Author's response to reviews

Title: Case management for patients with chronic systolic heart failure in primary care: The HICMan exploratory randomised controlled trial

Authors:

  Frank Peters-Klimm (frank.peters@med.uni-heidelberg.de)
  Stephen Campbell (stephen.campbell@manchester.ac.uk)
  Katja Hermann (katja.hermann@med.uni-heidelberg.de)
  Cornelia U Kunz (kunz@imbi.uni-heidelberg.de)
  Thomas Müller-Tasch (thomas.mueller-tasch@med.uni-heidelberg.de)
  Joachim Szecsenyi (joachim.szecsenyi@med.uni-heidelberg.de)

Version: 3 Date: 22 April 2010

Author's response to reviews: see over
to the editors of
"Trials"

22/04/2010

Manuscript: "Case management for patients with chronic systolic heart failure in primary care: The HICMan exploratory randomised controlled trial"

Dear editor(s),

Many thanks for reviewing the previous manuscript and for giving us the opportunity to send a revised manuscript giving a point-by-point response to the issues raised.

We are grateful for the comments of the reviewers. We have revised the manuscript in response to these helpful comments.

We hope these changes meet with your approval. Should you require further modifications, please do not hesitate to get in touch.

Many thanks in advance.

Sincerely yours,

Frank Peters-Klimm, MD
<table>
<thead>
<tr>
<th>Reviewers’ comments</th>
<th>Response to comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General</strong></td>
<td></td>
</tr>
<tr>
<td>Editorial board: We would be grateful if you could address the comments in a revised manuscript and provide a cover letter giving a point-by-point response to the concerns. Please also highlight all changes made when revising the manuscript to make it easier for the Editors to give you a prompt decision on your manuscript.</td>
<td>We decided to colour the changes red as the manuscript needed some extensions and a shortening in the Discussion section (Summary of main findings). We think, together with the point-by-point responses, this approach is most appropriate.</td>
</tr>
<tr>
<td>R1: This is a moderately-sized (for this type of intervention) RCT of case-management in a German primary health-care setting. The study results are equivocal which the authors acknowledge... In summary, a useful ‘pilot’ study. Would have been good to exclude highly stable patients with normal NT-proBNP (which I suspect most had). Needs some language corrections before being published.</td>
<td>Reviewers’ comments were sometimes strikingly different. R1 and 3 perceived we were concluding appropriately while R 2 seemed to have the impression that we were presenting the results unduly positive. We are well aware of the limitations and strengths of the trial - as to the limited power but also as to the comprehensiveness of reported health and care outcomes, which are still appropriate and useful for an exploratory trial.</td>
</tr>
<tr>
<td>Reviewer</td>
<td>Comment</td>
</tr>
<tr>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td>R2: The manuscript report on patients in general practice randomised to guideline directed treatment by doctors assistants compared to usual office care. The study is small and limited by funding restrictions. The endpoints are the results from 5 different versions of quality of life and behaviour assessments. There are no effects on QoL outcomes in this study, but the conclusions are presented on an unduly positive note. Probably because the authors find that that the patients score better on behaviour and response to the services provided. The discussion is very limited on the interpretation of the relevance of the different tests in this context. It is well known that patients on placebo benefit to the special care delivered in clinical trials. To the average cardiology reader the clinical relevance of the different qualitative scoring systems are controversial. The manuscript could be shortened. Needs some language corrections before being published.</td>
<td>We feel that we presented and discussed the results of the trial appropriately by means of transparency and of the conclusions that we have drawn and feel confirmed by the comments of R 1 and R 3. We do not perceive that we overstate the found results regarding patient-reported outcomes (HF self-care, quality of chronic care) and clinical outcomes (HRQOL, mortality or hospital admission rates). As requested, we provide now additional data (NT-proBNP levels of the patients, see below and cost of time of Case Management performed by the Doctors’ assistants). The manuscript was proofread again by two native speakers.</td>
</tr>
<tr>
<td>R3: This trial is well done and described. .... Acceptable quality of written English.</td>
<td></td>
</tr>
</tbody>
</table>

**Abstract**

R1: English expression in the abstracts is ‘mangled’. Fairly clear what is meant but room for much improvement. For instance:- “In this sample, with little room for improvement due to high baselines,” – now what exactly are ‘high baselines’? We specify now on evidence-based pharmacotherapy and CHF self-care.

R2: P 2, results: high guideline adherence must be due to selection bias with reference to how the patients were selected. The reviewer may be right. A crucial inclusion criterion was ‘ascertained systolic heart failure’. Therefore, it can be assumed, that this confirmation can lead to a good guideline adherence (see also Discussion: Strengths and Limitations)

**Background**

No reviewers’ comments

**Methods**

R 1: no comments
R 2:

- Design: The study was part of a subproject of quality of Life within the German “Competence Network Heart Failure” but there is no information on the overlap or relationship to the overall project.

Participants - recruitment and assignment:
- P 7, l 9 from bottom: The randomisation procedure should be explained. I do not understand what is meant by a weekly basis. Considering the low number of patients per practice randomisation on a weekly basis may suggest that the doctors could choose between several options.
- According to my understanding of the protocol, patients in both treatment arms were seen by the same doctor, thus there must be some carry over effects between the two cohorts.

- Outcome measures:
  P 9, Top: There is overlap between SF-36 and KCCQ. Was there any indication for this in the study and why was both methods used in the study. I get a feeling that the patients were over exposed to questionnaires with a possibility of trial fatigue.

- Data Collection and management:
  P10, bottom: Please explain why not all patients adhered to the questionnaires reporting.
- Statistical Methods:
  P 11, l 6 from top: Why was a one sided t-test applied and not a two sided?

- The “Competence Network Heart Failure” is a nationwide research network that bundles the scientific expertise in a large-scale research network. Its aims are the coordination of basic and applied clinical research as well as dissemination of findings into clinical practice in order to consolidate and perpetuate the achieved improvements. Each project has its own trials or studies with no overlap or contamination. In the Method section “Data collection and Management” details of the “relationship” (support) are already given.

- During the enrolment phase (2 months), in all practices for each enrolled patient a pseudonymised randomisation document was sent per fax immediately to the CCCT. At the end of the week, the accumulated randomisation documents were assigned as described. Each practice was sent back the randomisation document with the result of the randomisation.
- R2 refers to possible contamination on which we comment already in detail in the Discussion section (S+L).

- HRQOL is a multidimensional concept comprised of several domains, including physical/biological factors, symptom status, functional status, health perceptions, and overall well-being [1]. In research, the use of generic and disease-specific instruments to assess HRQOL is the recommended standard [2,3].
- A trial fatigue suspected by R2 can neither be excluded nor be proven. However, except for the SF-36 where an imputation algorithm is included in
the syntax for building the composite scores, missing data on item level is an explaining factor. Generally, if including losses to follow-up due to other reasons (e.g., death) as it is the case in our study, response rates between 70 and >90% can be regarded as fairly well.

- According to the phase of the development and evaluation of the complex intervention, we used a one-sided t-test as the background implicated a positive effect of CM on HRQOL. The results of this ‘pilot’-trial should be seen as the base for or against further trials with eventual modifications of the intervention, of the targeted patients and of targeted outcomes.

### Results

| R1: Patients enrolled were much younger than epidemiologically expected and with rather good renal function. What were the NT-proBNP levels (included in list of abbreviations but no data supplied). | We greatly appreciate the answers of the reviewers:
  Two crucial inclusion criteria for enrolment should be reminded (for details see study protocol): left ventricular ejection fraction ≤45% (assessed within the last 24 months) and dyspnea (NYHA II-IV or NYHA I, if hospital admission because of CHF within the last 24 months).
  We provide now the data on NT-proBNP, and obviously patients did not have ‘mostly’ normal values. Indeed, some had low levels (under level-lowering treatment!): 11 patients had values between 50 and 100 pg/ml; another 24 between 101 and 300 pg/ml.
  We agree with the conclusion of R1 that it would have been good to exclude highly stable patients, although the use of BNP as risk stratifier was not in the design of this trial, but could be considered in the future. |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>NT-proBNP is included in the abbreviation list but no data are given in the paper. I suspect these data are being withheld – probably because the patients have mostly normal values – casting doubt on the accuracy of the diagnosis and accounting for the good prognosis. NT-proBNP is an excellent diagnostic and risk stratifier in this population. These data should be supplied.</td>
<td></td>
</tr>
<tr>
<td>R2: NT-proBNP is listed in the list of abbreviations. The numbers are not given but would be helpful in understanding the degree of heart failure.</td>
<td></td>
</tr>
<tr>
<td>R2: ([P 6, bottom (?)]) How many patients had previously been hospitalised for heart</td>
<td>As already provided in Table 4 (see additional file), 23 patients of the IG</td>
</tr>
</tbody>
</table>
failure. Also how many patients had a previous myocardial infarction? Infarct before the age of 60 has limited information in this context.

caused 36 HF admissions and 29 patients of the CG caused 35 HF admissions pre-observation period. 7 patients in the IG (and 5 in CG) had previous myocardial infarction. 17 cases were missing, and in 168 patients it was documented as either “no or not known” without further possibility to discriminate. In this case, reporting of this variable seems not to be informative. “Infarct before the age of 60” is one of the 5 reported ‘cardiovascular risk factors’ more below in table 2 and is not intended as a substitute for Hx of MI.

R1: Of the 10,397 who did not meet the inclusion criteria, how many were thought to have heart failure? Presumably a lack of available LVEF or LVEF >45% were the main reasons for exclusion.

Literature suggests a 2% overall prevalence of HF patients in the population, approximately 50% with LVSD and 50% with HF with preserved systolic function. Considering the selection effect from population to primary care we conservatively expected that our multifaceted case finding strategy would double the overall population prevalence, which was the case, finally. A lack of available LVEF was the crucial “barrier” in these cases, where all other inclusion criteria were not violated (altogether 8) and the exclusion criteria (altogether 14, e.g. dementia, residency in nursing home, addictive disorder) were not fulfilled and the GP could not indicate an echocardiography as part of the routine care. We cannot quantify this proportion, but it was reported to be a rare case (personal communication to the study nurse). (In a previous research project, we were able to recruit approx. 5 patients with LVSD per practice, this time with the improved case finding almost 7 patients per practice. It would be interesting to validate this selection of patients to obtain a ‘real practice prevalence’, e.g. by
<table>
<thead>
<tr>
<th>R1: Two thirds of the population in this study is in functional class I and II and limitations in QoL are small and any improvement may accordingly be small. This raises questions about a possible selection bias in the authors “brainstorming” case finding strategies. I understand that most patients in functional class III and IV are cared for by specialists in Germany.</th>
<th>Two thirds were in functional class II and almost one third is in functional class III. As outlined above, the enrolment strategy was to include all eligible patients. Patients in functional class IV are practically always cared for in secondary care. As our study shows, also patients in functional class II are seen by specialists (see Table 4). There are no robust data from health services research to comment on functional class III. The multicentre HF ACTION trial (Flynn et al., Ref. 38 in the manuscript), however, in the US, but in an ambulatory secondary care setting, revealed the same proportions of NYHA functional status as our trial. Maybe this reflects another selection bias, but it could also indicate the prevalence of disease stages of CHF in ambulatory care.</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1: What was the travel time incurred to and from home visits. This is an often forgotten aspect/cost to home visits.</td>
<td>R1 is right. We completed the time cost of the DAs for performing the CM and add now the travel time. Furthermore, we calculated the total time DAs spent on each patient and present it as mean hours according to the frequency of CM (which was stratified according to NYHA functional status).</td>
</tr>
<tr>
<td>R2: The time spent on patient control and doctor reports by the assistants is difficult to understand. Please add up the time and convert into total working hours per patients in order to understand the implications of this care. ... (P15, l 12 from bottom) The feasibility to conduct should be documented by how many hours or days totally were spent during the year on each patient.</td>
<td>Specified, also in the table</td>
</tr>
<tr>
<td>R2: P14, l 3 from bottom: What is meant by practice attendance of 27.6 and 23.9 Is it days, percent or times. Please be specific.</td>
<td></td>
</tr>
</tbody>
</table>

**Discussion**

| R1: The discussion does not need to re-summarise the results and the abstract and could be usefully abbreviated. It gets very repetitive in parts, restating what has already been said three times already. | We shortened the summary, but as health and care outcomes changed on different levels, we think that it is helpful to structure the discussion as it is common in many journals. |
### R1: The authors state that the study is skewed towards lower classes but it looks to me as though it is skewed towards the middle class – but then most people in Germany might be middle class. This needs to be clarified.

Social inequality is less in GER than for example in the UK. According to the German Index for social class, most people belong to the middle class: According to the German National Health Interview and Examination survey from 1998, in the western part (applicable to our study) males belong to the lower, middle and upper class in 19.4, 55.2 and 25.5%. Accordingly, females in 25.7, 54.8 and 19.5%. Compared to our sample with >70% males, there is approximately a 10% difference in the upper and the lower class.

### R2: The discussion is very limited on the interpretation of the relevance of the different tests in this context. It is well known that patients on placebo benefit to the special care delivered in clinical trials. To the average cardiology reader the clinical relevance of the different qualitative scoring systems are controversial… Twice as many patients were hospitalized for heart failure during the study. Although the numbers are small, this is disturbing. The authors claim it may be due to better care, but the alternative that over focusing may create anxiety and undue admissions should be discussed. This is also supported by slightly more patients withdrew from the study in the assistant managed group.

See above

Definitive interpretations on reasons for hospitalisation cannot be given from this study, but we agree with the reviewers’ perspective and another potential interpretation.

### R1: Page 17 does ‘negative’ mean neutral or truly negative (ie harm).

‘Negative’ was not the correct term. ‘Neutral’ is the correct term.

### R3: My only comment would be that the issue of adherence to drug treatment may be commented upon since in both groups the number of patients on various drugs decreased.

Except for the combination of ACE inhibitor/A2RA and β-blocker in the control group, we cannot see big decreases on a still high level of adherence. One might consider the loss to follow-up and intolerance of the medication, which cannot be verified as part of this project.

### Conclusions
**R2: P 20, l 5 from top:** I disagree with that this trial demonstrated positive effects of clinical relevance for patients with mild heart failure. I suggest that the conclusion should be softened.

<table>
<thead>
<tr>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>R2: An incomplete reference (22) is given.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Table 1</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>R2: Please explain patients per quarter. Is it per doctor or per practice or something else?</td>
</tr>
</tbody>
</table>

We agree that in the Methods section, we mainly give the references to open access publications where the intervention is described in detail. As R2 stated that the article should be shortened, we think we should stress the references.

As indicated, the given numbers are per practice, which is in our study makes almost no difference as only two practices participated with two physicians.

Reference List

