

RENEWING HeALTH

REGIoNs of Europe Working
toGether for HEALTH

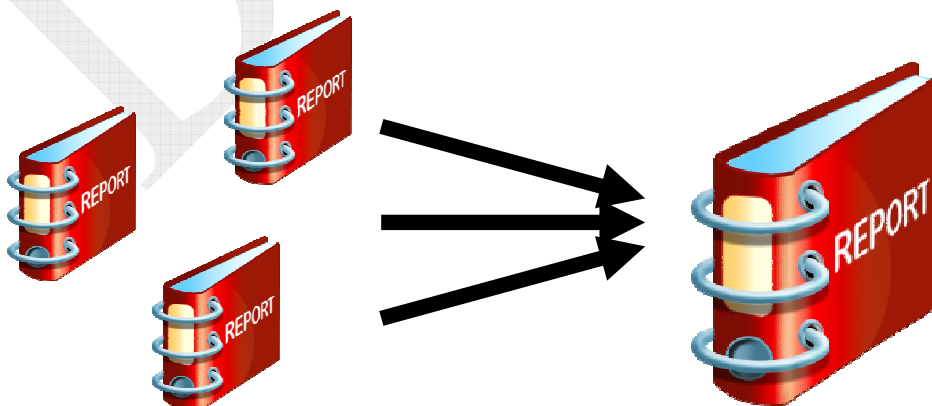
Guideline on reporting of results at cluster level in Renewing Health

Kristian Kidholm (editor), Vassilis Aletras, Alice Stærdahl Andersen,
George Dafoulas, Anne-Kirstine Dyrvig, Lise Kvistgaard Jensen,
Anna Kotzeva, Inger Lindberg, Anna-Leena Orsama, Claus Duedal Pedersen,
Panagiotis Stafylas, Christina Wanscher.

First version, October 2013

FROM PILOTS

TO A CLUSTER



CONTENT

1. Introduction

2. Reporting of results within each domain

2.1. Domain 1: Description of the health problem and characteristics of the application

2.2. Domains 2 and 3: Safety and clinical effectiveness

2.3. Domain 4: Patient perspectives

2.4. Domain 5: Economic aspects

2.5. Domain 6: Organisational aspects

2.6. Domain 7: Socio-cultural, ethical and legal aspects

References

Appendix 1. Codes for updating of STATA

1. Introduction

The Renewing Health project consists of 21 trials or pilots of telemedicine applications for patients with diabetes, COPD or CHF. These pilots are divided into 10 groups (called clusters) and within each cluster similar telemedicine solutions are tested on similar patients groups with either diabetes, COPD or CHF. To ensure the similarity within each cluster, a common protocol has been made for each cluster describing the common outcome measures, eligibility criteria etc.

During the summer and fall of 2013 each pilot has produced a report describing the results at pilot level. These reports are similar to the deliverables DX.7-Y (e.g. for the diabetes trial in Veneto D8.7-2). All reports describing the results at pilot level are structured and produced in accordance with the "Guideline on analysis and reporting of results from the pilots in Renewing Health".

Based on these pilot reports the cluster managers for all clusters with more than 1 pilot (that is clusters 1, 2, 4, 5, 6, and 7) will produce a deliverable during October and November 2013, describing the results at cluster level in which the pilot results are compared and analysed together. Comparing the results from several countries will enable the cluster managers to assess the generalisability and transferability of the results from Renewing Health.

This guideline describes how the cluster managers in Renewing Health should carry out the reporting of the results at cluster level based on the results from the pilots in each cluster.

The guideline presents a minimum level of analysis for all clusters to carry out, mainly on the basis of the minimum dataset collected by all pilots (see below). Similarly, the guideline includes information about statistical analysis, but the cluster managers are recommended to consult local experts in choice of statistical tests and use of multivariate analysis. The guideline can also be of assistance to the cluster teams, for publishing their cluster level results in scientific journals.

This guideline is mainly based on the corresponding guideline for analysis at cluster level and the framework for assessment (called MAST) used in Renewing Health (see Kidholm et al. 2012). The guideline describes how cluster level analysis can be done within each of the seven MAST domains. Please notice, however, that the results within each domain are closely related to the description of the other domains, e.g. the clinical impact (domain 3) depends on the content of the telemedicine service (domain 1) and the organisation (domain 6) and all of these determine the economic results (domain 5). Therefore, the description of the individual domains should consider the implications of the content of the other domains.

Finally, notice that no definitions of terms and concepts (e.g. the domains and outcome measures) are included in this report, but can be found in the pilot guideline, see Kidholm et al. 2012. Similar the use of statistical tests and estimates (e.g. the use of mean versus median values) must be found in the pilot guideline and is not described here.

2. Reporting of results within each domain

2.1. Domain 1: Description of the health problem and characteristics of the application

The aim of the description of the health problem and the telemedicine application at cluster level is to describe similarities and differences between the pilots in the cluster. All relevant information should be included in the reporting from the pilots and the cluster protocol.

2.1.1. Health problem

Description of the health problem

Based on the pilot reports describe the health problem of the patients in the pilot studies. Generally the patients should be similar in the pilots and only minor differences are expected.

Quantification of the burden of the disease/health problem

There may be differences in the epidemiology of the health problem in the different countries. Please describe differences.. As an example the number of patients (prevalence or incidence) or the mean age of the patients may differ. If such differences exist, the implication for the transferability of the results for Renewing Health must be considered.

Current management of the health problem (usual care)

Present a description of disease management in the current clinical practice (usual care) in the local context of the pilots and clarify differences between the pilots. This aspect is essential since these differences may explain differences in the pilot results.

2.1.2. Telemedicine application (technology)

The purpose of the telemedicine application is similar for all pilots and should be described based on the common protocol. However, the devices and technical solutions may vary between pilots. Describe the differences based on the pilot reports. For example description should be made of the level(s) of health care (primary, secondary, tertiary care, emergency unit, home care, etc) where the new telemedicine-based service is introduced, as well as the professional profiles involved. Figures or diagrams from the pilot report may be used for this purpose.

Technical characteristics of the application

Please write a short technical description of the type of device(s), operations, software and hardware, imaging and other available features and describe differences between the pilots.

Similarly describe the life cycle of the technology and the level of maturity of the technology, as well as the actual market situation.

Requirements for the use of the telemedicine application

In order to describe the requirements for successful use of the telemedicine application description should be made of differences in the investments made in each pilot. This includes all material investments necessary for the functioning of the system have to be defined including e.g. devices, machinery, computer programs, necessary facilities either in the clinical setting or in the patient's home (e.g. independent room for teleconsultations, broadband connection, etc).

Discuss differences in the need for a common infrastructure and interoperability within the pilots. Include differences in the needs of training for health care professionals and patients between pilots.

2.2. Domains 2 and 3: Safety and clinical effectiveness

When looking at the results at cluster level please focus on describing the methodological quality of the RCT, the estimated clinical results, a meta-analysis of the results across pilots and a discussion of the reasons for differences between pilots and the generalisability of the results.

2.2.1. Methodological quality

Information about the quality of the individual pilot RCT can be found in the pilot reports. Please include a summary in the cluster level report describing the degree to which the RCT was carried out as planned. This could include information about:

- Planned sample size?
- How randomisation was done in practise?
- Degree of blinding of patients, staff, outcome assessor and analyst?
- Recruitment of patients: Number of patients tested for eligibility, declining to participate, lost to follow up and number of patients included in the analysis?
- Comparability of intervention and control groups at baseline – e.g. have statistical differences been found?

This information is important because differences in the methodological quality of the trials may explain differences in the results.

2.2.2. Estimated clinical outcomes

Based on the pilot reports please produce a table describing the effects on the primary outcome in the cluster. If common secondary outcomes exist these could also be described. This can be done similarly to the table below. Describe also differences between the countries if these are found. Notice that if data have been collected for a period of more than 6 months, these should be reported.

Table 1. Results of analyses of the primary outcome in all pilots in the cluster

Pilot	Intervention Baseline	Intervention After 6 month	Control Baseline	Control After 6 month	Mean difference after 6 month between groups (CI 95%)
Country no. 1.	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (CI to CI)
Country no. 2	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (CI to CI)
Country no. 3	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (CI to CI)

2.2.3. Meta-analysis

A meta-analysis must be carried out in order to estimate the overall effect on the primary outcome in all pilots at cluster level.

If scientific publications are planned, the strategy should include thoughts on possible confounders and which variables to include in meta-regressions if the level of heterogeneity between studies is high.

Necessary data

For the safety and clinical effectiveness, include the following data:

- Primary outcome according to group (intervention/control)
- Other common outcomes across the cluster, depending on the aim of the analyses. For scientific publications more analyses are necessary than for reporting to the Commission
- Adverse events, e.g. deaths, technical difficulties according to group

Analysing data

Meta analyses made by use of the STATA software:

For each pilot site, input the variables:

Events_{exposed}

NoEvents_{exposed}

Events_{unexposed}

NoEvents_{unexposed}

Based on necessity, RR, OR, log OR and standard errors can be calculated. These are used in the meta analysis.

Basic meta analyses are calculated in STATA with the command **metan Events exposed NoEvent exposed Event unexposed NoEvent unexposed**. The default model is fixed effects, but in most cases, a random effects model will be preferred. If that is the case, use the suffix **, random**.

STATA can be used for this purpose but other statistical software can also be used. STATA will provide a table showing the results and a graph presenting the forest plot. From these, it can be interpreted whether there is an overall effect of telemedicine on the chosen outcome for the disease in question. Also, it will appear if the effect is statistically significant.

In addition, the I^2 (along with p-value) is part of the output. That is an indication of the between-study variance. As a rule of thumb, if the I^2 is below 50, the studies are quite homogenous, and a fixed effects meta-analysis is appropriate. If the value is above 50, a random effects model should be used due to heterogeneity between studies. Keep in mind that the random effects model does NOT adjust for heterogeneity – it only allows for the presence of it! If the heterogeneity is above 80, there is reason to discuss the appropriateness of carrying out the meta analysis at all. Also, in these cases, a meta regression should be carried out to investigate the causes of heterogeneity.

Reporting results

The output from STATA looks like this:

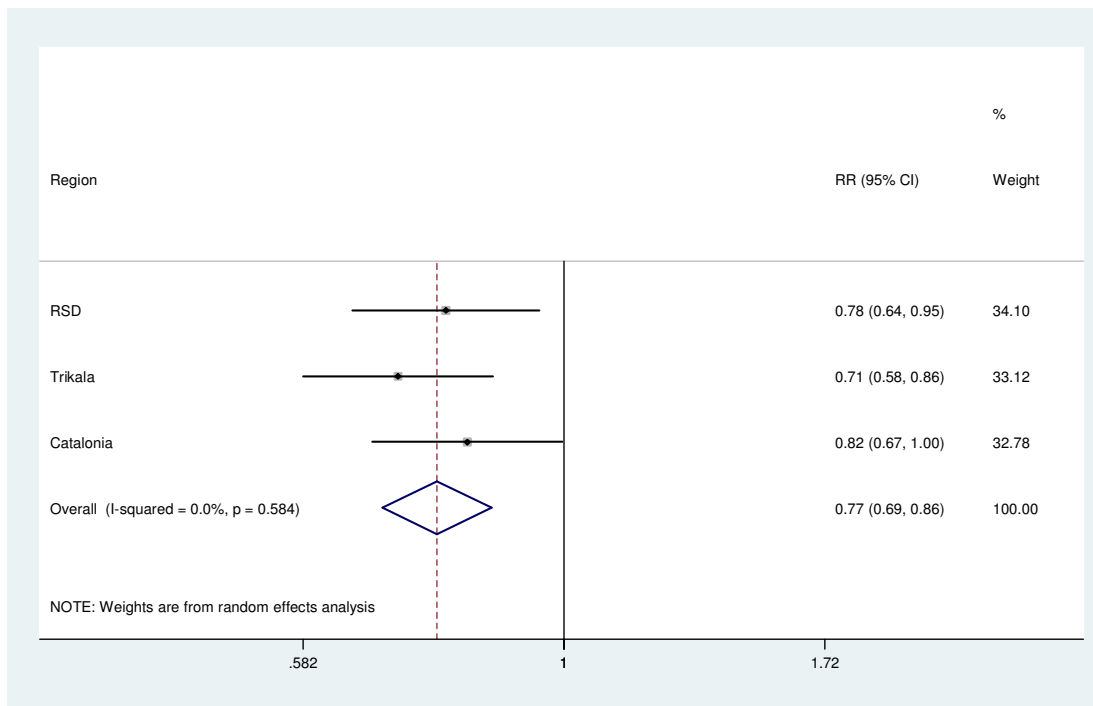
Study	RR	[95% Conf. Interval]		% Weight
RSD	0.783	0.645	0.950	34.10
Trikala	0.708	0.582	0.863	33.12
Catalonia	0.818	0.671	0.997	32.78
D+L pooled RR	0.768	0.686	0.861	100.00

Heterogeneity chi-squared = 1.08 (d.f. = 2) p = 0.584
I-squared (variation in RR attributable to heterogeneity) = 0.0%
Estimate of between-study variance Tau-squared = 0.0000

Test of RR=1 : z= 4.56 p = 0.000

So, what the output describes is the relative risks (RR) for each setting, 95% confidence intervals (CI) and the % weight given to each study. In this example, all studies have positive effects, i.e. the intervention protects the patients from having an event. All effects are statistically significant. The text below the table describes the level of heterogeneity, i.e. the level of variance between the studies. The I^2 is usually reported along with the p-value. In this case, $I^2 = 0.0\%$, $p=0.5$.

In addition to the table and explanatory text, STATA develops a graph (unless it is told not to by the command **nograph**). The graph appears in a separate window, and looks like this:



Please note that text sizes have been changed. Otherwise the graph looks like the default output from STATA. For each region, the RR and confidence intervals are presented graphically. The size of the box on each horizontal line depicts the weight given to each study. Since the studies in this example are of similar size, the weights are close to equal and the boxes are of similar size. The star below the horizontal lines is the summary measure. The width represents the confidence interval, and the corners of the height show the summary estimate.

2.2.4. Discussion of differences

After presenting description of the quality of the RCT, the estimated effects on outcome measures and safety at pilot and cluster level, the results should be discussed with regard to similarities and differences among the pilots. It is important also to include discussion of the generalisability of the results, e.g. whether other patient groups (with higher or lower degree of severity of the disease) would benefit from the intervention.

2.3. Domain 4: Patient perspectives

In most pilots patient perception is assessed by use of the Service User Technology Acceptability Questionnaire (SUTAQ) from the WSD-project. Therefore the cluster level analysis should compare the results from the different pilots and countries and try to give an overall assessment of the patients' perception of the technology. In addition to this the qualitative studies of patient perception (that has been made in a few pilots) should be briefly described with regard to methods and main results.

Before presenting the results, write a short description of the size of the total sample of patients and the response rate. Subsequently produce the subscale from each pilot by use of a table similar to the one below. Notice that median values may be more relevant if the distribution is not-normal as described in the pilot guideline.

Table 2. Estimated subscales based on respondents answers to SUTAQ

Subscale	Intervention After 3 months	Intervention After 6 month	Mean difference after 6 months (CI 95%)
Enhanced care	Mean (SD)	Mean (SD)	Mean (CI to CI)
Increased accessibility	Mean (SD)	Mean (SD)	Mean (CI to CI)
Privacy and discomfort	Mean (SD)	Mean (SD)	Mean (CI to CI)
Care personal concerns	Mean (SD)	Mean (SD)	Mean (CI to CI)
Kit as substitution	Mean (SD)	Mean (SD)	Mean (CI to CI)
Satisfaction	Mean (SD)	Mean (SD)	Mean (CI to CI)

Make a short description of differences and similarities between pilots based on the tables. Notice that differences could be caused by differences in the technologies used, differences in the organisation of the services or differences in the patients (e.g. differences in familiarity with computers or cultural differences).

If a cluster is able to merge data from the different pilots into one database it is also possible to do statistical analysis on the combined dataset, and test relation to explanatory variables e.g. in order to identify the patients with the highest degree of acceptability, but this is not mandatory.

2.4. Domain 5: Economic aspects

In the description of the economic aspects at cluster level please focus on the methodological quality of the economic analysis, the results at pilot level, an assessment of the results across pilots and a discussion of the reasons for differences between pilots and the generalisability of the results. Describe the results from the societal economic evaluations as well as the business cases..

2.4.1. Methodological quality

As described in the "Guideline on analysis and reporting of results from the pilots in Renewing Health", a very detailed plan for the analysis of the economic aspects at pilot level has been made. However, in practise there may be differences in the pilots' ability to obtain the data needed and produce the tables requested. Therefore a description should be made of the quality of the data and results from the relevant pilots. This could include:

- Are the same types of resources included in all pilots (look at Table 9 in the pilot guideline)?
- Are the dataset complete or are data missing?
- Are investment, uncertainty, overhead costs and the costs of use of time for patients and relatives included in the same way in accordance with the pilot guideline?
- Is the price-year 2011 in all pilot analysis as described in the guideline?
- Recruitment of patients: Number of patients tested for eligibility, declining to participate, lost to follow up and number of patients included in the analysis?
- Comparability of intervention and control groups at baseline – e.g. has statistical differences been found?

2.4.2. Estimated economic outcomes at pilot level

Subsequently describe the results at pilot level. This should include information about the average use of resources per patient in the intervention and control group (Table 12 in the pilot guideline), the average costs of the treatment per patient in the intervention and control group (Table 15 in the pilot guideline), and the estimated ICER. Here copies of the relevant tables from the pilot reports could be copied.

2.4.3. Assessment of overall economic results at cluster level

If data from the different pilots are merged into one database, analysis of the average use of resources and the costs per patient can be estimated at cluster level. This way statistical analysis of differences between the control group and the intervention group and the different pilots can be made e.g. by use of ANOVA or regression analysis.

If the data from the pilots are not merged into one database, a more qualitative approach must be taken in which the results from the different pilots are compared with regard to similarities and differences. An example can be found in Drummond et al. (1992) in which economic results from pilots in 4 counties are described separately.

As an alternative Meta-analysis of differences in the use of health care services can be made, as described in section 3.2.

2.4.4. Discussion of finding

Drummond et al. (1992) have identified a number of reasons why the costs of medical technologies may vary from country to country. Some of these are relevant for the costs of telemedicine services:

- Differences in the demography and epidemiology of the disease, e.g. number of patients and the age of the patients.
- Differences in the relative price level for health resource e.g. the price of nurses and doctors per hour or the price of transportation of patients per kilometre (see Table 10 in the pilot guideline).
- Differences in the availability of health care resources, e.g. the number of kilometres to the nearest hospital.
- Differences in the clinical practise and conventions, e.g. the number of hospitalisations per year per patient in usual care.
- Differences in the incentives of health care professionals, e.g. the payment system for GPs.
- Differences in the data on costs available for the analysis, e.g. whether information on use of GP and municipality services can be found and included in the analysis.

Based on the results described at pilot level, the analysis at cluster level should consider whether one or more of these explanations are possible reasons for the differences found in the economic analysis. Afterwards discuss the generalizability of the results from Renewing Health to other countries

With regard to the business case please consider to what extent variation in the reimbursement of the telemedicine service has an impact on the result of the business case in each country. Based on this discuss the generalizability of the business case to other countries, e.g. by pointing out the factors most important for the result of the business case.

2.5. Domain 6: Organisational aspects

Even though the primary outcome, the patient groups and the telemedicine devices are fairly similar in the pilots included in each cluster, large differences may exist in the organisation of the work related to the service because the local organisation depends on the general organization of health care in the different countries. Therefore a comparison of the organisational aspects in the pilots is important for an understanding of the differences in the clinical and economic outcomes and an assessment of the generalisability of the overall results.

As described in the pilot guideline all pilots should assess the organisational aspects by covering three topics:

- Process
- Structure
- Culture

2.5.1. Process

Based on the description of the pilots, please write a short description of the changes in work flow, patient pathways, staff changes and communication caused by the implementation of the telemedicine.

Please focus on differences and similarities between pilots with regard to:

- Changes in the use of health care professionals, e.g. task shifting
- Changes in the patient pathways, e.g. expected reduction in inpatient or outpatient visits
- Additional education or training for health care professionals.
- Changes in communication.

2.5.2. Structure

Structural outcomes are related to how the structures of work flow or the involvement of different organisational units changes as a result of the introduction of telemedicine. Since the structure of the health care sector varies between countries, differences in the structural changes in the pilots must be expected. Based on the pilot descriptions focus on the cluster level should be at:

- Description and number of units collaborating in the production and delivery of the service
- Changes in organisation of generalist and specialist tasks
- Changes in geographical spread
- Changes in time spent on travel

2.5.3. Culture

The cultural aspects should be assessed by interviews with or questionnaires given to the health care professionals in all pilots. In these interviews or questionnaires ten mandatory issues (regarding technical difficulties, usability, task shifting, communication and satisfaction) should have been assessed, as described in the pilot guideline.

Based on the answers to the ten questions make an overall assessment of the views of the health care professionals of the telemedicine services.. Notice that differences in the view of the respondents in the different pilots may be caused by differences in the structure of the health care system in the pilots, the resource level of the health care system or the education level of the health

care staff using the telemedicine services. Remember also to include a short description of the data collection method and the respondents included in each pilot.

2.5.4. Discussion

Based on the above the assessment of the organisational aspects at cluster level should end up in a discussion of the generalisability of the results. Please include assessment of which organisational preconditions that are important for the success of the implementation of the telemedicine service. Similarly, it should include consideration of the extent to which the variation of the organisation in the different pilots have had an impact on the clinical and economic outcomes and the patients perception.

DRAFT

3.6. Domain 7: Socio-cultural, ethical and legal aspects

The pilot reports are expected to include description of the socio-cultural, ethical and legal aspects.

Since the perception of whether a telemedicine service may involve ethical, legal or socio-cultural challenges probably varies between countries, the description of the relevant aspects within this domain is expected to vary between the pilots in each cluster.

The cluster reports are expected to assess the similarities and differences between the pilots with regard to the socio-cultural, ethical and legal aspects and to focus on the following issues:

- Socio-cultural issues:
 - Does the use of the application affect the way in which the patient can work or interact socially?
 - Will the service influence the general model for the delivery of healthcare service if deployed?
 - Are the patients and/or relatives capable of handling their responsibility?
 - Has the service any consequences on the position of gender?
- Ethical issues:
 - Does the use of the application challenge cultural, religious, or moral beliefs?
 - Does the use of the application give new responsibilities to patients?
 - Can patients exercise their patient's rights of privacy and confidentiality adequately?
 - Justice – does the application improve/ decrease access to care?
- Legal issues:
 - Are devices involved that need to be certified as compliant with legal/technical standards?
 - How is data protection and privacy assured?
 - How is appropriate access to data facilitated?
 - How are patients' rights of consent and control in use of the application assured?
 - What are the responsibilities of the different health care professionals involved?

Based on the answers to the issues above and differences between the countries a short discussion of the generalisability of the results from Renewing Health to other countries should be made.

References

Drummond MF, Bloom BS, Carrin G, Hillman AL, Hutchings HC, Knill-Jones RP, Pouvoirville G, Torfs K. Issues in the cross-national assessment of health technology. *International Journal of Technology Assessment in Health Care* 8:4 (1992): pp 671-682.

Kidholm K, Pedersen CD, Jensen LK, Ekeland AG, Bowes A, Flottorp S, Bech M. A model for assessment of telemedicine applications – MAST. *International Journal of Technology Assessment in Health Care*, 28:1, January 2012, 44-51.

Kidholm K (editor), Aletras V, Bech M, Bowes A, Christiansen EK, Dafoulas G, Dyrvig AK, Jensen LK, Anna Kotzeva, Inger Lindberg, Anna-Leena Orsama, Claus Duedal Pedersen, Reinhard Prior, Janne Rasmussen, Panagiotis Stafylas, Siv Söderberg, Petra Wilson. Guideline on analysis and reporting of results from the pilots in Renewing Health. 2013. Internal deliverable in the Renewing Health project. Available at www.renewinghealth.eu.

DRAFT

Appendix 1. Codes for updating stata:

```
. net from www://www.stata-press.com/data/mais
```

```
. net install mais
```

```
. spinst_mais
```

This should enable you to use the codes for meta-analysis. Please note it is necessary to be online while updating stata.

DRAFT