guide their decision-making through myriad decisions every day which might impact their long-term health and/or clinician management of their condition. The power of a pedometer or blood pressure monitor is less about a short-term impact on obesity or blood pressure, but rather as one additional contextual (at times clinical) data point among many for healthcare providers that also helps a patient stay connected to their health and may influence them to make more beneficial decisions than harmful ones. It is also of note that for some of the chronic conditions included in this evidence review, there are no treatments or therapies, even those approved by FDA, that demonstrate long term health outcomes. Instead, FDA uses what these Draft Technical Brief calls "surrogate" or "non-health" outcomes (both terms seem to be used interchangeably in the structured abstract, evidence summary, and introduction) for device and therapy reviews.	We agree wholeheartedly that devices may help a patient stay connected to their health and may influence them to make more beneficial decisions than harmful ones. The question, then, is whether trials have actually shown improvements in health outcomes, ostensibly based on those beneficial decisions, or other factors.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #4 (Rob Havasy, Personal Connected Health Alliance)	General	The evidence review identified traditional health outcomes such as mortality and hospitalization as its impact focus. For two conditions, hypertension and obesity, the evidence review found few, if any, published studies demonstrating surrogate or non-health outcomes - reduced BMI when pedometers alone are used to treat obesity and change in mortality when blood pressure monitoring is used to lower blood pressure. In both cases, the evidence review fails to note that many other widely deployed and accepted interventions, from nutritional counseling for obesity to the use of beta blockers (FDA approved) to manage hypertension have the same limited effect on outcomes. We believe this highlights that the proper use of connected health technology is as part of a comprehensive system of patient support that brings value by making clinician's more efficient and effective and/OR through long term, gradual, support of consumers. To solely measure connected care technology as a means to displace other common treatments with something less expensive, and disconnected from health care, ignores the evidence of connected care technologies documented by prior Technical Briefs.	Whether non-PGHD interventions influence health was not in the scope of our report. Note that we did state, for hypertension, there is a possible positive effect on the surrogate outcome of blood pressure. We understand that PGHD device live in a complex system of interventions, and this underscores the difficulty of conducting trials that actually help one determine whether the PGHD device is a necessary component.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #4 (Rob Havasy, Personal Connected Health Alliance)	General	Connected digital technologies must be evaluated without the assumption that they should only be accepted if they show health outcome effects greater than the current standard treatments. For many chronic conditions small changes, sustained over time, yield better quality of life and management of symptoms. Further, in many chronic conditions, these technologies can improve existing standard treatment through efficiencies and/or improved quality. Digital interventions are uniquely positioned to help patients make better decisions and slowly adjust their behaviors.	Nowhere in the report did we assert that PGHD technologies should only be accepted if they show health outcome effects greater than the current standard treatments. We agree that these technologies have potential. This report documented several cases where they have shown possible positive effects on health outcomes.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #4 (Rob Havasy, Personal Connected Health Alliance)	General	For example: · A pedometer by itself produces only a number of steps, which, without context, can barely be considered useful information about a patient to inform anything. But monitored over time, activity monitoring can help any number of health coaches and providers guide patients towards evidence based interventions such as prescription medication and a more active lifestyle and find ways to improve their physical activity levels, which have many more benefits than merely reducing BMI. Paired with a smartphone app and cloud-based systems which can adjust their based on myriad factors such as a patient's location and even a local weather forecast, these systems can incent increases in activity, changes in eating, or modifications of prescriptions which can be sustained beyond any methods we currently have available. There is a developing evidence base that activity monitoring paired with software-based communications providing individually tailored recommendations and messages can improve outcomes.	We understand that pedometers simply measure steps. Our goal was to see if their use caused better health outcomes.
Public Reviewer #4 (Rob Havasy, Personal Connected Health Alliance)	General	· A blood pressure monitor by itself only produces a biophysical reading. But, when that information is communicated to a health care provider it enables accurate diagnosis, medication management, and referrals for behavioral counseling for diet and physical activity. Over time, this information can enable consumers to seek professional care for a 'silent' disease. Further coupled with a two-way communication system or remote monitoring, medication adherence is improved dramatically.	Again, we wanted to see if all of these intermediate changes resulted in better health outcomes compared to no self-monitoring

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #4 (Rob Havasy, Personal Connected Health Alliance)	General	· Weight scales by themselves produce simply a weight reading. But for consumers with heart failure, when it is coupled with daily reporting to a care management team, tailored patient education, and two-way communication (remote monitoring), hospitalization rates are dramatically reduced.	Again, we wanted to see if all of these intermediate changes resulted in better health outcomes compared to no self-monitoring
Public Reviewer #4 (Rob Havasy, Personal Connected Health Alliance)	General	We urge the authors include a clear explanation of the context and scope of this evidence review: The report's context is not clear in the structured abstract, evidence summary and introduction. Specifically, the very limited focus of the evidence review is unclear. We urge the authors to clearly state in the abstract, the evidence summary and the introduction that this review is limited to short term results on consumer use of physical devices that provide digital data to a consumer. And, communicate that the review excluded: ♦ use of the device when it was wrapped with a health care service(s); ♦ software or 'apps' only devices; ♦ long term impact as many of these digital devices are designed to deliver data to support slow changes by consumers in concert with many other factors that lead to impact over a long time.	We did not limit to short-term outcomes or to physical devices (we included apps). Our inclusion criteria notes that we set no restrictions regarding time points. Specifically, we said "No limitations on timing". We also did not exclude studies where the device was combined with other health care services. Such combination designs do make it difficult to determine the impact of the device, and we did separate them from other studies that focused on device impact.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #4 (Rob Havasy, Personal Connected Health Alliance)	General	· We urge the authors to note the real-world data contribution of patient-generated health data (PGHD) from all types of devices for research and evaluation: The report does not note the vast repositories of data being created by consumer PGHD and all connected devices that are useful for research and real-world data. We urge the authors to acknowledge the role and need for real-world data that consumer PGHD offer to evaluate health services, interventions, and the reproducibility problem in peer reviewed research (https://www.nature.com/news/reproducibility-a-tragedy-of-errors-1.19264).	Additional remarks have been included in next steps to highlight this limitation of the current evidence base.
Public Reviewer #4 (Rob Havasy, Personal Connected Health Alliance)	General	· We urge the authors to provide information in a manner that communicates clearly to consumers: Improving information and the evidence base available to consumers for evaluation of digital devices used to manage chronic conditions is vitally important. We are concerned that this draft technical brief may not provide information and analysis helpful to consumers, instead it presents evidence in a manner that could be understood by researchers, but also focused solely on the evidence about the contribution of a connected device absent clinician, software-based, or coaching based chronic disease counseling.	Many of the included trials did involve physicians and/or software and/or coaching, not just the devices themselves, and this information was included. Such trials do make it harder to determine what contribution the device makes to outcomes.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #4 (Rob Havasy, Personal Connected Health Alliance)	General	We urge the authors to note the overwhelming evidence that PGHD automated transfer to health care staff is associated with improved non-health outcomes for a large number of the conditions included in this review: The inclusion criteria excluded devices that automatically transmit data to health care staff, yet, it is also evaluating chronic condition management which is typically done in concert with health care advice, counsel, and care planning. It is vital that consumers understand that when automated data transfer of PGHD from devices to health care staff (remote monitoring of diabetes, heart failure, or patients with multiple chronic conditions) or to evidence based coaching (e.g. Weight Watchers online which was extensively discussed in the Technical Brief) is an evidence proven component to chronic condition management. And, some of the devices included in this study have shown efficacy as part of a chronic condition management provided by health care staff.	We agree that there may be evidence on the relationship of PGHD transmission and effectiveness of patient outcomes. This is however outside of the scope of our report. We've highlighted this in the Next Steps section as a limitation.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #4 (Rob Havasy, Personal Connected Health Alliance)	General	· We urge the authors to make clear that the health outcomes of focus in this evidence review are aspirational and is not the approach typically used for FDA review of therapies or devices used for these conditions. In many cases evidence on how to manage the underlying chronic condition to achieve the defined health outcomes is not clear or known. Some of these chronic conditions, like obesity and pre-diabetes, are complex multi-factor diseases with genetic, biologic, environmental and behavioral components. For many of these chronic conditions, symptoms management ("surrogate outcomes" or "non-health outcomes") is a proxy for outcomes. FDA review on devices and therapies to address these chronic conditions use the very "non-health outcomes" that are not the "overall focus" of this evidence review. Why would this technical brief set a higher standard for consumer PGHD devices than is set for FDA reviewed devices?	The primary purpose of the report was to assess the evidence on health outcomes, and whether those are considered aspirational, or whether they are required by FDA, are not our concern. Payers and clinicians/guideline groups look for the types of evidence that we sought in this review. We do understand that surrogate outcomes are a proxy for health outcomes, and probably this is why the term "surrogate" is used.
Key Informant #2	General	Thank you for the opportunity to review this very thorough technical brief on consumer devices that capture PGHD. The use of technology to improve quality and efficiency of person-oriented care is an area of interest for health care stakeholders yet much is unknown about the extent to which such technologies achieve this without causing undue harm or added costs (without value).	Thank you for your comment.
Key Informant #2	General	The strength of this report is organization by conditions and the guiding questions to help frame the report. Further, the detailed information and organization of the tables in the report and appendices is helpful to garner a sense of the studies conducted and high-level findings.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Key Informant #2	General	While specific recommendations for improvement follow organized by section of the report, this reviewer would offer a couple of overarching comments for consideration. It is this reviewer's understanding that a goal of the technical brief is to provide a potential framework for assessing the applications and implications of the technology under study. It would be interesting and helpful to see to what extent the published studies reported the underlining framework or conceptual models that guided the study. In particular, the rationale for studying the technology and how it might modify/improve the outcomes of interest.	Few trials used a framework for evaluation, and given time constraints, we did not extract per-study information on their frameworks. We did include in the revised background a general framework for PGHD devices.
Key Informant #2	General	Clearly defining interoperability up front (and perhaps in the summary/implications sections) would be helpful. There are different levels of interoperability that have direct implications for workflow. While I believe this wasn't a topic well addressed throughout the majority of studies, it is an important aspect to clarify when reviewing the ability to a technology to integrate with electronic health records. A basic level is the ability of a device to generate a PDF report that can be incorporated as a file into the EHR. While this may fit the criteria, it does not allow for easy access/recall within an EHR as compared to a more advanced integration where data is stored as discrete data points based on common data standards.	We agree about the importance of interoperability concerns with PGHD but have focused this document on the effectiveness. Note that the revision contains interoperability information, separately for each clinical condition.

Commentator & Affiliation	Section	Comment	Response
Key Informant #2	General	The rationale for the incorporation of the usability testing for the Weight Watchers app. The Weight Watchers app functions quite differently than a consumer tracking device, in particular with regards to the requirements of users to login and input data. This type of study seems more relevant to a review of healthcare apps to support health behavior and tracking. Similarly, the rationale and methods for the analyses on the two types of consumer PGHD devices conducted by device engineers is unclear. Clarifying the goals and intent for both of these studies, the methods, and implications for future work would strengthen their inclusions. If this content remains in the technical report this reviewer recommends either pulling the details into the appendices or separating from the main results into separate results sections following the review of published literature.	We have added text to the methods section about why we conducted this usability analysis, how we chose that particular app, and how it fits in to the overall picture.
Key Informant #3	General	Overall, this technical brief was well-written, easy to follow and informative. The main issue is the inclusion of information on the Weight Watchers app, which I also discuss in my comments for the methods section.	We have added text to the methods section about why we conducted this usability analysis, how we chose that particular app, and how it fits in to the overall picture.
Key Informant #4	General	Comprehensive review of technologies published in the literature. Given the increasing importance of PGHD in care and trial paradigms, this report is timely.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Key Informant #4	General	It might be helpful to acknowledge the importance of including demographically diverse populations. There was acknowledgment of the pediatric population but having data on how well these devices collect information on children, large adults, and those with different skin pigmentation might be important if they are assessing outcomes.	We have added a statement about the importance of demographic diversity at the end of the Next Steps section.