2019 EULAR points to consider for the assessment of competences in rheumatology specialty training

Francisca Sivera ,[1,2] Alessia Alunno ,[3] Aurélie Najm ,[4,5] Tadej Avcin,[6] Xenofon Baraliakos ,[7,8] Johannes W Bijlsma,[9] Sara Badreh,[10] Gerd Burmester,[11] Nada Cikes,[12] Jose AP Da Silva ,[13,14] Nemanja Damjanov,[15] Maxime Dougados,[16] Jean Dudler,[17] Christopher J Edwards,[18] Annamaria Iagnocco,[19] Frédéric Lioté,[20,21] Elena Nikiphorou ,[22] Marloes van Onna,[23,24] Simon R Stones,[25] Dimitrios Vassilopoulos,[26] Catherine Haines,[27] Sofia Ramiro [28,29]

Handling editor Josef S Smolen

For numbered affiliations see end of article.

Correspondence to
Dr Francisca Sivera, Dpt Clinical Medicine, Miguel Hernandez University of Elche, 03550 Elche, Valenciana, Spain; fransimas@yahoo.es

Received 20 May 2020
Revised 23 June 2020
Accepted 14 July 2020

ABSTRACT

Background and aim Striving for harmonisation of specialty training and excellence of care in rheumatology, the European League Against Rheumatism (EULAR) established a task force to develop points to consider (PtCs) for the assessment of competences during rheumatology specialty training.

Methods A systematic literature review on the performance of methods for the assessment of competences in rheumatology specialty training was conducted. This was followed by focus groups in five selected countries to gather information on assessment practices and priorities. Combining the collected evidence with expert opinion, the PtCs were formulated by the multidisciplinary task force, including rheumatologists, medical educationalists, and people with rheumatic and musculoskeletal diseases. The level of agreement (LoA) for each PtC was anonymously voted online.

Results Four overarching principles and 10 PtCs were formulated. The overarching principles highlighted the importance of assessments being closely linked to the rheumatology training programme and protecting sufficient time and resources to ensure effective implementation. In the PtCs, two were related to overall assessment strategy (PtCs 1 and 5); three focused on formative assessment and portfolio (PtCs 2–4); three focused on the assessment of knowledge, skills or professionalism (PtCs 6–8); one focused on trainees at risk of failure (PtC 9); and one focused on training the trainers (PtC 10). The LoA (0–10) ranged from 8.75 to 9.9.

Conclusion These EULAR PtCs provide European guidance on assessment methods throughout rheumatology training programmes. These can be used to benchmark current practices and to develop future strategies, thereby fostering continuous improvement in rheumatology learning and, ultimately, in patient care.

To cite: Sivera F, Alunno A, Najm A, *et al*. *Ann Rheum Dis* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/annrheumdis-2020-218015

INTRODUCTION

Rheumatology specialty training is the educational process required for a physician to formally become a specialist in rheumatology. It is defined by an officially approved training programme which aims to bring physicians to an agreed standard of proficiency regarding the management of people with rheumatic and musculoskeletal diseases (RMDs). The definition of the aims, structure and contents of each country's rheumatology training programme is under the exclusive domain of national authorities. However, the harmonisation of specialist training in Europe is deemed essential to ensure equity of access to high standards of care for all people with RMDs and to support the movement of rheumatology specialists across countries.[1] Available data on training programmes in Europe show a wide heterogeneity on their length, structure and content.[2 3]

For decades, educationalists have highlighted the relationship between learning and assessment.[4] Indeed, learning is often driven by assessment.[5] Assessment during training has apowerful impact on motivating learners on their path towards assessment for certification purposes. Regular and repeated testing can increase the retention of knowledge[6] and the skill performance[7] in undergraduate medical students. Even though evidence is scarce, the same paradigm is thought to apply to other types of assessment within higher education, such as specialty medical training.

The aim of this task force was to develop European League Against Rheumatism (EULAR) points to consider (PtCs) for the assessment of competences during rheumatology specialty training with the broader goals of enhancing learning during rheumatology specialty training, contributing to the harmonisation of training outcomes across Europe and improving the care provided to people with RMDs.

METHODS

After approval by the EULAR Executive Committee, the convenor (FS) and the methodologists (CH and SR) led a multidisciplinary task force guided by the 2014 updated EULAR standardised operating procedures.[8] The task force consisted of 23 members, including rheumatologists with an interest in medical education (two of them also representing the Emerging Eular Network), a methodologist, a medical educationalist, and two people with RMDs, from 12 different countries. Two face-to-face meetings of the task force were held in November 2018 and October 2019. Two fellows (AA and AN), guided by the methodologists, performed a systematic literature review (SLR), retrieving individual studies on methods of assessment in rheumatology specialty training and SLRs of studies from other

Table 1 Glossary of terms related to the assessment of competences

Term	Definition
Assessment	A systematic process of gathering and analysing information on competences in order to measure a learner's achievement
CanMEDS framework	The most widely accepted and applied physician competency framework in the world, using a framework to explicit the knowledge, skills and behaviours associated with specific competences across seven roles: medical expert, professional, communicator, health advocate, collaborator, scholar and leader
Competence	An observable ability of a physician related to a specific ability that integrates knowledge, skills and behaviours and that develops through the stages of expertise from novice to master clinician
Curriculum	The description of the outcomes required, and the activities and experience prescribed to develop and demonstrate those outcomes
Direct observation of procedural skills (DOPS)	A workplace-based assessment to evaluate the competence in performing a required technical skill.
Feedback	A process whereby an individual is given information about their performance in order to help them learn and progress
Formative assessment	Information about a learner's performance or understanding, which is provided to the learner as part of the learning process so that they are stimulated to improve their performance and progress towards the required level of competence
Mini clinical examination (mini-CEX)	A workplace-based assessment to evaluate how effectively a clinician interacts with a patient
Multisource feedback	A system that collects the anonymous appraisal of the trainee's performance in an everyday clinical setting, by a variety of coworkers, from mentors to colleagues, nurses and patients; this tool is especially valuable to address issues related to professionalism
Objective structured clinical evaluation (OSCE)	A carefully designed examination circuit of different time-limited stations, each dedicated to the assessment of performance at a particular simulated task
Portfolio	A repository for multiple formative assessments, reflections and records of achievements
Professionalism	A set of values, behaviours and relationships that underpins the trust that the public has in doctors; as professionals, physicians are committed to the health and well-being of individual patients and society through ethical practice, high personal standards of behaviour, accountability to the profession and society, physician-led regulation and maintenance of personal health
Summative assessment	A measure of a learner's performance or understanding which sums up and grades whether the learner has succeeded in reaching the required level of competence

related medical specialties.[9] As published evidence on assessment methods was limited, a qualitative study using focus groups in five European countries (Denmark, Netherlands, Slovenia, Spain and UK) gathered insights into current practices and priorities.[10] These countries were selected to provide a representation of different assessment structures and cultures. The SLR and qualitative study are published separately; however, they form an integral part of the project.

Based on the presented evidence and expert opinion, and following a process of iterative discussion, the overarching principles and PtCs were developed across two 1-day task force meetings. For every statement, formulations were presented, discussed and voted on (informal voting). The statements were accepted if at least 75% of the task force approved the wording in the first round. If this was not reached, further discussion ensued and wording was refined. At least a 67% approval rate was required in the second voting round. If a third voting round was necessary, a simple majority was sufficient. Prompted by discussions during the meetings, the task force felt the need to develop a glossary (table 1) in order to standardise terminology.

After the meeting and once the PtCs were finalised, the level of evidence supporting each statement and the grade of recommendation was assigned following the Oxford Centre for Evidence-Based Medicine procedures.[11] Finally, each task force member anonymously indicated their level of agreement (LoA) with each PtC online (numerical rating scale ranging from 0='do not agree at all' to 10='fully agree'). In addition, based on the limited nature of the available evidence and the issues raised among the task force, a research agenda was formulated.

The final manuscript was reviewed and approved by all task force members, followed by ratification by the EULAR Executive Committee and the Rheumatology Section and Board of the European Union of Medical Specialists (UEMS).

RESULTS

Four overarching principles and 10 PtCs were formulated (table 2). Overall, the evidence underpinning these PtCs in the rheumatology setting is scarce, so the emphasis was placed on general expertise and consensus.

Overarching principles

Rheumatology training should generate rheumatologists capable of and committed to delivering the best care to people with RMDs

During rheumatology training, the physician should acquire the knowledge, skills and professional behaviours necessary to ensure delivery of optimal care to people with all types of RMDs throughout their careers.

Assessment of competences is vital to guide learning and to guarantee quality of care

In the past decades, there has been a move towards 'assessment for learning', in which the assessment environment encourages trainees to feel responsible for driving and appraising their own learning.[12]

Assessment is an integral part of training and must be guided by and aligned with a clear set of educational objectives established by the curriculum

The task force agreed on the need for assessments to be embedded into a structured strategy conveyed by the overall training programme. The curriculum provides the framework of learning objectives, each corresponding to adequate methods of teaching, learning and assessment. National curricula are available in most countries. Additionally, the UEMS Rheumatology Section and Board provides a European curriculum.[13]

Effective assessment requires protected time and resources

One of the key barriers to adequate assessment, identified by trainees and trainers alike throughout Europe, is the lack of protected time for this purpose.[10] In order to improve the clinical learning environment, it is essential that educational supervisors, programme directors and national authorities recognise this need and identify and provide the necessary resources.[14]

Table 2 Overarching principles and points to consider for the assessment of competences in rheumatology specialty training, with LoA and for the specific points, levels of evidence

Overarching principles	LoA, mean (SD)	
1. Rheumatology training should generate rheumatologists capable and committed to deliver the best of care to people with rheumatic and musculoskeletal diseases.	9.9 (0.45), 100%≥8	
2. Assessment as an integral part of training must be guided by and aligned with a clear set of educational objectives established by an official/national/accepted curriculum.	9.8 (0.52), 100%≥8	
3. Assessment of competences is vital to guide learning and to guarantee quality of care.	9.8 (0.41), 100%≥8	
4. Effective assessment requires protected time and resources.	9.7 (0.73), 95%≥8	
Points to consider	**LoE**	**LoA, mean (SD)**
1. Assessment of competences should be a structured and continuous process regularly carried out throughout the training period.	5	9.75 (0.55), 100%≥8
2. Formative assessment with constructive feedback should be frequently performed and with a greater frequency than summative assessment.	5	9.4 (0.82), 100%≥8
3. Feedback should aim to stimulate reflections by the trainee on how to achieve standards of competence and professional behaviour.	5	9.65 (0.67), 100%≥8
4. Trainees should maintain an updated portfolio, including feedback and evidence of self-reflection, to be used as part of the assessment process.	5	9.4 (0.75), 100%≥8
5. Different methods of assessment should be carried out throughout training as multiple methods of assessment can provide a complete overview of a trainee's competence.	5	9.75 (0.64), 100%≥8
6. Multiple-choice case-based questions should be the preferred form of knowledge assessment.	5	8.75 (1.83), 75%≥8
7. Clinical skills should be assessed either in the workplace (direct observation of procedural skills or the mini-clinical examination exercise) and/or in a simulated context (observational structured clinical examination)	5	9.35 (0.81), 100%≥8
8. Competences related to professionalism should be formally assessed using multisource feedback/360° method.	5	9.25 (0.97), 95%≥8
9. The training programme should incorporate predefined processes to identify and support trainees at risk of failure.	5	9.6 (0.75), 100%≥8
10. Trainers should receive continuous training in assessment methods and strategies, particularly in providing constructive feedback.	5	9.4 (1.23), 95%≥8

Numbers in the column 'LoA' indicate the mean (SD) of the LoA and the percentage of task force members with an LoA of at least 8 (0–10). LoE: based on the Oxford Centre for Evidence-Based Medicine classification, with 'level 1' corresponding to meta-analysis or RCTs or high-quality RCTs; 'level 2' corresponding to lesser quality RCTs or prospective comparative studies; 'level 3' corresponding to case–control studies or retrospective studies; 'level 4' corresponding to case series without the use of comparison or control groups; and 'level 5' corresponding to case reports or expert opinion.[11]
LoA, level of agreement; LoE, level of evidence; RCT, randomised controlled trial.

Points to consider

Assessment of competences should be a structured and continuous process, regularly carried out throughout the training period

Assessments should not be performed at a unique time point (eg, final examination); rather, they should be spaced out throughout the training, allowing the trainee to identify areas for improvement before a final summative assessment. Providing a specific recommendation on minimum or optimal assessment frequency was discussed in depth by the task force; however, it was thought that this needed to be flexible enough to be applied in different national contexts. Frequency should be enough to provide adequate feedback and to guide learning throughout the training programme. Some types of assessment, such as the identification of unprofessional behaviours with appropriate feedback, should take place continuously.

Formative assessment with constructive feedback should be regularly performed and with a greater frequency than summative assessment

Assessments can be performed with a formative or a summative aim. Summative assessment assigns grades to trainee performance at designated points in the curriculum, allowing comparison with established standards or between trainees, and a pass/fail decision. For example, an examination at the end of medical training, on which the decision to qualify for medical practice hinges, is a summative assessment. On the other hand, formative assessment is designed as an ongoing part of the instructional process to support and enhance learning and reflection. Formative assessment aims to stimulate the trainee to identify areas for improvement and to provide a plan to that purpose. Frequent, high-quality discussions about current performance, together with expert and customised suggestions for improvement, are associated with more effective learning and higher satisfaction in trainees.[15]

Feedback should stimulate reflection by the trainee on how to achieve the standards of competence and professional behaviour

Feedback is a core component of effective assessment, informing trainees of their progress (or lack of), observed learning needs (and available resources to facilitate learning) and providing motivation to undertake appropriate learning activities.[16] Feedback has the potential to change physicians' behaviour in different environments,[17] including clinical performance and professional conduct. Feedback should prompt self-reflection and management of the weaker aspects of performance. In the focus groups, both trainees and trainers identified feedback as a priority.[10]

Trainees should maintain an updated portfolio, including feedback and evidence of self-reflection, to be used as part of the assessment process

Portfolios are instruments used to collect and assess evidence of a trainee's experience and progression in tasks and competences.[18] They provide a key connection between learning at individual and organisational levels. The implementation of portfolios throughout Europe varies, and there is no consensus on their aims, design and content. The task force felt that portfolios should extend beyond a 'logbook' list of patients managed, procedures performed, courses attended or research performed. Rather, they should be an integral part of the continuous formative process and self-learning; as such, they should include examples of assessors' feedback and trainees' self-reflection. In order to promote honesty and self-critique, reflections included in the portfolio should be kept private and should not be misused or misconstrued in legal contexts. Use of electronic portfolios and, even better, integration within e-learning platforms increase their utility and address one of the key complaints of trainees[10]—the excessive time spent in their compilation. The EULAR portfolio task force has developed a portfolio structure which can be considered for uptake in different countries.[19]

Different methods of assessments should be carried out throughout training, as no single method of assessment can provide a complete overview of trainee competences

During training, rheumatologists acquire a wide variety of competences ranging from the ability to independently manage people with different forms of RMDs to the performance of specific skills (eg, aspirating a knee joint) or the acquisition of professional attitudes (eg, commitment to lifelong learning). No

single method of assessment can properly evaluate all competences. For example, written exams are unable to assess how a trainee works within a multidisciplinary team or whether they can perform a joint injection. In fact, the correlation between scores from assessment tools evaluating different competences is very weak.[20–24] On the other hand, different assessment methods can be used to assess a single competence, providing complementary information. For instance, the ability to aspirate a joint can be assessed with a mannequin (simulation) or on a real person with an RMD (eg, workplace-based via direct observation of procedural skills (DOPS)). In the second instance, assessment of a trainee's skill in patient communication or respect for patient autonomy can be included. When implementing a training programme and designing a local assessment strategy, thought must be given as to how each competence is assessed.

Multiple-choice case-based questions should be the preferred form of knowledge assessment

Each competence is composed of the integration of specific knowledge, skills and professional behaviours.[25] When assessing knowledge, we are therefore not assessing overall competence, but one of the core pillars that support it. Emphasis was placed on the fact that multiple choice questions (MCQs) should be based on a clinical scenario (case-based), allowing the assessment of complex clinical reasoning rather than the mere memorisation of specific facts. Oral examinations, commonly known as *vivas*, are not recommended as the inclusion of few examiners, the sampling of limited content and the use of a global judgement result in poor reliability.[26]

Clinical skills should be assessed in the workplace (DOPS) or the mini-clinical examination exercise (mini-CEX) and/or in a simulated context (observational structured clinical examination (OSCEs))

Clinical skills and competences can be assessed in a simulated environment or directly in the workplace. In simulated environments, the recommended assessment tool is the OSCE. An OSCE consists of multiple, time-limited stations where trainees perform specific tasks, under structured assessment. At each station, trainees are marked against standardised scoring checklists by trained assessors. In this manner, an OSCE can assess many competences. Typical competences assessed in this manner include performing a site-specific clinical examination, discussing treatment options or skills such as the identification of crystals in a synovial fluid sample.[20 21] Patient experts can be trained to role-play a patient with a given disease. In workplace assessments, a trainer observes the trainee interacting with a patient around a clinical task (mini-CEX) or a procedure (DOPS). The trainer uses a structured form to assess and provide feedback to the trainee. Encounters can take place in a variety of settings (inpatient, outpatient, emergency room) and contexts (initial or follow-up visit). The mini-CEX can be used to assess competences such as taking a focused history or performing a physical examination, while the DOPS is tailored for procedures such as joint aspiration, crystal identification or joint ultrasonography. Overall, each patient encounter takes 15–30 min followed by 5–10 min of feedback. It is expected that trainees are assessed several times throughout the year of training, with different trainers and in different clinical situations or with different focuses, so that different competencies are assessed.[5] A similar case may be especially dedicated to assess clinical examination or management planning, for example. The EULAR portfolio task force is developing forms for both the mini-CEX and the DOPS.[19]

Competences related to professionalism should be formally assessed using multisource feedback (MSF)/360° method

Professionalism is key to a good clinical practice and should be part of training and assessment. However, the assessment of professionalism is hampered by varying definitions and the difficulty in transforming the elements of professionalism into aspects that can be taught and measured.[27] It is beyond the scope of these PtCs to establish which aspects of professionalism should be assessed; these could include areas such as ethical practice, effective interaction with patients and relatives, working effectively with other health professionals, health authorities and other stakeholders, reliability and commitment to continuous improvement.[28] The MSF, also known as the 360° evaluation, allows the systematic collection of data on a trainee's performance, acquired anonymously from a variety of coworkers. Typically, 10–20 assessors comment on a specified range of that person's functioning. The assessors may include trainers, physicians, trainees, nurses, medical students, health professionals, patients and administrative personnel. MSF is especially useful in assessing actual behaviours in the workplace which are difficult to measure, or which can be concealed under formal assessment conditions. The results from the MSF should be discussed with the trainee in order to promote reflection.

The training programme should incorporate a predefined process to identify and support trainees at risk of failure

The identification of trainees who are at risk of failure within training programmes is a challenge.[29] Some trainers feel unprepared and/or unwilling to report a trainee's underperformance. Barriers include lack of documentation, lack of knowledge of what to document, anticipation of an appeal process and lack of remediation options.[30] Assessor development programmes, a strong assessment system with clear standards to be achieved at different training levels and a support system that offers guidance to the failing trainee are deemed essential.[31]

Trainers should receive regular training in assessment methods and strategies, particularly in providing constructive feedback

The existence, depth and scope of development programmes in assessment methods vary widely among countries[10] and can even be training centre-specific. Accepted training and assessment methods evolve with time as new evidence accrues. Continuous professional development in assessment methods and strategies should be encouraged by relevant stakeholders. Of special importance is training in providing constructive feedback,[32] a far more complex competence than it may seem. There is a recognised gap between the feedback given and what is perceived by the trainee. Feedback is effective when it leads to an improvement in the performance of the trainee. Both the skills of the person selecting and providing the feedback and the willingness and ability of the recipient to engage with it can modulate its effectiveness.

DISCUSSION

These are the first EULAR-endorsed PtCs for the assessment of competences in rheumatology specialty training. Their aim is to serve as a benchmark and an inspiration toinvolved stakeholders. In total, 41 EULAR countries provide rheumatology specialty training. Each country has its own training structure, curriculum and assessment strategy, resulting in a wide heterogeneity.[2] Some countries provide a comprehensive list of assessments to be undertaken, while some provide national, summative final examinations, and others provide broad statements. Overall,

Box 1 Research agenda

Barriers and enablers
- What are the key features of an assessment strategy that impact the professional development of trainees in a rheumatology training programme?

Competency components
- Which competences should be subjected, as a minimum, to formative assessment during the specialty training programme?

Frequency
- How often should formative (and summative) assessments take place?
- How often should each assessment method (eg, mini-CEX and DOPS) be performed?

Impact, value and outcomes
- How does a structured assessment of competences throughout training impact on training/learning outcomes and on care delivery outcomes?
- What is the impact of the use of a proper portfolio on training/learning outcomes and on care delivery outcomes?
- What is the added value of a summative assessment in the presence of a structured formative assessment programme?
- Do improvements in the quality of assessments translate into better outcomes and satisfaction for trainees and especially for patients?

Validity and reliability
- What is the validity of mini-CEX, DOPS and MSF in a rheumatology setting?
- What are the minimum requirements for an OSCE to be valid and reliable in a rheumatology training programme?

DOPS, direct observation of procedural skills; mini-CEX, mini clinical examination; MSF: multisource feedback; OSCE, objective structured clinical evaluation.

the specific implementation remains largely dependent on each centre's culture and attitude. These PtCs in no way attempt to undermine local regulations. Rather, they seek to provide recommendations of good practice, which can help stakeholders analyse their own assessment strategy and inspire positive change, where appropriate.

Many practising physicians are involved in assessing the competences of trainees. However, some are not as comfortable using educational assessment tools as they are managing patients with RMDs.[33] Assessment tools can measure knowledge or demonstrate competence in a simulated or in a 'real-life' setting.[34] Written examinations with MCQs can assess pure knowledge, but they are best employed in assessing its application to clinical problems; for this purpose, context-specific questions, based on a clinical scenario should be used. OSCEs can evaluate a trainee's skills and competences in a simulated environment. OSCEs can be used for both common or rarer diseases, highlighting the need for systematic assessment that might provide clues for the differential diagnosis, while rare diseases might be difficult to evaluate in workplace-based assessments. However, in order to evaluate what a trainee actually does, assessment needs to take place within the workplace by direct observation of a trainee's performance in a 'real-life' setting. Implementing a structure and effective assessment strategy within a busy clinic is a challenge, highlighting one of the barriers to workplace assessment. Tools such as the mini-CEX or the DOPS facilitate the standardisation of the assessment and feedback of clinical encounters and procedures. Professionalism is key to becoming an effective physician but is one of the most difficult aspects to define and measure. While some aspects of professionalism can be assessed in a simulated context (eg, efficient patient communication in an OSCE), most should be explored in the workplace. The major barrier for effective implementation of this multimodal assessment strategy is lack of time and resources (eg, trained trainers). Support from training centres, institutions and national authorities is key.

Even though specific evidence from rheumatology studies supporting these PtCs was scarce, the LoA with the PtC was very good. Published evidence identified in the SLR[9] was limited to the evaluation of some aspects of validity or reliability of a few assessment tools (OSCE, mini-CEX and DOPS). Indirect evidence, stemming from other medical specialties, provides additional support, but its applicability is varied, given the different contexts. As per EULAR standard operation procedures, the Oxford Levels of Evidence have been applied.[11] In medical education, quantitative evidence is scarce; specifically, evidence assessing the impact of different tools or strategies is lacking. Research allowing rheumatologists to implement best practices supported by consistent evidence would be welcome and is the basis of the proposed research agenda (box 1). While we await this, the high level of consensus that these recommendations provided is reassuring as to its cross-national validity.

In conclusion, these EULAR PtCs provide European guidance on assessment tools and strategies to be used throughout rheumatology training programmes. Given the relationship between learning and assessment, the harmonisation of assessment strategies could impact rheumatology training, encouraging stakeholders to strive for excellence and thereby optimise the future care delivered to people with RMDs.

Author affiliations
[1]Department of Clinical Medicine, Miguel Hernandez University of Elche, Elche, Spain
[2]Department of Rheumatology, Hospital General Universitario Elda, Elda, Spain
[3]Department of Medicine, Rheumatology Unit, University of Perugia, Perugia, Italy
[4]INSERM UMR1238, University of Medicine, CHU Nantes, Nantes, France
[5]Institute of Infection, Immunity and Inflammation, University of Glasgow College of Medical Veterinary and Life Sciences, Glasgow, UK
[6]Department of Allergology, Rheumatology and Clinical Immunology, University Children's Hospital Ljubljana, Ljubljana, Slovenia
[7]Rheumazentrum Ruhrgebiet, Herne, Germany
[8]Ruhr University Bochum, Bochum, Nordrhein-Westfalen, Germany
[9]Department of Rheumatology and Clinical Immunology, UMCUtrecht, Utrecht, Netherlands
[10]EULAR Patient Research Partner, Stockholm, Sweden
[11]Rheumatology and Clinical Immunology, Charité University Hospital, Berlin, Germany
[12]Division of Clinical Immunology and Rheumatology, University Hospital Centre Zagreb, Zagreb, Croatia
[13]Reumatologia, Centro Hospitalar e Universitário de Coimbra, Coimbra, Portugal
[14]Coimbra Institute for Clinical and Biomedical Research (iCBR), Faculty of Medicine, University of Coimbra, Coimbra, Portugal
[15]Institute of Rheumatology, University of Belgrade School of Medicine, Belgrade, Serbia
[16]Rheumatologie B, Hopital Cochin, Paris, Île-de-France, France
[17]Service de Rhumatologie, HFR Fribourg, Hôpital Cantonal, Fribourg, Switzerland
[18]Musculoskeletal Research Unit, NIHR Clinical Research Facility, University Hospital Southampton NHS Foundation Trust, Southampton, UK
[19]Academic Rheumatology Center, Università degli Studi di Torino, Torino, Italy
[20]Department of Rhumatologie, Hôpital Lariboisière, Paris, France
[21]INSERM UMR-1132, University of Paris, Paris, France
[22]Centre for Rheumatic Diseases, King's College London, London, UK
[23]Department of Medicine, Division of Rheumatology, Maastricht University Medical Center, Maastricht, The Netherlands

[24]School for Public Health and Primary Care (CAPHRI), Maastricht University, Maastricht, The Netherlands
[25]EULAR Patient Research Partner, Manchester, UK
[26]2nd Department of Medicine and Laboratory, Clinical Immunology-Rheumatology Unit, Athens University School of Medicine, Athens, Greece
[27]Center for Teaching and Learning, University of Oxford, Oxford, UK
[28]Rheumatology, Leiden University Medical Center, Leiden, The Netherlands
[29]Rheumatology, Zuyderland Medical Center, Heerlen, The Netherlands

Twitter Francisca Sivera @FranciscaSivera and Elena Nikiphorou @ElenaNikiUK

Acknowledgements We thank Laure Gossec for the thoughtful review of the manuscript and invaluable input throughout the project. We also thank Louise Falzon for the literature searches used in the systematic literature review.

Contributors FS, AA, AN, CH and SR participated in the design of the study. All authors participated in the development of the project, the interpretation of the data and the manuscript preparation, and approved the current version of the manuscript. CH and SR share last authorship.

Funding This study was funded by European League Against Rheumatism.

Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, conduct, reporting or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article.

ORCID iDs
Francisca Sivera http://orcid.org/0000-0002-3414-1667
Alessia Alunno http://orcid.org/0000-0003-1105-5640
Aurélie Najm http://orcid.org/0000-0002-6008-503X
Xenofon Baraliakos http://orcid.org/0000-0002-9475-9362
Jose AP Da Silva http://orcid.org/0000-0002-2782-6780
Elena Nikiphorou http://orcid.org/0000-0001-6847-3726
Sofia Ramiro http://orcid.org/0000-0002-8899-9087

REFERENCES

1 Faarvang KL, da Silva JA. Competencies in rheumatology: a European framework. *Best Pract Res Clin Rheumatol* 2009;23:145–60.
2 Sivera F, Ramiro S, Cikes N, *et al*. Differences and similarities in rheumatology specialty training programmes across European countries. *Ann Rheum Dis* 2015;74:1183–7.
3 Sivera F, Ramiro S, Cikes N, *et al*. Rheumatology training experience across Europe: analysis of core competences. *Arthritis Res Ther* 2016;18:213.
4 Schuwirth LWT, van der Vleuten CPM. General overview of the theories used in assessment: AMEE guide No. 57. *Med Teach* 2011;33:783–97.
5 Norcini J, Burch V. Workplace-place based assessment as an educational tool: AMEE guide No. 31. *Med Teach* 2007;29:855–71.
6 Larsen DP, Butler AC, Roediger HL. Test-enhanced learning in medical education. *Med Educ* 2008;42:959–66.
7 Kromann CB, Jensen ML, Ringsted C. The effect of testing on skills learning. *Med Educ* 2009;43:21–7.
8 van der Heijde D, Aletaha D, Carmona L, *et al*. 2014 update of the EULAR standardised operating procedures for EULAR-endorsed recommendations. *Ann Rheum Dis* 2015;74:8–13.
9 Alunno A, Najm A, Sivera F, *et al*. Assessment of competences in rheumatology training: results of a systematic literature review to inform EULAR points to consider. *Ann Rheum Dis* 2019;78:A1440.
10 Najm A, Alunno A, Sivera F, *et al*. Strategies for the assessment of competences during rheumatology training across Europe: results of a qualitative study. *RMD Open* 2020;6:e001183.
11 CEBM. Oxford centre for evidence-based medicine levels of evidence, 2016. Available: https://www.cebm.net/2016/05/ocebm-levels-of-evidence/
12 Heeneman S, Oudkerk Pool A, Schuwirth LWT, *et al*. The impact of programmatic assessment on student learning: theory versus practice. *Med Educ* 2015;49:487–98.
13 UEMS/EBR. Training requirements for the specialty of rheumatology. Available: http://www.eular.org/myUploadData/files/European-Training-Requirements-in-Rheumatology%20endorsed%20UEMS%20April%2012%202014.pdf [Accessed 2 Jan 2020].
14 Doherty M. Specialist training in rheumatology--the need for a curriculum and assessment. *Ann Rheum Dis* 1994;53:215–7.
15 Bing-You R, Hayes V, Varaklis K, *et al*. Feedback for learners in medical education: what is known? A scoping review. *Acad Med* 2017;92:1346–54.
16 Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE guide No. 31. *Med Teach* 2007;29:855–71.
17 Veloski J, Boex JR, Grasberger MJ, *et al*. Systematic review of the literature on assessment, feedback and physicians' clinical performance: BEME guide No. 7. *Med Teach* 2006;28:117–28.
18 Heeneman S, Driessen EW. The use of a portfolio in postgraduate medical education - reflect, assess and account, one for each or all in one? *GMS J Med Educ* 2017;34:Doc57.
19 van Onna M, Ramiro S, Haines C, *et al*. THU0586 establishing the key components of a EULAR portfolio for training in rheumatology: a EULAR school of rheumatology initiative. *Ann Rheum Dis*;79:535.
20 Pascual Ramos V, Medrano Ramírez G, Solís Vallejo E, *et al*. Performance of an objective structured clinical examination in a national certification process of trainees in rheumatology. *Reumatol Clin* 2015;11:215–20.
21 Berman JR, Lazaro D, Fields T, *et al*. The new York City rheumatology objective structured clinical examination: five-year data demonstrates its validity, usefulness as a unique rating tool, objectivity, and sensitivity to change. *Arthritis Rheum* 2009;61:1686–93.
22 Pascual-Ramos V, Guilaisne Bernard-Medina A, Flores-Alvarado DE, *et al*. The method used to set the pass mark in an objective structured clinical examination defines the performance of candidates for certification as rheumatologists. *Reumatol Clin* 2018;14:137–41.
23 Kissin EY, Grayson PC, Cannella AC, *et al*. Musculoskeletal ultrasound objective structured clinical examination: an assessment of the test. *Arthritis Care Res* 2014;66:2–6.
24 Humphrey-Murto S, Khalidi N, Smith CD, *et al*. Resident evaluations: the use of daily evaluation forms in rheumatology ambulatory care. *J Rheumatol* 2009;36:1298–303.
25 Frank JR, Snell L, Sherbino J. *CanMEDS 2015 physician comepetency framework*. Ottawa: Royal College of Physicians and Surgeons of Canada, 2015.
26 Daelmans HEM, Scherpbier AJJA, Van Der Vleuten CPM, *et al*. Reliability of clinical oral examinations re-examined. *Med Teach* 2001;23:422–4.
27 Woolf AD. How to achieve and enhance professionalism in rheumatology. *Best Pract Res Clin Rheumatol* 2009;23:127–44.
28 Wilkinson TJ, Wade WB, Knock LD. A blueprint to assess professionalism: results of a systematic review. *Acad Med* 2009;84:551–8.
29 Dacre J, Haq I. Assessing competencies in rheumatology. *Ann Rheum Dis* 2005;64:3–6.
30 Dudek NL, Marks MB, Regehr G. Failure to fail: the perspectives of clinical supervisors. *Acad Med* 2005;80:S84–7.
31 Yepes-Rios M, Dudek N, Duboyce R, *et al*. The failure to fail underperforming trainees in health professions education: a BEME systematic review: BEME guide No. 42. *Med Teach* 2016;38:1092–9.
32 Lefroy J, Watling C, Teunissen PW, *et al*. Guidelines: the do's, don'ts and don't knows of feedback for clinical education. *Perspect Med Educ* 2015;4:284–99.
33 Epstein RM. Assessment in medical education. *N Engl J Med* 2007;356:387–96.
34 Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990;65:S63–7.