PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Patients' experiences regarding self-monitoring of the disease course: an observational pilot study in patients with Inflammatory Rheumatic Diseases at a rheumatology outpatient clinic in the Netherlands
AUTHORS	Renskers, Lisanne; Rongen-van Dartel, Sanne; Huis, Anita; van Riel, Piet

VERSION 1 – REVIEW

REVIEWER	Annette de Thurah Department of Rheumatology, Aarhus University Hospital, Denmark
REVIEW RETURNED	01-Sep-2019

GENERAL COMMENTS	It was interesting to read this article about Influencing Factors Regarding Self-monitoring in Patients with Inflammatory Rheumatic Diseases. The study has interesting aspects, and I fully agree that challenges in the future health care systems calls for new solutions, but I think the present study suffer from methodological flaws, and issues that needs clarification. 1. On p 5 it is stated that one of the aims is to evaluate the correlations between the PROMs and the Disease Activity Score 28 (DAS28). In the method and result section is is called 'congurence'. I find it very difficult to understand the way in which these analaysis are obtained. I would expect some sort of concurrent validation with calculations of e.g. Spermanns correlation coefficients. 2. When and how was data about DAS28 obtained. For obvious reasons this can not be done by self report. This needs clarification and further explanations. 3. Which cutt-offs are used to establish 'poor' and 'good'? 4. Further 'adherence to e-mail alerts' makes no sense to me when the assessment interval is based on the patients own preferences. The definition of adherence (WHO) is the extend to which the patients behavior matches the (describers) recommendation. Please explain. 5. Please indicate treatment (csDMARD/bDMARD) in Table 1 6. Please explain the reason for the (very frequent) self assessment (on weekly/monthly basis). What is the purpose? Even though these diseases fluctuates this is overdoing things. 7. Please discuss feasibility. Who should give feedback to patients at least 12 times a year (2/3 of the patients) ? Is it realistic 8. Please discuss clinical relevant changes/algorithms for alerts. Patients can indicate small changes from one self-assessment to an other, but do all these small changes need full attention?

	9. Please discuss the possibility that these frequent self assessments can also have a negative impact on patients (worries, dependence on HPRs, constant attention to the disease) 10. Please also discuss the potential commercial interests a company like Phizer could have in patient self-assessment of symptoms. You might consider including some of these references: Thurah A, et al: Tele-Health Followup Strategy for Tight Control of Disease Activity in Rheumatoid Arthritis: Results of a Randomized Controlled Trial. 2018 Knudsen L, et al: Experiences With Telehealth Followup in Patients With Rheumatoid Arthritis Knudsen. A Qualitative Interview Study .Arthritis Care Res (Hoboken). 2018 Schougaard LM, et al; AmbuFlex: tele-patient-reported outcomes (telePRO) as the basis for follow-up in chronic and malignant diseases. Qual Life Res. 2016 Calvert, M et al:. Maximising the impact of patient reported outcome assessment for patients and society. BMJ (Clinical research ed.), 2019 Greenhalgh, J et al. The applications of PROs in clinical practice: what are they, do they work, and why?. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation, 2009

REVIEWER	Gunnhild Berdal, PT, PhD, Post doc Diakonhjemmet Hospital, Dept. of Rheumatology, National Advisory Unit on Rehabilitation in Rheumatology, Oslo, Norway
REVIEW RETURNED	23-Sep-2019

GENERAL COMMENTS	Title of manuscript: Influencing factors regarding Self-monitoring in Patients with Inflammatory Rheumatic Diseases This is an interesting research article covering an important subject, as increased knowledge about this patient groups' self-monitoring of disease impact potentially may lead to enhanced patient engagement in disease self-management, better disease control and thus improved patient health, and perhaps also more cost-effective utilization of health care services. However, I have several comments, which I hope will improve the manuscript. 1. Is the research question or study objective clearly defined? The study aims are quite clearly defined in the abstract, as well as in the introduction. However, the second/third aim (p 5, line 109-110) requires additional information: a) Information about what type of PROMs (disease specific, standardised) and what these instruments measure (disease impact, disease activity etc.) is lacking and should be provided previously in the introduction. The concept of PROM is broad and manifold, and as it is used in the introduction and the aims, it is too broad, and need to be narrowed down and specified. When that is taken care of, the term "PROM" could for instance be replaced by "patient reported measures of disease impact" in the 3rd aim. b) The rationale for comparing these particular PROMs with the DAS28 should be explained in the Introduction.

	c) The DAS28 is mentioned for the first time in aim 3, assuming that the reader is familiar with the instrument. The purpose of using the DAS28 and what it specifically measures should be provided previously in the introduction. d) The aim need to be reformulated as “to assess”, to replace “were assessed”. e) Further, related to the study objective (page 5, line 106): Information about the reason for pilot-testing the online self-monitoring program should be provided. Does this pilot study precede a larger scientific study, and/or is the pilot intended as a feasibility study of the online self-monitoring program before implementation in clinical practice? I recommend that the relevant information provided in the discussion (line 381-83) is moved to the end of the introduction. 2. Is the abstract accurate, balanced and complete? The abstract is easy to read and provides an overview that roughly shows what the study is about. Unlike the introduction, the abstract provides information that disease-specific PROMS were used in the online patient self-monitoring program. However, the abstract is incomplete as the methods section lacks information about: a) Study design b) Recruitment procedure and geographical setting c) Qualitative research method used to analyze the qualitative data; data which allegedly were collected through focus group discussions and telephone interviews d) How adherence to remainder e-mails was measured (provided in the main manuscript) e) Statistical methods used to analyze the quantitative data. I recommend that this information is added. Further, the results section in the abstract does not balance with the methods section as: f) “Facilitators” is not addressed g) The “RAID” is introduced in results as “best congruent with DAS28”, and thus allegedly compared to several other PROMs, while no other PROMs are mentioned in the methods section of the abstract. h) The “mean adherence to PROM reminder e-mails” expressed as percentage is difficult to interpret (percentage of what?) Furthermore: i) The conclusion needs to address “patients’ experiences with online self-monitoring” which was stated as a study objective. I think that the sentence in line 56; “Patients were predominantly positive about the concept of self-monitoring”, does not cover that objective sufficiently. j) I recommend that the reason why “self-monitoring has the potential to contribute to a more efficient allocation of outpatient consultations” is addressed in light of this particular study’s results. 3. Is the study design appropriate to answer the research question? The study design is not described. This information should be provided. 4. Are the methods described sufficiently to allow for the study to be repeated? The methods descriptions are to some extent described sufficiently enough to allow for repetition. However;

	a) The purposive sampling technique (line 146-47) is insufficiently described. b) The method to determine “congruence between DAS28 scores and PROM values” (line 165-169) and what the authors mean by “congruence”, as well as what they mean by “good” and poor” congruence (line 167-68), is unclear. Did they relate “good” and “poor” to a cut-off value? c) Neither the PROMs nor DAS28 are defined or described in the manuscript. As a consequence, a reader does not know what is measured and compared, or whether the instruments are compatible. d) Furthermore, the rationale for letting the patients choose which and how many PROMs to complete, is not given. e) The iMonitor program could have been more sufficiently described and in accordance with the stated reference. f) The semi-structured topic guide (line 175-77) used to “identify facilitating factors and barriers regarding online or remote monitoring” is not provided. This reduces the transparency. g) The qualitative methods used to generate the themes and subcategories, and to code the qualitative data is poorly described (line 184) with no reference to relevant methodological literature. h) Further, a better organization of the methods section in the manuscript is required. Section headings need to be revised to fit the corresponding text (e.g. line 116 “Study participants” should be replaced by “Inclusion criteria”), and/or text content need to be revised to fit the headings. Unnecessary repetitions should be removed (e.g. line 160 and line 200-01). Information regarding recruitment rate (line 368-69) should be moved from the discussion to the results. 5. Are research ethics (e.g. participant consent, ethics approval) expressed appropriately? Ethics approval is confirmed (line 407-08). A participant consent form is reported utilized (line 155) and privacy concerns are reported taken care of (line133). It is stated that “Patients were able to withdraw from the program at any time point” (line 123), which may indicate that patients have been informed of the right to abstain from participation in the study or withdraw consent to participate at any time without reprisal, in accordance with the Helsinki declaration. If this interpretation is correct, the sentence could be revised accordingly. It is stated in line 404 that the study was funded by Pfizer. This company has (according to line 131) developed the self-monitoring program that is being tested in the study. Nevertheless, (in line 400-01) “No conflict of interest” is declared. Although providing this information allows for transparency, I think that some sort of additional research ethical elaboration on this financial bond is required. 6. Are the outcomes clearly defined? None of the Patient Reported Outcome Measures that are utilized in this study are described or defined in the manuscript. Only the full forms of their acronyms are provided. More details on the outcome measures are needed. 7. If statistics are used are the appropriate and described fully? The applied statistical analyses are descriptive and quite simple. a)The appropriateness of using counts and percentages to produce useful information on adherence to e-mail reminders, as long as the number of e-mail-reminders is not taken into account (line 226-28), is questionable. Or is it so that the number of e-mail reminders equals “the number of PROM-assessments that should

	have been completed according to the chosen PROM-frequency"? That is, was the number of reminder e-mails defined by the patients, by using iMonitor? In that case, to clarify, the text needs to be revised accordingly. b) The exclusion of patients who only completed PROM(s) once (line 164-65) may lead to loss of information on (inadequate) adherence. I therefore suggest these should be included in the analyses. c) The method used to measure "correlation" (line 40 and line 109) or "congruence" (line 165) between the PROMS and the DAS28 requires a fuller description. 8. Are the references up to date and appropriate? The references in the first section of the introduction are quite old. I think more recent publications exist on patient-centred approaches, self-management support, shared decision-making and patient involvement in chronic healthcare. Otherwise, the references seem appropriate. 9. Do the results address the research question or objective? a) The first study aim "to obtain experiences (facilitating factors and barriers) regarding online remote monitoring in patients with IRDs" is only partly addressed, because "facilitating factors" is not integrated into the results. Further, what is meant by "facilitating factors" is not defined in the manuscript. b) To the best of my understanding, the second aim, to assess "information about adherence to PROM reminder emails", is not addressed, referring to the way adherence is calculated (subtext Table 2, line 226-28 and 162-64). This may be solved by revision to a more precise language. c) The third study aim, to assess "correlations between the PROMS and the Disease Activity Score 28 (DAS28)" (line 109) is not met either, because statistical methods for calculating correlations have not been used. Further, in the results section, the term "correlations" is replaced by "congruence", without the meaning of this word being explained. Furthermore, in results, the "congruence" with DAS28 is presented only for only two PROMS (Table 3), while five PROMs have been completed by the participants according to Table 1. 10. Are they presented clearly? a) Table 2 and Table 3 are unclear and need revision. b) Table 4 also need revision; it is difficult to grasp the meaning of the themes and subcategories in relation to the table heading. Generally, the themes may need further condensation and more precise descriptions of what they refer to (meaningful content) in the context of "facilitators" and "barriers". c) I recommend replacing "Influencing factors" with "Patient experiences", alternatively just adding these words to the subheading on line 243. This may also apply to the manuscript title. e) Further, should not (properly timed) e-mail alerts be considered as facilitators? 11. Are the discussion and conclusions justified by the results? The discussion and the conclusions in the main manuscript are quite reader-friendly and largely justified by the results, particularly with respect to the elaborations on the patients' experiences with the self-monitoring program. 12. Are the study limitations discussed adequately? The authors adequately refer to the relatively small number of participants as a factor that limits the generalization of the results (line 67 and 369-70). Further, due to the recruitment procedure, the study population may have been prone to selection bias (line

	368). No other study limitations are addressed. Potential study limitations associated with the qualitative method used should be addressed. 13. Is the supplementary reporting complete (e.g. trial registration; funding details; CONSORT, STROBE or PRISMA checklist)? Funding details is (partly) reported (line 404). Trial registration: not applicable. The overall reporting is considered incomplete. The STROBE checklist is recommended for better reporting of the quantitative part of the study. The COREQ checklist (1) and the SRQR recommendations (2) is recommended to guide a better reporting of the qualitative part of the study. I would suggest the authors to read these checklists carefully. A few important aspects that may have influenced the qualitative findings should be reported. For instance, the relationship of the researchers to the participants; were they involved in treatment as well? During the qualitative data analysis, were the themes derived solely from the data, or were they also inspired by theory? Were the findings influenced by the topic guide (line 175)? Personal characteristics (occupation, experience, and training) and perspectives of the members of the research team that may have influenced the data analysis should be described. 14. To the best of your knowledge is the paper free from concerns over publication ethics (e.g. plagiarism, redundant publication, undeclared conflicts of interest)? There may be an undeclared conflict of interest related to Pfizer, as described above (point 5). 15. Is the standard of written English acceptable for publication? No. I generally recommend a thorough examination of the written language (proof reading) to clarify the message in this manuscript. 1. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. International Journal for Quality in Health Care. 2007;19(6):349-57. 2. O'brien CB, Harris BI, Beckman JT, Reed AD, Cook AD. Standards for Reporting Qualitative Research: A Synthesis of Recommendations. Academic Medicine. 2014;89(9):1245-51.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1 – Annette de Thurah

It was interesting to read this article about Influencing Factors Regarding Self-monitoring in Patients with Inflammatory Rheumatic Diseases. The study has interesting aspects, and I fully agree that challenges in the future health care systems callsfor new solutions, but I think the presentstudy suffer from methodological flaws, and issues that needs clarification.

1. On p 5 it is stated that one of the aims is to evaluate the correlations between the PROMs and the Disease Activity Score 28 (DAS28). In the method and resultsection isis called 'congurence'. I find it very difficult to understand the way in which these analaysis are obtained. I would expect some sort of concurrent validation with calculations of e.g. Spermanns correlation coefficients.

Response: We completely understand that using the terms ‘correlations’ and ‘congruence’ interchangeably may have caused confusion. We did not have enough data to accurately conduct Spearman correlation tests, so we subjectively determined the agreement between the PROM- and

DAS28 scores (see also point 3). Action: We changed ‘correlations/congruence’ into ‘agreement between DAS course and PROM scores’ consistently throughout the document.
2. When and how was data about DAS28 obtained. For obvious reasons this can not be done by self report. This needs clarification and further explanations.
Response: DAS28-scores were obtained during outpatient consultations, about 3 to 4 times a year. These scores are kept in the electronic medical files from the hospital. The DAS28-scores used in this study are certainly not based on self-reports. Action: We added thisinformation to the method section (page 8, line 196).
3. Which cutt-offs are used to establish 'poor' and 'good'?
Response: We subjectively determined agreement between the DAS28 course and PROM scores (see
point 1), so we used no real cut-off ‘score’. Instead, we used 2 options: ‘Good’: cases where the DAS28
course and the PROM scores showed the same direction (for example: DAS28 course increased and PROM scores as well). ‘Poor’: when the DAS28 course and PROM scores showed opposite directions
(DAS28 course increased and PROM scores decreased (or the other way around). Only PROM-values
within 14-day-window from DAS28 assessment were used.
2
Action: For clarification, we extended the methodology section with more information about this procedure (page 10, line 231).
4. Further 'adherence to e-mail alerts' makes no sense to me when the assessment interval is based on the patients’ own preferences. The definition of adherence (WHO) is the extend to which the patients behavior matches the (describers) recommendation. Please explain.
Response: We apologize for possible misunderstanding regarding the adherence. First, we would like to make clear that patients chose their own preferred interval before they started with the program. Forcing patients to complete PROMs at intervals set by researchers does not support our idea of self-management and might impede patients’ motivation. Second, we wanted to know if patients did really stick to their preferred interval. Action: We will use the term ‘adherence to the predefined PROM frequency’ throughout the document. We also gave an example of calculating adherence (page 9 line 223). Regarding the patients’ preferred intervals, we mentioned this more specifically in the method section (page 8 line 176). Next to this, we emphasized the role of the patient regarding self-management behavior more explicitly in the method and discussion (pages 8 and 23, line 175 and 450-
452).
5. Please indicate treatment (csDMARD/bDMARD) in Table 1
Response and action: We added csDMARD/bDMARD to table 1 (page 14).
6. Please explain the reason for the (very frequent) self-assessment (on weekly/monthly basis). What is the purpose? Even though these diseases fluctuates this is overdoing things.
Response: Weekly self-assessment might seem very frequent, but we would like to mention that this was preferred by patients. They told they felt supported in handling their disease on a weekly basis. We think that it is very important to comply with their wishes, since we want to stimulate self-management and keeping our patients satisfied. Next to this, we think that especially those patients with an unstable disease course might benefit from weekly assessment. When remission is reached (or stable disease course), other intervals (e.g. monthly or quarterly assessment) are more realistic. Action: we mentioned the role of patient self-management more explicitly in the discussion (see also point 4): page 23 line 450-454).
7. Please discuss feasibility. Who should give feedback to patients at least 12 times a year (2/3 of the patients) ? Is it realistic
Response: We do not think that giving feedback 12 times a year is very realistic, it will be very time-consuming for HCPs. Instead, we think that feedback should be provided during outpatient

consultations. During these consultations, the patient receivesfeedback about the disease course from
the previous period. Action: We have explained this more thoroughly in the method section (page 9 line 211) and in the discussion on (page 21 line 416).
8. Please discuss clinical relevant changes/algorithms for alerts. Patients can indicate small changes from one self-assessment to an other, but do all these small changes need full attention?
3
Response: No, we do not believe that all small changes need full attention. HCP were expected to give
feedback to the patients about a patient's disease course during outpatient consultations. Next to this, patients were explicitly asked to contact the outpatient clinic in case they were worried about their values. Action: Additional information has been added with regard to giving feedback by HCP and the message option and telephone (page 8 line 180-181).
9. Please discuss the possibility that these frequent self assessments can also have a negative impact on patients (worries, dependence on HPRs, constant attention to the disease)
Response: Ensuring the wellbeing from our patient was very important while conducting this pilot study. Patients had the option to contact the outpatient clinic at any time by email (less urgent) or by telephone (urgent). This was explicitly emphasized during instruction classes. Dependence on HCPs did
not seem to be an issue in our study, since patients explicitly mentioned that by using iMonitor they felt less dependent from HCP in handling their disease and felt to have more control. Lastly, we believe
that constant attention to the disease was not an issue, since patients were able to choose their own preferred interval and they were not forced or imposed by the researchers. Action: We elaborated on this in the method section (page, 8 line 176-181) and discussion (page 23 line 452).
10. Please also discuss the potential commercial interests a company like Pfizer could have in patient self-assessment of symptoms.
Response: We do not know if Pfizer had commercial interests. Pfizer was able to access patient data, but they were certainly not involved in the medical decisions and encounters. After the pilot study, we did not continue with their system. Instead, a non-commercial system was as monitoring system. Next to this, we think that self-monitoring has the potential to lead to more efficient healthcare utilization. Also, it might diminish outpatient consultations and might also contribute to patients who are able to self-manage their disease, leading to optimal use of medication. Action: We added the information described above in 'Conflicts of Interest' section (page, 25 line 502).
11. You might consider including some of these references:
- Thurah A, et al: Tele-Health Followup Strategy for Tight Control of Disease Activity in Rheumatoid Arthritis: Results of a Randomized Controlled Trial. 2018
- Knudsen L, et al: Experiences With Telehealth Followup in Patients With Rheumatoid Arthritis Knudsen. A Qualitative Interview Study .Arthritis Care Res (Hoboken). 2018
- Schougaard LM, et al; AmbuFlex: tele-patient-reported outcomes (telePRO) as the basis for follow-up in chronic and malignant diseases. Qual Life Res. 2016
- Calvert, M et al:. Maximising the impact of patient reported outcome assessment for patients and society. BMJ (Clinical research ed.), 2019
- Greenhalgh, J et al. The applications of PROs in clinical practice: what are they, do they work, and why?. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation, 2009
4
Response: thank you for these suggestions. Action: we used the studies by Knudsen and de Thurah et
al. in the discussion (pages 21 and 22, lines 413 and 427) as supportive literature and Calvert et al in the introduction (page 6 line 139).

Reviewer 2: Gunnhild Berdal
Thisis an interesting research article covering an importantsubject, asincreased knowledge about this patient groups' self-monitoring of disease impact potentially may lead to enhanced patient engagement in disease self-management, better disease control and thus improved patient health, and perhaps also more cost-effective utilization of health care services. However, I have several comments, which I hope will improve the manuscript.
1. Is the research question or study objective clearly defined?
The study aims are quite clearly defined in the abstract, as well as in the introduction. However, the second/third aim (p 5, line 109-110) requires additional information:
a. Information about what type of PROMs (disease specific, standardised) and what these instruments measure (disease impact, disease activity etc.) is lacking and should be provided previously in the introduction. The concept of PROM is broad and manifold, and as it is used in the introduction and the aims, it is too broad, and need to be narrowed down and specified. When that is taken care of, the term "PROM" could for instance be replaced by "patient reported measures of disease impact" in the 3rd aim
Response: Thank you for this valuable suggestion. We agree that this information is needed. Most PROMs were disease-specific, whereas HAQ is non-disease specific. Action: We added this information
about the PROMs earlier in the introduction (page 5 line 109-113). We used RAID and RADAI-5 as disease-specific PROMs to determine their agreement with the DAS28 course, and we changed the text into '..agreement between the disease course assessed with disease-specific PROMs (RAID and RADAI-5)..' (page 6 line 143-144).
b. The rationale for comparing these particular PROMs with the DAS28 should be explained in the Introduction
Response: We agree with that. The rationale is to determine a patient's ability to monitor his/her disease by comparing the agreement between the disease-specific PROMs measuring disease impact/activity (patient) and the objective measure for disease activity (DAS28, by HCPs). Action: We added this rationale more specifically in the introduction (page 6 line 125-126 and 143-144).
c. The DAS28 is mentioned for the first time in aim 3, assuming that the reader is familiar with the instrument. The purpose of using the DAS28 and what it specifically measures should be provided previously in the introduction
Response: We agree with this suggestion, since BMJ Open readers might not be familiar with DAS28. Action: We added this information in the introduction (page 5 line 116-118).
5
d. The aim need to be reformulated as "to assess", to replace "were assessed".
Response and action: We changed the aim into 'to asses' (page 6 line 143 & 145).
e. Further, related to the study objective (page 5, line 106): Information about the reason for pilot-testing the online self-monitoring program should be provided. Does this pilot study precede a larger scientific study, and/or is the pilot intended as a feasibility study of the online self-monitoring program before implementation in clinical practice? I recommend that the relevant information provided in the discussion (line 381-83) is moved to the end of the introduction.
Response: The program wasintended to test the feasibility of online remote monitoring in daily clinical practice. With our insights, a new study about self-management was conducted. Action: We totally agree with your suggestion and we moved the information to the introduction and added the information about the feasibility (page 6 line 140-142).
2. Is the abstract accurate, balanced and complete?
The abstract is easy to read and provides an overview that roughly shows what the study is about. Unlike the introduction, the abstract provides information that disease-specific PROMS were used in the online patient self-monitoring program. However, the abstract is incomplete as the methods section lacks information about:
a) Study design
b) Recruitment procedure and geographical setting

c) Qualitative research method used to analyze the qualitative data; data which allegedly were . collected through focus group discussions and telephone interviews
d) How adherence to remainder e-mails was measured (provided in the main manuscript)
e) Statistical methods used to analyze the quantitative data. I recommend that this information is added. Further, the resultssection in the abstract does not balance with the methodssection as:
f) “Facilitators” is not addressed
g) The “RAID” is introduced in results as “best congruent with DAS28”, and thus allegedly compared to several other PROMs, while no other PROMs are mentioned in the methods section of the abstract.
h) The “mean adherence to PROM reminder e-mails” expressed as percentage is difficult to interpret (percentage of what?). Furthermore:
i) The conclusion needs to address “patients’ experiences with online self-monitoring” which was stated as a study objective. I think that the sentence in line 56; “Patients were predominantly positive about the concept of self-monitoring”, does not cover that objective sufficiently.
j) I recommend that the reason why “self-monitoring has the potential to contribute to a more efficient allocation of outpatient consultations” is addressed in light of this particular study’s results.
Response: We agree that the abstractsection isincomplete atsome point and we changed it according to your suggestions. Actions: a). Observational study using qualitative and quantitative research methods b). We gave information about the geographical setting. c, d, and e). We gave more
6
information about the research methods used for quantitative and qualitative data analysis and how we measured adherence f). We understand that using the word ‘facilitators’ is confusing. We changed it into: patients’ experiences g). next to RAID, we also mentioned RADAI-5. h). We changed it into: Adherence to the predefined PROM frequency. We consistently changed it throughout the document and gave an example of calculating adherence (see method section). i). We agree with that and we gave examples of positive experiences j). Thank you for this recommendation. We changed it into: ‘Patients were able to and willing to self-monitor their disease, which could contribute to a more efficient allocation of outpatient consultation in the future’ (see abstract pages 2 and 3).
3. Is the study design appropriate to answer the research question?
The study design is not described. This information should be provided.
Response: We added ‘observational study’ in the title, abstract, and the methodology section.
4. Are the methods described sufficiently to allow for the study to be repeated?
The methods descriptions are to some extent described sufficiently enough to allow for repetition. However;
a) The purposive sampling technique (line 146-47) is insufficiently described.
Response and action: We elaborated on our purposive sampling technique (page 9 line 202-203), which
will hopefully be more specific now. See also ‘limitations’ in the discussion section (page 23 line 468).
b) The method to determine “congruence between DAS28 scores and PROM values” (line 165-169) and what the authors mean by “congruence”, as well as what they mean by “good” and poor” congruence (line 167-68), is unclear. Did they relate “good” and “poor” to a cut-off value?
Response and action: We apologize for the misunderstanding. Action: We provided additional information in the method section (page 10 line 231-235) and changed ‘congruence/correlations’ consistently into ‘agreement between the DAS28 course and PROM values. See also point 3 from reviewer 1 on page 1 and 2 in this document.
c) Neither the PROMs nor DAS28 are defined or described in the manuscript. As a consequence, a reader does not know what is measured and compared, or whether the instruments are compatible.

Response and action: We totally agree with this and we gave descriptions of these concepts in the introduction (page 5 line 109 and 116), and method section (page 8 line 183).
d) Furthermore, the rationale for letting the patients choose which and how many PROMs to complete, is not given.
Response: Letting patients choose contributes to the concept of self-management and patient involvement. We think that imposing rules on patients has a contrary effect on them, they might be
7
less motivated to comply. Action: we explained these reasons more specifically in the method section (pages 7-8 line 174-177) and discussion (page 23 line 450-453).
e) The iMonitor program could have been more sufficiently described and in accordance with the stated reference.
Response: Thank you for this suggestion. Action: We provided more information about the program, also in relation to patients' preferences and self-management (method, pages 7-8 lines 174-181), see page 3 point 8.
f) The semi-structured topic guide (line 175-77) used to "identify facilitating factors and barriers regarding online or remote monitoring" is not provided. This reduces the transparency.
Response and action: We included the topic guide as an appendix (page 31).
g) The qualitative methods used to generate the themes and subcategories, and to code the qualitative data is poorly described (line 184) with no reference to relevant methodological literature.
Response: We agree with this comment. Action: we added a paragraph 'Qualitative analysis' and we have described the methods for data analysis more specifically (page 11, line 253) and we used references.
h) Further, a better organization of the methods section in the manuscript is required. Section headings need to be revised to fit the corresponding text (e.g. line 116 "Study participants" should be replaced by "Inclusion criteria"), and/or text content need to be revised to fit the headings. Unnecessary repetitions should be removed (e.g. line 160 and line 200-01). Information regarding recruitment rate (line 368-69) should be moved from the discussion to the results.
Response and action: We changed the method section as follows: We used 'Inclusion criteria' (instead
of Study participants). Moreover, we added a paragraph 'Patient Reported Outcome Measures (PROMs)' in order to explain the different types of PROMs used in ourstudy. We divided the paragraph
'Qualitative' into 'Qualitative methods' and 'Qualitative analysis' to facilitate the readability and comprehensibility. We removed unnecessary repetitions and we moved the information about the recruitment rate to the results.
5. Are research ethics (e.g. participant consent, ethics approval) expressed appropriately?
Ethics approval is confirmed (line 407-08). A participant consent form is reported utilized (line 155) and privacy concerns are reported taken care of (line133). It is stated that "Patients were able to withdraw from the program at any time point" (line 123), which may indicate that patients have been informed of the right to abstain from participation in the study or withdraw
8
consent to participate at any time without reprisal, in accordance with the Helsinki declaration. If this interpretation is correct, the sentence could be revised accordingly.
Response: Yes, this is correct. Action: We revised this section with inclusion of the Helsinki declaration
(page 25 line 511).
It is stated in line 404 that the study was funded by Pfizer. This company has (according to line 131) developed the self-monitoring program that is being tested in the study. Nevertheless, (in line 400-401)

“No conflict of interest” is declared. Although providing thisinformation allowsfor transparency, I think that some sort of additional research ethical elaboration on this financial bond is required.
Response: We totally agree with this suggestion, and we elaborated on this point (page 25 line 503) (see also outlined on page 3 point 10 in this document).
6. Are the outcomes clearly defined?
None of the Patient Reported Outcome Measures that are utilized in this study are described or defined
in the manuscript. Only the full forms of their acronyms are provided. More details on the outcome measures are needed.
Response: We understand that the lack of information about the PROMs may have caused confusion. Action: We added a new paragraph in the method section (see also point 4h) and we provided more information about the PROMs in the introduction.
7. If statistics are used are the appropriate and described fully?
The applied statistical analyses are descriptive and quite simple.
a) The appropriateness of using counts and percentages to produce useful information on adherence to e-mail reminders, as long as the number of e-mail-reminders is not taken into account (line 226-28), is questionable. Or is it so that the number of e-mail reminders equals “the number of PROM-assessments that should have been completed according to the chosen PROM frequency”? That is, was the number of reminder e-mails defined by the patients, by using iMonitor? In that case, to clarify, the text needs to be revised accordingly.
Response: We would like to mention that the number of e-mail reminders was taken into account. Indeed, it equals the number of PROM-assessments thatshould have been completed. Action: to solve
the misunderstanding, we deleted redundant information (i.e. ‘person time frequency’) and we revised the text accordingly (page 15 line 310 and page 9 line 223).
b) The exclusion of patients who only completed PROM(s) once (line 164-65) may lead to loss of information on (inadequate) adherence. I therefore suggest these should be included in the analyses.
9
Response: We totally agree that this may lead to loss of information. Our initial idea was to exclude them. However, we decided to include them in the analyses but we accidentally forgot to revise the text. Action: we deleted the sentence.
c) The method used to measure “correlation” (line 40 and line 109) or “congruence” (line 165) between the PROMS and the DAS28 requires a fuller description.
Response and action: We have changed congruence and correlation consistently into ‘agreement between..’ see also point 1, page 1.
8. Are the references up to date and appropriate?
The referencesin the firstsection of the introduction are quite old. I think more recent publications exist on patient-centred approaches, self-management support, shared decision-making and patient involvement in chronic healthcare. Otherwise, the references seem appropriate.
Response and action: Thank you for the suggestion. Action: We added the following references: Patient-centred care in established RA (Voshaar 2015), de Wit: EULAR recommendations for the inclusion of patient representatives in scientific projects (2011) and Calvert: Maximising the impact of Patient-reported outcome assessment for patients and society (2019) (introduction lines 99 and 139).
9. Do the results address the research question or objective?
a) The first study aim “to obtain experiences (facilitating factors and barriers) regarding online remote monitoring in patients with IRDs” is only partly addressed, because “facilitating factors” is not integrated into the results. Further, what is meant by “facilitating factors” is not defined
in the manuscript.
Response: We agree that these are not specifically mentioned. Action: We changed ‘facilitators and barriers’ into ‘Patients’ experiences’. We also gave a detailed description of how we established the topic guide with relevant factors (page 10 lines 243-245). We revised the title of our manuscript.

b) To the best of my understanding, the second aim, to assess “information about adherence to PROM reminder emails”, is not addressed, referring to the way adherence is calculated (subtext Table 2, line 226-28 and 162-64). This may be solved by revision to a more precise language.
Response and action: We agree and we have described adherence more carefully in both the method section and the subtext in Table 2 (page 15 line 310). We also gave an example of calculating the adherence to facilitate the comprehensibility (page 9 line 223).
c) The third study aim, to assess “correlations between the PROMS and the Disease Activity Score 28 (DAS28)” (line 109) is not met either, because statistical methodsfor calculating correlations have not been used. Further, in the results section, the term "correlations" is replaced by "congruence", without the meaning of this word being explained. Furthermore, in results, the
10
“congruence” with DAS28 is presented only for only two PROMS (Table 3), while five PROMs have been completed by the participants according to Table 1.
Response: We totally understand that using the words ‘correlations’ and ‘congruence’ have caused confusion. Action: Therefore, we will use the word ‘agreement’ consistently throughout the manuscript,see also point 1 from reviewer 1. Regarding this agreement, we were only able to use RAID
and RADAI-5 (two disease-specific PROMs, because only these two PROMs measure disease activity/impact and can therefore be used to measure agreement with objective measures; the DAS28 course in our study. We added this information on page 10 line 230-231.
10. Are they presented clearly?
a) Table 2 and Table 3 are unclear and need revision
Response and action: we changed table 2 and 3 according to your feedback about the way in which we
calculated adherence, how we determined ‘poor’ and ‘good’ classifications.
b) Table 4 also need revision; it is difficult to grasp the meaning of the themes and subcategories in relation to the table heading. Generally, the themes may need further condensation and more precise descriptions of what they refer to (meaningful content) in the context of “facilitators” and “barriers”.
Response and action: We changed barriers and facilitators into patients’ experiences. Inclusion of the topic guide (appendix) will hopefully increase comprehensibility and interpretation of the themes.
c) I recommend replacing “Influencing factors” with ”Patient experiences”, alternatively just adding these words to the subheading on line 243. This may also apply to the manuscript title.
Response and action: We totally agree with this recommendation. We changed the title and we used patients’ experiences consistently throughout the manuscript.
d) Further, should not (properly timed) e-mail alerts be considered as facilitators?
Response: Yes, properly timed email-alerts can function as facilitators. However, this was not specifically mentioned by patients. They experienced that alerts were sent at unfortunate time points (page 19 line 377). Action: We consistently used the term ‘Patients’ experiences” throughout the manuscript (see also point 9A and point 10C).
11. Are the discussion and conclusions justified by the results?
The discussion and the conclusionsin the main manuscript are quite reader-friendly and largely justified
by the results, particularly with respect to the elaborations on the patients’ experiences with the self-monitoring program.
Response: thank you for this compliment. Action: in order to elaborate a little more on patient involvement and self-management, we have added a little more information in the discussion (page 23 line 449) and also with regard to your suggestion regarding the study limitations (point 12).
11
12. Are the study limitations discussed adequately?
The authors adequately refer to the relatively small number of participants as a factor that limits the generalization of the results (line 67 and 369-70). Further, due to the recruitment procedure, the study

population may have been prone to selection bias (line 368). No other study limitations are addressed.
Potential study limitations associated with the qualitative method used should be addressed.
Response: We agree that limitations regarding qualitative methods are not sufficiently addressed.
Action: we added the following limitations to the discussion:
- Data interpretation is difficult (did we really grasp what patients were thinking/feeling)
- Moderator was main care provider for some patients
- Purposive sampling: selection bias (page 23 line 468)
13. Is the supplementary reporting complete (e.g. trial registration; funding details; CONSORT, STROBE or PRISMA checklist)?
Funding details is (partly) reported (line 404). Trial registration: not applicable. The overall reporting is considered incomplete. The STROBE checklist is recommended for better reporting of the quantitative
part of the study. The COREQ checklist (1) and the SRQR recommendations (2) is recommended to guide a better reporting of the qualitative part of the study. I would suggest the authors to read these checklists carefully. A few important aspects that may have influenced the qualitative findings should be reported. For instance, the relationship of the researchers to the participants; were they involved in treatment as well? During the qualitative data analysis, were the themes derived solely from the data, or were they also inspired by theory? Were the findings influenced by the topic guide (line 175)? Personal characteristics(occupation, experience, and training) and perspectives of the members of the
research team that may have influenced the data analysis should be described.
Response: thank you for this suggestion. Action: We added more information about the checklist (COREQ), see method section (page 11 line 262) and the content analysis – thematic approach (line 256).
 14. To the best of your knowledge is the paper free from concerns over publication ethics (e.g. plagiarism, redundant publication, undeclared conflicts of interest)?
There may be an undeclared conflict of interest related to Pfizer, as described above (point 5).
Response and action: we added more information to the Conflicts of Interest paragraph.
 15. Is the standard of written English acceptable for publication?
No. I generally recommend a thorough examination of the written language (proof reading) to clarify the message in this manuscript.
Response and action: proof-reading was performed by a member of the editorial team from our institution. We hope that the written English has improved with the suggestions in this document.

1. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. International Journal for Quality in Health Care. 2007;19(6):349-57.
2. O'brien CB, Harris BI, Beckman JT, Reed AD, Cook AD. Standards for Reporting Qualitative Research: A Synthesis of Recommendations. Academic Medicine. 2014;89(9):1245-51.

VERSION 2 – REVIEW

REVIEWER	Gunnhild Berdal Norwegian National Advisory Unit on Rehabilitation in Rheumatology, Department of Rheumatology, Diakonhjemmet Hospital, Norway
REVIEW RETURNED	26-Nov-2019

GENERAL COMMENTS	Title of manuscript: Patients' experiences regarding self-monitoring of the disease course: an observational study in patients with Inflammatory Rheumatic Diseases General: Thank you for the revised manuscript, which I enjoyed reading. It is considerably improved and addresses the previous response. However, I still have some comments. • Title: I recommend including the word "pilot" in the manuscript title; "an observational pilot study". • Article summary (bullet points): First bullet point: I recommend replacing the word "understanding" (line 74) with "description". Third bullet point: As the PROMs selected for this study are standardised and disease specific (i.e. not individualized and patient specific), they are not designed to capture individual preferences. I therefore recommend replacing the phrase "what really matters to patients" (line 78) with "patient reported measures of disease activity and impact of rheumatic disease", or a similar expression of what these instruments measure. • Introduction, line 92-94: Is the current scientific support strong enough to justify this claim? If not, please moderate the statement (i.e. use "may" instead of "can"). • Introduction, (page 5). In this study, patient reported measures of 'disease activity' (RADAI-5) and 'disease impact' (RAID) is compared with HCP assessments of 'tender and swollen joint counts, acute phase response and a patient's general health assessment' (DAS28). However, as far as I can see, no reason is provided for the choice of these particular instruments by reference to research that reports whether these instruments measure the same constructs or correlate. Please provide relevant references, for example (1-3). • Methods, subheading "Procedure" (line 178) is incomplete. I recommend adding "Sampling" or "Recruitment" to the subheading. • Results, Table 1; please remove the explanation of the subtext acronym PROM (not relevant). • Results: Direct quotes are given in quotation marks (e.g. line 322-323, line 331-332), but reference to the (anonymous) sources are not provided. Please provide or clarify. • Discussion: The discussion is mainly about the qualitative part of the study. The discussion should be expanded to cover all the 3 aims of the study: • Discussion: The quantitative part of the study related to aim 2 (line 128-129) is not discussed at all in the manuscript. Please discuss the related results, as well as potential strengths and weaknesses of the method used. • Discussion: As regards the discussion of the results and methods related to aim 3 (line 422-432), I still find it a bit unclear. The intention with measuring 'adherence to the predefined PROM frequency was very nicely put in words in the authors' response, but I think it can be further clarified in the manuscript discussion. It

	is described that the patients defined in advance the frequency by which they would prefer to complete the PROMs to monitor their disease, and then the researchers later counted the actual number of times they completed the forms, before calculating the adherence to the predefined frequency. What does the result (68.1%) mean (beyond the interpretation "reasonable high" adherence)? Please elaborate/clarify the usefulness of this information (i.e. does it relate to feasibility/usability of the system/patient ability/user engagement?). The sentence on line 447 can probably be moved to this section. • Discussion, line 435-36: "This resulted in a thorough understanding of the concept of self-monitoring, (..)". However, the 'concept' of self-monitoring is not defined or examined in the manuscript. Do the authors mean that the qualitative methods provided information about patients' experiences with 'the practicalities' of self-monitoring? Which then again was used to guide further development of the system before implementation in clinical practice? Please clarify. • Discussion of limitations: The high risk of selection bias in this study is presented two times as two different study limitations (line 436-37 and line 440-41). Please revise to avoid repetition in the text. The manuscript further lacks some critical reflections of the results in light of the analytical approaches applied (quantitative and qualitative). Please discuss (pros/cons of) the method used to answer aim 3 (optionally in light of other adequate quantitative approaches). I also recommend mentioning the common and probable impact of the researches themselves (e.g. in terms of professional backgrounds and scientific theoretical perspectives) on the results of the qualitative study, in the process of constructing the themes and categories. • Standard of written English: I recommend to correct inconsistent use of verb tenses (e.g. line 139-142), and in general a second round of proof reading to clarify the message in this manuscript. 1. Salaffi F, Di Carlo M, Vojinovic J, Tincani A, Sulli A, Soldano S, et al. Validity of the rheumatoid arthritis impact of disease (RAID) score and definition of cut-off points for disease activity states in a population-based European cohort of patients with rheumatoid arthritis. Joint Bone Spine. 2018;85(3):317-22. 2. Leeb BF, Haindl PM, Brezinschek HP, Nothnagl T, Rintelen B. RADAI-5 to monitor rheumatoid arthritis. Clinical and experimental rheumatology. 2014;32(5 Suppl 85):S-55-8. 3. Leeb BF, Brezinschek HP, Rintelen B. RADAI-5 and electronic monitoring tools. Clinical and experimental rheumatology. 2016;34(5 Suppl 101):S5-s10

VERSION 2 – AUTHOR RESPONSE

Title of manuscript: Patients' experiences regarding self-monitoring of the disease course: an observational study in patients with Inflammatory Rheumatic Diseases

General: Thank you for the revised manuscript, which I enjoyed reading. It is considerably improved and

addresses the previous response. However, I still have some comments.

1. Title: I recommend including the word “pilot” in themanuscripttitle; “an observational pilot study”

Response: We think this is a good suggestion, and it may even give a better description of our study.

Action: We added the word ‘pilot’ to the title.

2. Article summary (bullet points): First bullet point: I recommend replacing the word

“understanding” (line 74) with “description”. Third bullet point: As the PROMs selected for this

study are standardised and disease specific (i.e. not individualized and patient specific), they are

not designed to capture individual preferences. I therefore recommend replacing the phrase

“whatreally mattersto patients” (line 78) with “patientreported measures of disease activity and

impact of rheumatic disease”, or a similar expression of what these instruments measure.

Response: The word ‘understanding’ might have been stated too firmly. We agree that ‘description’

might be a better word. Regarding the phrase ‘what really matters to patients’: indeed, these PROMs are

not designed to measure individual preferences. Action: we changed ‘understanding’ into ‘description’

and we changed ‘matters to patients’ into ‘patient-reported measures of disease activity and disease

impact’ (see article summary).

3. Introduction, line 92-94: Isthe currentscientific supportstrong enough to justify this claim? If not,

please moderate the statement (i.e. use “may” instead of “can”).

Response: We think that the role of self-monitoring in IRDs is more and more recognized. There are

studies showing its potential (amongst others our study). We think that the evidence is not that strong

yet, so we think that ‘may’ is a better word. Action: we changed it accordingly (page 5 line 94).

4. Introduction, (page 5). In this study, patient reported measures of ‘disease activity’ (RADAI-5 and

‘disease impact’ (RAID) is compared with HCP assessments of ‘tender and swollen joint counts,

acute phase response and a patient’s general health assessment’(DAS28). However, asfar asI can

see, no reason is provided for the choice of these particular instruments by reference to research

thatreports whether these instrumentsmeasure the same constructs or correlate. Please provide

relevant references, for example (1-3).

Response: We agree that the reasons for using these PROMs have not been sufficiently explained. These

PROMs are however validated, they correlate with objective measures such as the DAS28. Action: In order

to make our choice for these PROMs more evident, we included the suggested references (page 5 line 96-

97).

2

Authors' responses – Patients' experiences regarding self-monitoring of the disease course

5. Methods, subheading "Procedure" (line 178) is incomplete. I recommend adding "Sampling" or

"Recruitment" to the subheading.

Response: We do agree that the word 'Procedure' is insufficient. Next to the sampling/recruitment, this

paragraph also describes the procedure after the recruitment of patients. Action: We added 'Recruitment'

to the subheading (page 8 line 179).

6. Results, Table 1; please remove the explanation of the subtext acronym PROM (not relevant).

Response: The explanation was explained several times before. Action: we removed the acronym.

7. Results:Direct quotes are given in quotationmarks(e.g. line 322-323, line 331-332), butreference

to the (anonymous) sources are not provided. Please provide or clarify.

Response: We understand that direct quotes without any context may be difficult to interpreted. Action:

We gave additional explanation to the quotations: sex and age of the respondent.

8. Discussion: The discussion is mainly about the qualitative part of the study. The discussion should

be expanded to cover all the 3 aims of the study

Response: We agree with your comment. Action: we expanded the discussion, which now covers all three

aims of the study (see discussion).

9. Discussion: The quantitative part of the study related to aim 2 (line 128-129) is not discussed at

all in the manuscript. Please discuss the related results, as well as potential strengths and

weaknesses of the method used.

Response: We agree that aim 2 was not discussed. Action: we added more information about the

agreement between the RADAI-5 en RAID scores in relation to the DAS28 scores and implications of using

this method (page 20 line 374, page 22 line 437).

10. Discussion: As regards the discussion of the results and methods related to aim 3 (line 422-432), I still find it a bit unclear. The intention with measuring 'adherence to the predefined PROM frequency was very nicely put in words in the authors' response, but I think it can be further clarified in the manuscript discussion. It is described that the patients defined in advance the frequency by which they would prefer to complete the PROMsto monitor their disease, and then the researchers later counted the actual number of times they completed the forms, before calculating the adherence to the predefined frequency. What does the result (68.1%) mean (beyond the interpretation "reasonable high" adherence)? Please elaborate/clarify the usefulness of this information (i.e. does it relate to feasibility/usability of the system/patient ability/user engagement?). The sentence on line 447 can probably be moved to this section.

Response: Patients defined their preferred frequency. Study duration was known for every patient. Ifstudy

duration was 1 year and PROM-frequency was one week, 52 reminders were sent. If this patient completed

40 assessments, adherence was (40/52 * 100 =)76.9% (see also example in method section). 68.1% means

the average adherence (so the average of all 47 patients). We agree that the interpretation of this

percentage might be difficult to understand or to interpret. Action: we provided more information about the interpretability and usefulness of the information regarding the average adherence rate of 68.1% (i.e.

system factors and patient factors) (page 22 line 431-436). We also moved line 447 to this section (page

22 line 430).

11. Discussion, line 435-36: "This resulted in a thorough understanding of the concept of selfmonitoring, (..)". However, the 'concept' of self-monitoring is not defined or examined in the manuscript. Do the authors mean that the qualitative methods provided information about patients' experiences with 'the practicalities' of self-monitoring? Which then again was used to guide further development of the system before implementation in clinical practice? Please clarify.

Response: We understand the confusion. The qualitative research methods gave us more information about different aspects, practicality is one of them. We also gained insight in patients' experiences

regarding their own ability to self-monitor, and the tools they needed to do so. Action: we deleted the words ‘concept of self-monitoring’ and changed in into ‘description of factors related to self-monitoring’ (page 22 line 442). These factors (such as technical factors, patient-professional interaction, etc.) are more extensive described in the result section.

12. Discussion oflimitations: The high risk ofselection biasin thisstudy is presented two times astwo different study limitations (line 436-37 and line 440-41). Please revise to avoid repetition in the text. The manuscript further lacks some critical reflections of the results in light of the analytical approaches applied (quantitative and qualitative). Please discuss (pros/cons of) the method used to answer aim 3 (optionally in light of other adequate quantitative approaches). I also recommend mentioning the common and probable impact of the researches themselves (e.g. in terms of professional backgrounds and scientific theoretical perspectives) on the results of the qualitative study, in the process of constructing the themes and categories.

Response: Thank you for these suggestions. Action: We deleted the redundant information regarding the

selection bias. We also provided more information concerning the impact of the subjective method to measure the agreement (see also point 9). Lastly, we mentioned the role of researchers in the interpretability and construction of qualitative themes and categories (page 23 line 447-451).

13. Standard of written English: I recommend to correctinconsistent use of verb tenses (e.g. line 139-142), and in general a second round of proof reading to clarify the message in this manuscript.

Response: We agree with these suggestions. Action: We corrected the inconsistent use of verb tenses

throughout the manuscript. The manuscript was edited by a member from the editorial board of our institute.

4

Authors’ responses – Patients’ experiences regarding self-monitoring of the disease course

- Salaffi F, Di Carlo M,Vojinovic J, Tincani A, Sulli A, Soldano S, et al. Validity of the rheumatoid arthritisimpact

of disease (RAID) score and definition of cut-off points for disease activity states in a population-based

European cohort of patients with rheumatoid arthritis. Joint Bone Spine. 2018;85(3):317-22.

- Leeb BF, Haindl PM, Brezinschek HP, Nothnagl T, Rintelen B. RADAI-5 to monitor rheumatoid arthritis.

Clinical and experimental rheumatology. 2014;32(5 Suppl 85):S-55-8.

- Leeb BF, Brezinschek HP, Rintelen B. RADAI-5 and electronic monitoring tools. Clinical and experimental rheumatology. 2016;34(5 Suppl 101):S5-s10.

VERSION 3 – REVIEW

REVIEWER	Gunnhild Berdal National Advisory Unit on Rehabilitation in Rheumatology, Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway.
REVIEW RETURNED	17-Apr-2020

GENERAL COMMENTS	Referee's third response to the authors Title of manuscript: Patients' experiences regarding self-monitoring of the disease course: an observational pilot study in patients with Inflammatory Rheumatic Diseases General: Thank you for the revised manuscript, which is further improved and mostly addresses the previous response. I have some minor comments, which may provide further improvement: 1. Article summary (bullet points): Second bullet point: I recommend adding some words about why the mentioned patient involvement is a strength of the study (p. 5, line 76), for example ".., ensuring clinical relevance according to a user perspective". Fourth bullet point: I recommend rephrasing the last part to a more general statement addressing generalizability or applicability (p. 5, line 81). 2. Methods, Qualitative analysis (p. 10, line 231): I find the meaning of "..to increase intercoder reliability" unclear in this context. Has the intercoder reliability been measured; e.g. in terms of percent agreement between the coders? If intercoder reliability has been measured in a test-retest procedure, please provide the numbers. If the point here is to show the benefit of using two coders, please rephrase. 3. Discussion, first paragraph: The DAS28 is described in the discussion as "objectively measured" (p. 21, line 375), while in the Introduction (p. 5, line 103-4) it is described as a "composite index", consisting of both objective and subjective assessments. Would it be more correct to state that "The patient-reported RAID showed best agreement with the DAS28 scored by rheumatologist" (or similar)? Please clarify.

VERSION 3 – AUTHOR RESPONSE

General: Thank you for the revised manuscript, which is further improved and mostly addresses the previous response. I have some minor comments, which may provide further improvement:
1. Article summary (bullet points): Second bullet point: I recommend adding some words about why the mentioned patient involvement is a strength of the study (p. 5, line 76), for example ".., ensuring clinical relevance according to a user perspective". Fourth bullet point: I recommend rephrasing the last part to a more general statement addressing generalizability or applicability (p. 5, line 81).

Response: Thank you for this valuable suggestion. Action 1: We added the following sentence: '… which
increases the clinical relevance according to a user perspective'.
Action 2: Regarding bullet point 4, we changed this sentence into: 'The selective and small study population might have influenced the generalizability and applicability of the study'.
2. Methods, Qualitative analysis (p. 10, line 231): I find the meaning of "..to increase intercoder reliability" unclear in this context. Has the intercoder reliability been measured; e.g. in terms of percent agreement between the coders? If intercoder reliability has been measured in a test-retest procedure, please provide the numbers. If the point here is to show the benefit of using two coders, please rephrase.
Response: We understand the confusion regarding 'increasing intercoder reliability'. We did not use a test-retest procedure. Instead, we wanted to indicate the benefit of using two coders for the coding process. Action: we changed the sentence into: 'One of the researchers (LR) and a research assistant independently coded the transcripts, in order to enhance the coding process, data interpretability, and trustworthiness'.
3. Discussion, first paragraph: The DAS28 is described in the discussion as "objectively measured" (p. 21, line 375), while in the Introduction (p. 5, line 103-4) it is described as a "composite index", consisting of both objective and subjective assessments. Would it be more correct to state that "The patient-reported RAID showed best agreement with the DAS28 scored by rheumatologist" (or similar)? Please clarify.
Response: Thank you for the alertness. By using the word 'objective' we wanted to point out that DAS28 is the most objective measure available in our study. Indeed, the DAS28 also contains a

subjective part (assessment of disease activity by patients using a 0-100 scale). Action: We changed the sentence into: 'The disease-specific and patient-reported RAID showed the best agreement with the DAS28 assessed by the rheumatologist (see also page 6, line 130).