Rare Diseases

ISSN: (Print) 2167-5511 (Online) Journal homepage: http://www.tandfonline.com/loi/krad20

Developing methodology for the creation of clinical practice guidelines for rare diseases: A report from RARE-Bestpractices

Menaka Pai, Alfonso Iorio, Joerg Meerpohl, Domenica Taruscio, Paola Laricchiuta, Pierpaolo Mincarone, Cristina Morciano, Carlo Giacomo Leo, Saverio Sabina, Elie Akl, Shaun Treweek, Benjamin Djulbegovic, Holger Schunemann & Consortium for the RARE-Bestpractices

To cite this article: Menaka Pai, Alfonso Iorio, Joerg Meerpohl, Domenica Taruscio, Paola Laricchiuta, Pierpaolo Mincarone, Cristina Morciano, Carlo Giacomo Leo, Saverio Sabina, Elie Akl, Shaun Treweek, Benjamin Djulbegovic, Holger Schunemann & Consortium for the RARE-Bestpractices (2015) Developing methodology for the creation of clinical practice guidelines for rare diseases: A report from RARE-Bestpractices, Rare Diseases, 3:1, e1058463, DOI: 10.1080/21675511.2015.1058463

To link to this article: http://dx.doi.org/10.1080/21675511.2015.1058463

Rare Diseases 3:1, e1058463; June 1, 2015; Published with license by Taylor & Francis Group, LLC

Developing methodology for the creation of clinical practice guidelines for rare diseases: A report from RARE-Bestpractices

Menaka Pai[1,2,*], Alfonso Iorio[1], Joerg Meerpohl[3], Domenica Taruscio[4], Paola Laricchiuta[4], Pierpaolo Mincarone[5], Cristina Morciano[4], Carlo Giacomo Leo[6,7], Saverio Sabina[8], Elie Akl[8], Shaun Treweek[9], Benjamin Djulbegovic[10], and Holger Schunemann[1,3] for the RARE-Bestpractices Consortium[†]

[1]McMaster University; Hamilton, Canada; [2]Hamilton Regional Laboratory Medicine Program; Hamilton, Canada; [3]Universitaetsklinikum; Freiburg, Germany; [4]National Center for Rare Diseases; Istituto Superiore di Sanità; Rome, Italy; [5]National Research Council; Institute for Research on Population and Social Policies; Rome, Italy; [6]National Research Council; Institute of Clinical Physiology; Lecce, Italy; [7]Tufts Medical Center; Boston, MA USA; [8]American University of Beirut; Beirut, Lebanon; [9]University of Aberdeen; Aberdeen, UK; [10]University of Florida; Gainesville, FL USA

[†]A list of members can be found in the Note section.

Keywords: access, clinical practice guideline, evaluation, evidence-based medicine, health care quality, rare diseases

Rare diseases are a global public health priority; they can cause significant morbidity and mortality, can gravely affect quality of life, and can confer a social and economic burden on families and communities. These conditions are, by their nature, encountered very infrequently by clinicians. Thus, clinical practice guidelines are potentially very helpful in supporting clinical decisions, health policy and resource allocation. The Grading of Recommendations Assessment, Development and Evaluation (GRADE) system is a structured and transparent approach to developing and presenting summaries of evidence, grading its quality, and then transparently interpreting the available evidence to make recommendations in health care. GRADE has been adopted widely. However, its use in creating guidelines for rare diseases – which are often plagued by a paucity of high quality evidence – has not yet been explored. RARE-Bestpractices is a project to create and populate a platform for sharing best practices for management of rare diseases. A major aim of this project is to ensure that European Union countries have the capacity to produce high quality clinical practice guidelines for rare diseases. On February 12, 2013 at the Istituto Superiore di Sanità, in Rome, Italy, the RARE-Bestpractices group held the first of a series of 2 workshops to discuss methodology for creating clinical practice guidelines, and explore issues specific to rare diseases. This paper summarizes key results of the first workshop, and explores how the current GRADE approach might (or might not) work for rare diseases. Avenues for future research are also identified.

*Correspondence to: Menaka Pai; Email: menakapai@mcmaster.ca

Submitted: 01/30/2015; Revised: 04/20/2015; Accepted: 05/29/2015

http://dx.doi.org/10.1080/21675511.2015.1058463

Introduction

The US. National Institutes of Health (NIH) Office of Rare Diseases Research considers a disease to be rare if it has a prevalence of fewer than 200,000 affected individuals in the United States (for a rate of 1 in 1,500 people), while the European Union defines rare diseases (RDs) as those affecting less than 1 in 2,000 people.[1,2] RDs are frequently life-threatening or seriously debilitating conditions, which can cause significant morbidity and mortality, can gravely affect quality of life, and can confer a social and economic burden on families and communities. They often have no satisfactory method of treatment.[2,3] Many are not the focus of research interest, market interest or public health policies. There are thought to be over 7,000 RDs, although the exact number is not known. Together, they affect over 25 million American citizens and 30 million European Union citizens of all ages. This underscores the fact that though individual RDs have a low prevalence, the total number of patients with RDs is large. For these reasons, RDs, taken as a whole, are a global public health priority.

In the European Union, special initiatives have been undertaken to improve research and health services for patients with RDs. A growing number of national centers of expertise have been established as a way to provide highly-specialized and complex care for these patients. However, the centers are named inconsistently; terms used include "centers of expertise," "centers of reference," "centers of excellence," and "network of reference." This heterogeneity reflects the divergent practices adopted by Member States for establishing centers of expertise. Some countries rely on informal processes, where centers are identified based solely on reputation, often in the absence of any system to monitor quality of care.

Other countries have well-established systems and procedures in place for designating centers of expertise and defining their scope, as well as providing mechanisms for quality assurance.[4] Acknowledgment of the pressing need to share knowledge and resources in the area of RDs has led to the establishment of European Reference Networks (ERNs) by several Member States. These are listed, together with existing centers of expertise, for a range of RDs in the Orphanet database (http://www.orpha.net). After ten years of extensive work at the European Union level, a firm basis for the creation and evaluation of ERNs was provided by Directive 2011/24/EU on patients' rights in cross-border healthcare, the Commission Implementing Decision (2014/287/EU): setting out criteria for establishing and evaluating European Reference Networks, and the Commission Delegated Decision 2014/286/EU: setting out criteria and conditions that European Reference Networks must fulfil.[4-7] The former European Union Committee of Experts on Rare Diseases (EUCERD), now the Commission Expert Group on Rare Diseases, contributed significantly to this work (http://ec.europa.eu/health/rare_diseases/expert_group/index_en.htm).

The International Rare Diseases Research Consortium (IRDiRC) is a further cooperative effort by the European Commission and the US. National Institutes of Health. IRDiRC is a network of 40 public and private sector organizations invested in RD research which aims to accelerate medical breakthroughs for people affected by RDs. The Consortium has 2 goals to achieve by 2020, namely to develop the means to diagnose most RDs and to deliver 200 new therapies for RDs. Clinical research is also strengthened by the Rare Diseases Clinical Research Network (RDCRN), a network of consortia conducting research in over 200 RDs.[8] A collaborative effort of the US. Office of Rare Diseases Research (ORDR) and several National Institutes of Health (NIH) Insitutes/Centers, RDCRN has overseen more than 120 studies since 2003, enrolling over 24,000 participants. Data management and coordination is centralized, facilitating epidemiological and survey research, and data sharing. The establishment of a patient Contact Registry – an online system to collect contact and diagnosis information volunteered by patients with RDs – has allowed RDCRN to connect with patients online, so it can offer more individuals access to research participation. ORDR also coordinates the Rare Diseases Human Biospecimens/Biorepositories (RD-HuB), a searchable database of biospecimens collected, stored, and distributed by biorepositories globally.[9] (https://biospecimens.ordr.info.nih.gov/) All specimens are linked to the Genetic and Rare Diseases (GARD) Information Center, which provides specific educational information on RDs to patients and health care providers. (http://rarediseases.info.nih.gov/gard) Another ORDR project on the horizon is the NIH/NCATS Global Rare Diseases Patient Registry Data Repository (GRDR), a web based resource that will securely aggregate and store de-identified patient information from many different RD registries.[10] (https://grdr.ncats.nih.gov/)

Though global efforts to address the problem of RDs are underway, clinical decision making for individual patients with RDs remains complex. Gaining both clinical experience and good evidence takes significantly more time for these conditions than for common diseases. Thus, clinical practice guidelines are potentially very helpful in supporting clinical decision making. Ideally, clinical practice guidelines summarize existing evidence for clinical outcomes, and use clear criteria to establish our confidence in effect estimates. In many methodological approaches, convincing (high quality) evidence often leads to strong treatment recommendations, while unconvincing (low quality) evidence often leads to weaker treatment recommendations. Systematically created guidelines which use an explicit approach are invaluable. They allow clinicians to make critical treatment decisions, and communicate these decisions transparently to patients and colleagues. Rigorous guidelines also allow policymakers to determine whether a given treatment provides value to stakeholders (so resources can be allocated appropriately), and can standardize care within and between centers of expertise.

The Grading of Recommendations Assessment, Development and Evaluation (GRADE) system helps guideline developers present summaries of evidence in a structured way, make transparent judgements about the quality of evidence, and then systematically move toward developing recommendations for rigorous clinical practice guidelines.[11] GRADE has been adopted by over 80 organizations worldwide, including WHO.[12,13] It has evolved to encompass not only guideline development for treatments, but also guideline development for diagnostic tests and strategies. However, the use of GRADE in creating guidelines for rare diseases has not yet been explored systematically and in detail.

RARE-Bestpractices (http://www.rarebestpractices.eu/) is a European Seventh Framework Program project that aims to develop a platform to facilitate information exchange, identifying and spreading best practices for the management of RDs. One of its goals is to address the paucity of rigorous clinical practice guidelines for RDs, and propose methods for guideline creation and dissemination. The authors of this paper are part of RARE-Bestpractices, of GRADE and DECIDE (a separate Seventh Framework Program project that builds on GRADE to develop new ways of presenting research information in guidelines (http://www.decide-collaboration.eu)).[12,13] On February 12, 2013 at the Istituto Superiore di Sanità, in Rome, Italy, the RARE-Bestpractices group held the first of a series of 2 workshops to discuss methodology for creating clinical practice guidelines, and explore issues specific to RDs. This paper summarizes key results of the first workshop. It reviews facilitators of and barriers to clinical care, research and development of clinical practice guidelines in RDs. It also discusses how the current GRADE approach might (or might not) work for these conditions. Avenues for future research are identified as well.

How Are Rare Diseases Different?

The RARE-Bestpractices Guidelines workshop began with an open forum, where participants discussed perceived characteristics that set rare disorders apart from common diseases. These characteristics can act as facilitators of or barriers to clinical care, research and guideline development in RDs (**Table 1**). These areas are clearly interrelated (**Fig. 1**).

An important theme is the general lack of public awareness about RDs. Although these conditions are considered to

Table 1. Facilitators of and barriers to clinical care, research and guideline development in rare diseases (RDs)

Phase	Facilitators	Barriers
Clinical Care	o Though individual RDs have low prevalence, total number of patients with RDs is large o In some jurisdictions, public funds available to ensure access to therapy o Existence of European reference networks, centres of expertise, and patient associations	o Not an actual priority for policymakers, funders o Lack of clinical expertise o Limited therapeutic options o Low availability and/or accessibility o Patient eligibility for treatment, dosing forms and administration guidelines vary worldwide o Patients and health care providers may be willing to accept treatments with greater risk and unclear benefits
Research	o Though individual RDs have a low prevalence, total number of patients with RDs is large o In some jurisdictions, public funds available for research o Existence of European reference networks, centres of expertise, and patient associations	o General lack of public awareness o Not an actual priority for funders, researchers o Perceived lack of clinical equipoise o Patients and health care providers may be willing to accept treatments with greater risk and unclear benefits o Patients and health care providers may be unwilling to accept placebo or comparator treatment o Impossible to calculate relative treatment effects if studies are single-arm (i.e., do not have a comparator) o Cannot control study results for baseline effects, as these are often unknown o Heterogeneity in studies o Aggregating data extremely challenging o Study enrolment difficult o Many RDs do not have clear diagnostic criteria o Patients not registered in databases in reliable, harmonised way o Dearth of epidemiologists and trialists capable of executing creative, methodologically sound studies for RDs
Guideline development	o In some jurisdictions, public funds available for methodologic research in guideline creation (e.g. RARE-Bestpractices) o Increasing uptake of GRADE system to summarize evidence, grade its quality, and transparently interpret it to make clinical recommendations o European Union directive on application of patients' rights in cross border healthcare supports European reference networks, which must have capacity to produce good practice guidelines	o Paucity of published data on RDs (and much of it is low quality) o Often no published evidence at all for critically important outcomes, or for patient values and preferences

be a global health issue, they are not an actual priority for policymakers, funders, and researchers in many countries. Pharmaceutical companies also de-emphasize RDs; thought they may invest significant financial resources in bringing orphan drugs to the market, they often struggle to recover costs through sales of these products.[14] Without Orphan Drug Regulations and cross-sector sponsorship initiatives, there are few economic incentives for the pharmaceutical industry to develop treatments for RDs.[15-19] Thus, patients with RDs generally have limited therapeutic options.

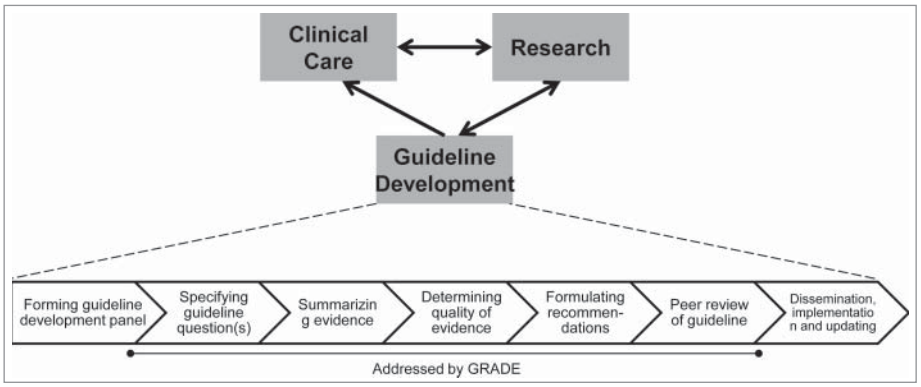

Figure 1. The relationship between clinical care, research and guidelines in rare diseases. Guideline development process adapted from Qaseem A et al. Ann Intern Med. 2012 Apr 3;156(7):525–31.[42]

If agreement is reached that a treatment may have utility in a RD, accessing treatment can still be prohibitive. Different jurisdictions have different regulations to help patients with RDs access treatment. Orphan drugs for RDs often have limited availability and/or accessibility, conditional approval, and partial or no funding coverage. Patient eligibility for treatments, dosing forms and administration guidelines can vary widely from country to country. This makes aggregating data for research purposes extremely challenging.

In the face of these limited therapeutic options, patients and clinicians may have a potentially unreasonable perception or expectation of benefit. That is, they may feel that any treatment is better than no treatment – a perception that dismantles the complex concept of clinical equipoise, and makes it more difficult to conduct properly controlled clinical trials.[20] Patients with RDs are often well informed, aware of the severity of their condition and willing to accept greater risk for possible benefit.[21] Clinicians may feel compelled to offer potentially harmful or ineffective treatments, due to a sense of urgency to treat the RD, as well as pressure from pharmaceutical companies and patients.[22,23] This can create inequity among patients, and restriction of professional autonomy.

Finally, even if there is a will to rigorously study the effects of treatments in RDs, performing properly sized, unbiased and unconfounded studies can be difficult. Enrolment of patients in clinical trials may be hampered by their desire to receive treatment (making placebo options unacceptable). Conversely, it may compel patients to pursue enrolment in very risky trials, so they have a chance at accessing treatment. Specific study designs, like the parallel RCT-cohort trial, sequential design, risk based allocation, hierarchical designs, placebo phase trials, Bayesian designs, and adaptive trials may be well suited to the study of RDs. They are also increasingly acceptable to both the US. Food and Drug Administration (FDA) and the European Medicines Agency (EMA).[24-26] However, these types of trials require specific expertise to design and run. Though pharmaceutical companies and large institutions are increasingly interested in RDs, there is still a dearth of epidemiologists and trialists working in this area who are capable of executing creative, methodologically sound studies. Indeed, widespread use of innovative trial design and analysis approaches is limited because trialists, peer-reviewers, regulatory authorities and other stakeholders are still unfamiliar with these methods, despite the promise they show with regard to efficiency gains in RD trials.[27]

Applying GRADE to rare diseases – the PICO question

GRADE methodology suggests that when framing questions addressing alternative management strategies in systematic reviews or guidelines, one must first specify the patient population, the intervention, the comparator, and the outcomes of interest.[28] This strategy, commonly known as "PICO" (patient, intervention, comparator, outcome), ensures that recommendations are focused and appropriate. When we consider RDs, every part of the PICO question can pose a challenge, as detailed below.

Population

Defining the population of interest for any research endeavor" Also covers: on RDs, including guideline creation, was identified as a major challenge. Many RDs do not have clear diagnostic criteria. Those that do may be difficult to diagnose, because testing strategies are inaccessible, risky or costly. This makes it difficult to identify groups of patients. It might be practical to use broad definitions of the population in RDs (e.g., incorporate closely related disease entities) to potentially increase the amount of data relevant to the PICO question.

Intervention and comparator

GRADE states that PICO questions must define both the intervention and the comparator. Participants noted that there is often only one treatment for any given RD, and in many cases, use of a placebo is not an option due to the severe course of the untreated disease. Thus, finding studies that use an active or inactive comparator can be difficult. Because patterns of practice can vary for RDs, treatments are also frequently underused or not used in a consistent way. It might be practical to use broad definitions of the intervention (e.g., a class of medication) to potentially increase the amount of data relevant to the PICO question.

Outcomes

A key step in creating evidence-based guidelines is determining critical disease outcomes – factors that are important not just to providers and health care systems, but also to patients. There are several challenges in defining relevant outcomes for RDs. First, outcomes relevant to a specific disease are typically characterized (in terms of severity and patterns of recurrence) in long-term prospective studies. Such studies are less likely to be available for RDs, making the selection of relevant outcomes and the determination of their relative importance challenging. Case reports and case series are more likely to present a biased view of any outcomes, as they tend to report unusual and severe events. Second, relying on surrogate outcomes can be more problematic in RDs than non-rare diseases; the pathophysiology and empiric evidence linking them to patient important outcomes are less likely to be well understood. Third, patient reported outcomes, such as disease-specific quality of life instruments, are less likely to be studied and validated in RDs. As a result, the most reliable outcomes might only be those related to morbidity and mortality. There are ongoing initiatives focused on outcome development. One such initiative is COMET (Core Outcome Measures in Effectiveness Trials).[29-32] (http://www.comet-initiative.org) COMET is concerned with the development and application of agreed standardised sets of outcomes, which represent the minimum that should be measured and reported in all clinical trials of a specific condition. These "core outcome sets" are also suitable for use in other study designs.

Applying GRADE to rare diseases – quality of evidence

GRADE requires guideline developers to make an overall rating of confidence in estimates of effect, based on the quality of evidence (high, moderate, low, or very low) for each selectedoutcome critical to the patient. The quality of evidence starts as high

for randomized trials, and low for observational studies. GRADE then assesses 5 factors that decrease confidence in effect estimates: study limitations; inconsistency of results; indirectness of evidence; imprecision; and publication bias.[33] GRADE also assesses 3 factors that increase confidence in effect estimates: large magnitude of effect; plausible confounding which would reduce a demonstrated effect; and a dose-response gradient.[33] An overall rating of confidence is based on the ratings of confidence for each critical outcome. Generally, it is based on the critical outcome that provides the lowest confidence.[34]

GRADE's process to determine confidence in estimates of effect has value, in that it is formal and transparent. However, evidence for most critical outcomes in RDs would be rated "very low," for a number of reasons. Poor data quality ratings may make it challenging to provide useful recommendations for end users of a clinical practice guideline.

Paucity of published data is a serious concern in RDs. Randomized controlled trials and large observational studies (when they do exist) are frequently plagued by study limitations. Guideline developers often have to "settle" for case studies or case series. Case series may still provide moderate or high quality evidence, if they report unique outcomes that are highly specific to the condition of interest. N of 1 trials may also be particularly useful in non-rapidly progressing chronic RDs, or recurrent, symptomatic conditions.[35] In any case, any systematic review of a RD must plan to identify observational data, in addition to trial/experimental data.

Studies of RDs may fail to report important outcomes, typically those for which no effect was observed, or those that capture important harms. Additionally, they may not be powered to detect statistically or clinically significant treatment effects.[36] Recruitment of patients into studies is difficult, as patients with RDs are often not registered in databases in a reliable, harmonised way.[23,37] If very few patients are affected by a given RD, it can be difficult to replicate study findings. There is also little clarity on how to incorporate qualitative research into guidelines, particularly if it is the only type of evidence available for a given RD. Qualitative studies are not part of GRADE's quality assessment of effect estimates. However, they can be used for the process evaluation of complex interventions, and thus may help interpret evidence for their effectiveness. They can also be used to assess patient values and preferences related to the disease, its outcomes, and interventions of interest.

Indirectness is another common problem in studies of RDs. Evidence quality decreases when there are differences between the population, intervention, comparator to the intervention, and outcome of interest, and those included in the relevant studies. It is difficult to carry out large studies in well-defined RD patient populations, so clinicians often use indirect evidence, extrapolating data from a population that shares some features. (For example, data from epilepsy – a common disease – is often used to treat patients with tuberous sclerosis – a rare disease.) Indirectness can even occur within a defined RD population. Many RDs have a clear, common genotype. However, the phenotype can vary widely. Similar indirectness for interventions, comparators and outcomes also occurs in RDs. There is often no evidence on the effect of treatments on critical outcomes in RDs, and use of surrogate outcomes abound. Multiple comparisons, where patients have several treatments administered simultaneously, are also a challenge in RDs. It can be methodologically difficult to interpret the resulting data, and separate out the individual effects of each treatment on an outcome.

Imprecision is a final area of concern that lowers the quality of evidence in RDs. When a study includes relatively few patients who have relatively few events, estimates of effect will have wide confidence intervals. Studies of RDs often do not even include precision estimates. When they are available, small patient and event numbers yield very imprecise estimates of effect. Networks like RDCRN, which uses an automated patient Contact Registry to improve patient access to studies, can potentially address the problem of imprecision, by increasing sample sizes for observational studies.[8] "Meta-registries" like GRDR also hold promise in aggregating information from far more RD patients than any one reference center or registry could access. GRDR also offers opportunities for cross-disease research on RDs that share common features.[10]

Applying GRADE to rare diseases – evidence profiles and summary of findings tables

In the GRADE system, evidence can be summarized in 2 ways: evidence profiles (EPs), and summary of findings (SoF) tables.[38] SoF tables succinctly present aggregate data for each important outcome, including the absolute risk of the outcome for the intervention and the comparator, the relative effect of the intervention, the number of participants and studies considered, a summary of the quality of evidence, and any other comments (**Table 2**). EPs contain similar information, but expand on the quality of evidence; they include an explicit judgment of each factor that determines the quality of evidence for each outcome (**Table 3**). EPs and SoF tables serve different purposes and are intended for different audiences. EPs are a detailed record of the judgments that go into developing a systematic review or guideline, and are intended for authors, reviewers and guideline panels. It is vital for these groups to understand the judgements underlying the overall quality assessments, which are then recorded in the SoF tables. SoF tables are more concise, and are intended for a broader audience, including end users of systematic reviews and guidelines.

EP and SoF tables for a RD may look very different from one for a common disease. Many studies in RDs have no comparator, so establishing relative effects is not possible. Treatment of RDs often considers "baseline effect" (e.g., spontaneous remission) – this is not traditionally captured in EPs and SOFs. Further, for some critically important outcomes in RDs, there is no published evidence at all, yielding "empty" rows.

Applying GRADE to rare diseases– evidence to recommendation tables

The GRADE system asks guideline creators to consider a number of factors when making recommendations, including: the overall quality of the evidence; the balance between benefits and risks; resource use; and patient values and preferences.[39,40]

Table 2. Sample GRADE Evidence Profile for inhaled antibiotics versus intravenous antibiotics for pulmonary exacerbations in cystic fibrosis[43]

Quality assessment							No of patients		Effect			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inhaled antibiotics	Intravenous antibiotics	Relative (95% CI)	Absolute	Quality	Importance
adverse events - renal toxicity												
2	randomised trials	serious[6]	no serious inconsistency	no serious indirectness	no serious imprecision	none	24/24 (100%)	22/22 (100%)	—	1000 fewer per 1000 (from 1000 fewer to 1000 fewer)	ÅÅÅOMODERATE	CRITICAL
lung function - FEV1[1] (measured with: litre or percent (%) predicted: absolute values or change values; Better indicated by lower values)												
2	randomised trials[2]	serious	no serious inconsistency	no serious indirectness	no serious imprecision	none	24	22	—	mean ranged from 0 to 0 higher	ÅÅOOLOW	CRITICAL
need for hospital admission (follow-up 14 days)												
1	randomised trials	serious[3]	no serious inconsistency	no serious indirectness	serious	none	2/16 (12.5%)	1/12 (8.3%)	RR 1.50 (0.15 to 14.68)	42 more per 1000 (from 71 fewer to 1000 more)		
								0%		—		
time off work or school - not reported												
0	—	—	—	—	—	none	0	—	—	—		IMPORTANT
quality of life - not reported												
0	—	—	—	—	—	none	0	—	—	—		
need for additional IV antibiotics												
1	randomised trials	serious[4]	no serious inconsistency	no serious indirectness	serious[5]	none	2/8 (25%)	0/10 (0%)	RR 6.11 (0.33 to 111.71)	—	ÅÅOOLOW	IMPORTANT
								0%		—		
								0%		—		
development of resistant organisms (follow-up 14 days)												
1	randomised trials[2]	serious[7]	no serious inconsistency	no serious indirectness	serious[8]	none[9]	3/16 (18.8%)	1/12 (8.3%)	RR 2.25 (0.27 to 19.04)	104 more per 1000 (from 61 fewer to 1000 more)	ÅÅOOLOW	IMPORTANT
								0%		—		
time to next pulmonary exacerbation[10] (follow-up 2 weeks; Better indicated by lower values)												
1	randomised trials	serious[11]	no serious inconsistency	no serious indirectness	no serious imprecision	none	0	—	—	MD 0 higher (0 to 0 higher)[12]	ÅÅÅOMODERATE	IMPORTANT

[1] there was insufficient information for meta-analysis: one study (Cooper) listed Before and After: Intervention B:42%, A:55% Control:B:39%, A:52% one study (Stephens) listed change in FEV1: Intervention: 6.7%, Control: 3.9% -> both reported no significant difference.
[2] 2 studies reported on this outcome: one reported RR and one only stated that in the inhaled group, 3 out of 39 strains were resistant, in the IV group 4 resistances (not listed in this evidence profile).
[3] unclear risk of random sequence generation, allocation concealment, no blinding, unclear risk of incomplete outcome data.
[4] unclear risk of random sequence generation, unclear risk of allocation concealment, no blinding etc.
[5] very wide CI.
[6] no blinding etc.
[7] no blinding etc. (see previous footnotes).
[8] very wide CI
[9] outcome was stated to be measured in 3 trials but reported in only 2 trials.
[10] single study in abstract form.
[11] unclear risk of random sequence generation, unclear risk of allocation concealment, no blinding.
[12] no numbers given: time to next exacerbation maximal in the once-daily inhaled antibiotic group, less in the twice daily IV antibiotic group, shortest time: group receiving IV antibiotics three-times daily.

Table 3. Sample GRADE Summary of Findings Table for oral oseltamivir versus no antiviral therapy in influenza[44]

Quality assessment		Summary of Findings: Study event rates (%)			Anticipated absolute effects Follow up	
Participants (studies)	**Overall quality of evidence**	**With no antiviral treatment**	**With oseltamivir**	**Relative effect (95% CI)**	**Risk with no antiviral treatment**	**Absolute effect with Oseltamivir (95% CI)**
			Mortality			
681 (3 studies)	⊕⊕⊖⊖ LOW[1]	59/242 (24.4%)	31/439 (7.1%)	Adjusted OR 0.23 (0.13 to 0.43)	240 deaths per 1000	172 fewer deaths per 1000 (from 120 to 201 fewer)
Hospitalisation						
150710 (4 studies)	⊕⊕⊖⊖ LOW[4]	1238/ 100585 (1.2%)	431/50125 (0.86%)	Adjusted OR 0.75 (0.66 to 0.89)	12 hospitalisations per 1000	3 fewer hospitalisations per 1000 (from 1 to 4 fewer)
			ICU admissions/mechanical ventilation/respiratory failure			
1032 (6 studies[5])	⊕⊖⊖⊖ VERY LOW[1,6] due to risk of bias, inconsistency	—	200/1032 (19.4%) Pooled Risk 13.0% (95% CI 11 to 15%)	—	—	
			Duration of hospitalisation (days)			
832 (5 studies)	⊕⊖⊖⊖ VERY LOW[1,5,6] due to risk of bias, inconsistency	—	832	—	—	The mean duration of hospital stay was 5.16 days (5.02 to 5.29)
			Duration of signs and symptoms (measured from onset of symptoms or treatment) (time to return to normal activity was not measured)			
5842 (6 studies)	⊕⊖⊖⊖ VERY LOW [1,6] due to inconsistency	449	5393	—	—	The mean time was 0.91 standard deviations lower (1.25 to 0.57 lower)[7]
			Complications - Pneumonia			
150466 (3 studies)	⊕⊖⊖⊖ VERY LOW[4,6] due to inconsistency	2111/ 100449 (2.1%)	647/50017 (1.3%)	Adjusted OR 0.83 (0.59 to 1.16)	21 pneumonias per 1000	4 fewer pneumonias per 1000 (from 9 fewer to 3 more)
			Critical Adverse Events			
104930 (5 studies)	⊕⊕⊖⊖ LOW[4]	60817	44113	Rate Ratio 0.76 (0.7 to 0.81)	420 adverse events per 1000 patient years	101 fewer adverse events per 1000 patient years (from 80 to 126 fewer)

[1] Although we did not downgrade, publication bias cannot be excluded.; [2] Studies not adjusted for potential confounding factors.; [3] Significant differences in effect for pandemic versus seasonal influenza (see subgroup analyses table).; [4] Publication bias a concern since large studies had for-profit funding and weighted heavily in analyses.; [5] No independent comparison group.; [6] High heterogeneity among studies.; [7] This translates to reduced symptom duration of approximately 33 hours (95% CI 21 to 45 hours). Despite the large effect we did not upgrade because there was important inconsistency across studies.

GRADE formalizes the evaluation of these factors using the Evidence to Recommendation (EtR) framework.[41] It can be challenging to move from evidence to recommendations, in light of the lack of high quality evidence for RDs. End users may not find a guideline useful if it summarizes evidence, but ultimately cannot make recommendations for or against any treatment. For RDs, where a bedrock of high quality evidence is not always available, it is important that guideline panels not resort to making no recommendation at all; a more pragmatic approach is to provide some guidance to end users, even if it is in the form of weak recommendations. Uncertainty about an intervention's benefit in a given RD may also be a valid reason for a panel to recommend against it, particularly if there is more compelling evidence of harm. Useful recommendations can also be made for RDs when considering treatment of symptoms (e.g., pain). The needed evidence for these recommendations may be derived from studies of symptomatic treatment in the setting of comparable non-rare diseases. Useful recommendations can similarly be made for complications of RDs that are themselves non-rare (e.g., chronic kidney disease, congestive heart failure); moderate or high quality evidence likely exists for these complications. RD guideline panels can also opt to make "recommendations for research" – recommendations to only use treatments as part of an evaluative process which collects information about benefits, adverse events, and patient values and preferences. This not only provides end users with helpful guidance, it encourages them to add to the body of knowledge. The need for published research on patient values and preferences, partnerships with patient groups, and representation of patients on guideline panels – shortcomings in both rare and non-rare disease – should also be identified.

Conclusions

Rigorous clinical practice guidelines are needed to improve the care of the millions of people worldwide who suffer from RDs, and fuel the work of Reference Networks and Centers of Expertise. The first RARE-Bestpractices Workshop identified key features that set RDs apart from common diseases, and make them more challenging to study in a rigorous way. We must keep these features in mind as we create methodology suitable to develop and update best practice guidelines in RDs. The application of a RD perspective may be relevant to genomic/genetic testing and, ultimately, may be applicable to the entire field of "personalized medicine."

Members of GRADE and RARE-Bestpractices will continue to work together to further test and adapt the GRADE system for the creation of guidelines for RDs. The goal is to formalize a systematic process that can be used to create evidence-based practice guidelines that are useful to patients, clinicians, researchers, industry, and policy makers in the RD community.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Funding

The research leading to these results has received funding from the (FP7/2007–2013), under grant agreement n° 305690 (RARE-Bestpractices project). The opinions, presented here reflect only the authors' views. The European Union is not liable for any, use that may be made of the information contained herein.

Note

Members of the RARE-Bestpractices Consortium: Domenica Taruscio (Istituto Superiore di Sanità), Cristina Morciano (Istituto Superiore di Sanità), Paola Laricchiuta (Istituto Superiore di Sanità), Joanne Auld (JAMARAU), Thomas Sejersen (Karolinska Institutet), Karen Ritchie (Healthcare Improvement Scotland), Michele Hilton Boon (Healthcare Improvement Scotland), Jan Manson (Healthcare Improvement Scotland), Panos G. Kanavos (London School of Economics and Political Science), David Tordrup (London School of Economics and Political Science), Elena Nicod (London School of Economics and Political Science), Pierpaolo Mincarone (National Research Council), Carlo Giacomo Leo (National Research Council), Saverio Sabina (National Research Council), Monica Ensini (EURORDIS, European Organization for Rare Diseases), Juliette Senecat (EURORDIS, European Organization for Rare Diseases), Graziella Filippini (Associazione per la Ricerca sull'Efficacia dell'Assistenza Sanitaria Centro Cochrane Italiano), Slivia Minozzi (Associazione per la Ricerca sull'Efficacia dell'Assistenza Sanitaria Centro Cochrane Italiano), Cinzia Del Giovane (Associazione per la Ricerca sull'Efficacia dell'Assistenza Sanitaria Centro Cochrane Italiano), Holger Schünemann (Universitaetsklinikum Freiburg), Joerg Meerpohl (Universitaetsklinikum Freiburg), Tsonka Miteva-Katrandzhieva (Bulgarian Association for Promotion of Education and Science), Rumen Stefanov (Bulgarian Association for Promotion of Education and Science), Georgi Iskrov (Bulgarian Association for Promotion of Education and Science), Pedro Serrano-Aguilar (Fundación Canaria de Investigación y Salud), Lilisbeth Perestelo-Perez (Fundación Canaria de Investigación y Salud), Maria M. Trujillo-Martín (Fundación Canaria de Investigación y Salud), Angela Brand (Universiteit Maastricht), Kate Bushby (Newcastle University Upon Tyne), Liesbeth Siderius (The European Academy of Paediatrics AISBL), Jose Ramet (The European Academy of Paediatrics AISBL), Manuel Posada (Istituto de Salud Carlos III), Menaka Pai (McMaster University), and Alfonso Iorio (McMaster University).

Authors' Contributions

MP, HJS, JM and AI drafted the manuscript. All authors contributed to the conception, design and revision of the manuscript. All authors read and approved the final manuscript.

References

1. Frequently Asked Questions – Office of Rare Diseases Research. National Institutes of Health.
2. Rare Diseases – Health-EU. European Union
3. Rare Diseases: understanding this Public Health Priority. Paris, France: EURORDIS (European Organization for Rare Diseases), 2005.
4. Palm W, Glinos IA, Rechel B, Garel P, Busse R, Figueras J. Building European Reference Networks in Health Care: Exploring concepts and national practices in the European Union. Copenhagen: World Health Organization, 2013; 83.
5. Montserrat MA, Waligóra J. The European union policy in the field of rare diseases. Public Health Genomics 2012; 16:268-77; PMID:24503587; http://dx.doi.org/oi.org/10.1159/000355930
6. Peeters M. Free movement of patients: directive 2011/24 on the application of patients' rights in cross-border healthcare. Eur J health Law 2012; 19:29-60; PMID:22428388; http://dx.doi.org/10.1163/157180912X615158
7. Abbing HDR. Eu cross-border healthcare and health law. Eur J Health Law 2015; 22:1-12; http://dx.doi.org/10.1163/15718093-12341348
8. Krischer JP, Gopal-Srivastava R, Groft SC, Eckstein DJ; Rare Diseases Clinical Research Network. The Rare Diseases Clinical Research Network's organization and approach to observational research and health outcomes research. J Gen Intern Med 2014; 29 Suppl 3: S739-44; PMID:25029976; http://dx.doi.org/10.1007/s11606-014-2894-x
9. Rubinstein YR, Groft SC, Bartek R, Brown K, Christensen RA, Collier E, Farber A, Farmer J, Ferguson JH, Forrest CB, et al. Creating a global rare disease patient registry linked to a rare diseases biorepository database: Rare Disease-HUB (RD-HUB). Contemp Clin Trials 2010; 31:394-404; PMID:20609392; http://dx.doi.org/10.1016/j.cct.2010.06.007
10. Forrest CB, Bartek RJ, Rubinstein Y, Groft SC. The case for a global rare-diseases registry. Lancet 2011; 377:1057-9; PMID:20674966; http://dx.doi.org/10.1016/S0140-6736(10)60680-0
11. Terracciano L, Brozek J, Compalati E, Schünemann H. GRADE system: new paradigm. Curr Opin Allergy Clin Immunol 2010; 10:377-83; PMID:20610980; http://dx.doi.org/10.1097/ACI.0b013e32833c148b
12. Oxman A. DECIDE: Developing and Evaluating Communication Strategies to Support Informed Decisions and Practices Based on Evidence.: European Commission under the Seventh Framework Programme, 2011-2015.
13. RARE-Bestpractices: Platform for sharing best practices for management of rare diseases. RARE-Bestpractices
14. Haffner ME. Adopting orphan drugs–two dozen years of treating rare diseases. N Engl J Med 2006; 354:445-7; PMID:16452556; http://dx.doi.org/10.1056/NEJMp058317
15. Fiorentino R, Liu G, Pariser AR, Mulberg AE. Cross-sector sponsorship of research in eosinophilic esophagitis: a collaborative model for rational drug development in rare diseases. J Allergy Clin Immunol 2012; 130:613-6; PMID:22857796; http://dx.doi.org/10.1016/j.jaci.2012.07.011
16. Dunoyer M. Accelerating access to treatments for rare diseases. Nat Rev Drug Discov 2011; 10: 475-6; PMID:21701499; http://dx.doi.org/10.1038/nrd3493
17. Oprhan Medicinal Products. European Commission
18. Orphan Drug Act. Food and Drug Administration, US. Department of Health and Human Services.
19. European Parliament and Council Regulation (EC) No 141/2000 of 16 December 1999 on orphan medicinal products. In: Council EPa, ed., 1999
20. Robinson EJ, Kerr CE, Stevens AJ, Lilford RJ, Braunholtz DA, Edwards SJ, Beck SR, Rowley MG. Lay public's understanding of equipoise and randomisation in randomised controlled trials. Health Technol Assess 2005; 9:1-192, iii-iv; PMID:15763039
21. Kahneman D, Tversky A. The psychology of preferences. Sci Am 1982; 246:160-73; http://dx.doi.org/10.1038/scientificamerican0182-160
22. Aartsma-Rus A. The risks of therapeutic misconception and individual patient (n=1) "trials" in rare diseases such as Duchenne dystrophy. Neuromuscul Disord: NMD 2011; 21:13-5; PMID:21051233; http://dx.doi.org/10.1016/j.nmd.2010.09.012
23. Groft SC, de la Paz MP. Rare diseases – avoiding misperceptions and establishing realities: the need for reliable epidemiological data. Adv Exp Med Biol 2010; 686:3-14; PMID:20824436
24. Elsasser A, Regnstrom J, Vetter T, Koenig F, Hemmings RJ, Greco M, Papaluca-Amati M, Posch M. Adaptive clinical trial designs for European marketing authorization: a survey of scientific advice letters from the European Medicines Agency. Trials 2014; 15:383; PMID:25278265; http://dx.doi.org/10.1186/1745-6215-15-383
25. Cook T, DeMets DL. Review of draft FDA adaptive design guidance. J Biopharm Stat 2010; 20:1132-42; PMID:21058109; http://dx.doi.org/10.1080/10543406.2010.514455
26. Chow SC, Corey R. Benefits, challenges and obstacles of adaptive clinical trial designs. Orphanet J Rare Dis 2011; 6:79; PMID:22129361; http://dx.doi.org/10.1186/1750-1172-6-79
27. Gupta S, Faughnan ME, Tomlinson GA, Bayoumi AM. A framework for applying unfamiliar trial designs in studies of rare diseases. J Clin Epidemiol 2011; 64:1085-94; PMID:21530171; http://dx.doi.org/10.1016/j.jclinepi.2010.12.019
28. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, Alderson P, Glasziou P, Falck-Ytter Y, Schünemann HJ. GRADE guidelines: 2. Framing the question and deciding on important outcomes. J Clin Epidemiol 2011; 64:395-400.
29. Gargon E, Gurung B, Medley N, Altman DG, Blazeby JM, Clarke M, Williamson PR. Choosing important health outcomes for comparative effectiveness research: a systematic review. PLoS One 2014; 9:e99111; PMID:24932522; http://dx.doi.org/10.1371/journal.pone.0099111
30. Gargon E, Williamson PR, Altman DG, Blazeby JM, Clarke M. The COMET Initiative database: progress and activities from 2011 to 2013. Trials 2014; 15:279; PMID:25012001; http://dx.doi.org/10.1186/1745-6215-15-279
31. Prinsen CA, Vohra S, Rose MR, King-Jones S, Ishaque S, Bhaloo Z, Adams D, Terwee CB. Core Outcome Measures in Effectiveness Trials (COMET) initiative: protocol for an international Delphi study to achieve consensus on how to select outcome measurement instruments for outcomes included in a 'core outcome set'. Trials 2014; 15:247; PMID:24962012; http://dx.doi.org/10.1186/1745-6215-15-247
32. Williamson P, Clarke M. The COMET (Core Outcome Measures in Effectiveness Trials) Initiative: Its Role in Improving Cochrane Reviews. Cochrane Database Syst Rev 2012; 5:ED000041; PMID:22592744
33. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schünemann HJ; GRADE Working Group. What is "quality of evidence" and why is it important to clinicians? BMJ 2008; 336:995-8; PMID:18456631
34. Guyatt G, Oxman AD, Sultan S, Brozek J, Glasziou P, Alonso-Coello P, Atkins D, Kunz R, Montori V, Jaeschke R, et al. GRADE guidelines: 11. Making an overall rating of confidence in effect estimates for a single outcome and for all outcomes. J Clin Epidemiol 2013; 66:151-7; PMID:22542023
35. Gabler NB, Duan N, Vohra S, Kravitz RL. N-of-1 trials in the medical literature: a systematic review. Med Care 2011; 49:761-8; PMID:21478771; http://dx.doi.org/10.1097/MLR.0b013e318215d90d
36. Lilford RJ, Thornton JG, Braunholtz D. Clinical trials and rare diseases: a way out of a conundrum. BMJ 1995; 311:1621-5; PMID:8555809; http://dx.doi.org/10.1136/bmj.311.7020.1621
37. Luisetti M, Campo I, Scabini R, Zorzetto M, Kadija Z, Mariani F, Ferrarotti I. The problems of clinical trials and registries in rare diseases. Respir Med 2010; 104 Suppl 1:S42-4; PMID:20413285; http://dx.doi.org/10.1016/j.rmed.2010.03.016
38. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, Norris S, Falck-Ytter Y, Glasziou P, DeBeer H, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011; 64:383-94; PMID:21195583; http://dx.doi.org/10.1016/j.jclinepi.2010.04.026
39. Andrews JC, Schunemann HJ, Oxman AD, Pottie K, Meerpohl JJ, Coello PA, Rind D, Montori VM, Brito JP, Norris S, et al. GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. J Clin Epidemiol 2013; 66:726-35; PMID:23570745; http://dx.doi.org/10.1016/j.jclinepi.2013.02.003
40. Andrews J, Guyatt G, Oxman AD, Pottie K, Meerpohl JJ, Coello PA, Rind D, Montori VM, Brito JP, Norris S, et al. GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. J Clin Epidemiol 2013; 66:719-25; PMID:23312392; http://dx.doi.org/10.1016/j.jclinepi.2012.03.013
41. Brozek JL, Akl EA, Compalati E, Kreis J, Terracciano L, Fiocchi A, Ueffing E, Andrews J, Alonso-Coello P, Meerpohl JJ, et al. Grading quality of evidence and strength of recommendations in clinical practice guidelines part 3 of 3. The GRADE approach to developing recommendations. Allergy 2011; 66:588-95; PMID:21241318; http://dx.doi.org/10.1111/j.1398-9995.2010.02530.x
42. Qaseem A, Forland F, Macbeth F, Ollenschläger G, Phillips S, van der Wees P, Board of Trustees of the Guidelines International Network. Guidelines International Network: toward international standards for clinical practice guidelines. Ann Intern Med 2012; 156:525-31; PMID:22473437; http://dx.doi.org/10.7326/0003-4819-156-7-201204030-00009
43. Ryan G, Jahnke N, Remmington T. Inhaled antibiotics for pulmonary exacerbations in cystic fibrosis. Cochrane Database syst Rev 2012; 12:CD008319; PMID:23235659
44. Hsu J, Santesso N, Mustafa R, Brozek J, Chen YL, Hopkins JP, Cheung A, Hovhannisyan G, Ivanova L, Flottorp SA, et al. Antivirals for treatment of influenza: a systematic review and meta-analysis of observational studies. Ann Intern Med 2012; 156:512-24; PMID:22371849; http://dx.doi.org/ 10.7326/0003-4819-156-7-201204030-00411