

ORIGINAL ARTICLE

Clinical reasoning evaluation using script concordance test in primary care residents[☆]



Carlos Iglesias Gómez^{a,*}, Ofelia González Sequeros^b, Diego Salmerón Martínez^{c,d,e}

^a Centro de Salud Alcantarilla-Sangonera, Servicio Murciano de Salud, Alcantarilla, Murcia, Spain

^b Departamento de Anatomía Humana y Psicobiología, Facultad de Medicina, Universidad de Murcia, Campus de Espinardo, Murcia, Spain

^c Departamento de Ciencias Sociosanitarias, Universidad de Murcia, Campus de Espinardo, Murcia, Spain

^d Centro de Investigación Biomédica en Red (CIBER) Epidemiología y Salud Pública (CIBERESP), Murcia, Spain

^e Instituto Murciano de Investigación Biomédica-Arrixaca, El Palmar, Murcia, Spain

Received 1 July 2021; accepted 30 September 2021

Available online 4 July 2022

KEYWORDS

Script concordance test;
Paediatric primary care;
Medical evaluation;
Training course;
Clinical reasoning

Abstract

Introduction: Actual unawareness about paediatric resident's residency program factual clinical reasoning precludes professional retraining directed to solve deficiencies. Script Concordance Test (SCT) evaluates clinical reasoning due to its orientation to usual clinical practice but surprisingly it has not been used in Spain for Paediatric Primary Care clinical reasoning evaluation so far. Due to this we consider it to be of relevance to design a Paediatric Primary Care SCT which meets validity, reliability and accessibility criteria described on bibliography.

Methods: Development, validation and application of an SCT questionnaire for clinical reasoning analysis in Paediatric Primary Care applied on a population of paediatric residents and which includes demographic and employment data in order to study possible relationship between them and achieved scores.

Results: Our SCT was approved by an experts committee. It met reliability and accessibility criteria and it allowed distinguishing experts from paediatric internal residents. No statistically significant differences were found concerning age, gender, type and duration of the training received in Primary Care, and the completion of a course on that training.

[☆] Please cite this article as: Iglesias Gómez C, González Sequeros O, Salmerón Martínez D. Evaluación mediante script concordance test del razonamiento clínico de residentes en Atención Primaria. An Pediatr (Barc). 2022;97:87–94.

* Corresponding author.

E-mail address: doctoriglesias84@gmail.com (C. Iglesias Gómez).

Conclusions: We developed a SCT that was approved by a Paediatric experts committee, it met reliability and accessibility criteria and it allowed distinguishing clinical reasoning from experts and paediatric internal residents. Except second year residency program, we didn't objectified relevant differences in residency program year, age, gender, duration and realization of Paediatric Primary Care rotation, and training course realization.

© 2021 Asociación Española de Pediatría. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

PALABRAS CLAVE

Script concordance test;
Pediatría de Atención Primaria;
Evaluación médica;
Reciclaje formativo;
Razonamiento clínico

Evaluación mediante script concordance test del razonamiento clínico de residentes en Atención Primaria

Resumen

Introducción: Actualmente desconocemos el razonamiento clínico objetivo de los residentes durante su formación, lo que impide orientar su reciclaje formativo a solventar los déficits detectados. El Script Concordance Test evalúa el razonamiento clínico planteando situaciones clínicas reales, pero aún no ha sido utilizado en España con contenidos propios de Atención Primaria. Por ello consideramos relevante diseñar un Script Concordance Test de Atención Primaria que cumpla los criterios de validez, fiabilidad y aceptabilidad descritos en la bibliografía.

Métodos: Elaboración, corrección y validación de un Script Concordance Test para evaluar el razonamiento clínico en Atención Primaria de los residentes y que incluya variables socio-laborales para estudiar su posible relación con la puntuación obtenida.

Resultados: Nuestro cuestionario fue aprobado por un comité de expertos, alcanzó una fiabilidad y accesibilidad adecuadas, y distinguió a los expertos de los residentes. No se observaron diferencias estadísticamente significativas en función de la edad, género, tipo y duración de la formación recibida en Atención Primaria, y la realización de un curso sobre dicha formación.

Conclusiones: Se construyó un Script Concordance Test que obtuvo la aprobación por parte de un comité de expertos, cumplió con los criterios de fiabilidad y accesibilidad, y permitió objetivar diferencias significativas en el razonamiento clínico de los expertos y residentes. Excepto en el 2º año de residencia; no se observaron diferencias estadísticamente significativas respecto al año formativo dentro de la residencia, la edad, el género, la realización y duración de la rotación en Atención Primaria y la realización de un curso sobre dicha formación.

© 2021 Asociación Española de Pediatría. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Specialised medical intern-resident (MIR) training programmes have barely changed in the past 30 years¹ and need to be revised to adapt them to advances in education and evaluation.

In Spain, training in paediatrics is mostly delivered in the hospital setting², even though health care organization is based on a primary care (PC) model. Primary care paediatrics (PCP) focuses on the paediatric population at every stage and therefore the rotation of residents through this setting is crucial for their training, regardless of the setting in which their professional career develops in the future³.

It was not until 2006 that the new paediatrics MIR curriculum included a compulsory 3-month rotation in PCP³; however, different paediatrics societies have demanded that training at the PC level be extended to 12 months, in order to balance what is currently a nearly exclusively hospital-based training⁴. The implementation of this demand should be associated with an ongoing internal evaluation of the MIR system to identify areas of improvement and adapt training to the current context of PCP¹.

The training of physicians is based on acquisition of theoretical knowledge, practical skills and clinical reasoning skills, of which the last two are acquired during graduate education (post-medical baccalaureate degree). Typically, assessment during graduate education has been based exclusively on written exams, including multiple-choice tests, and the subjective appraisal of evaluators⁵, thus assessing performance out of context in a manner that may not adequately reflect real-world clinical practice.

At present, new methods such as the *script concordance test* (SCT) allow an adequate and objective evaluation of clinical reasoning, as they provide hypotheses to the subjects of the evaluation and guide the collection of information and decision-making, so that the collected information is used to confirm or reject the hypotheses⁵, as would be done in real-world clinical practice.

The characteristics of this type of test are based on the assumption that structured knowledge resides in a complex network of neural connections that interrelate different clinical scenarios and allow professionals to solve problems in clinical practice. Thus, the professionals with the greatest

expertise would not be those holding the greatest amounts of knowledge, but those with the richest networks of mental associations⁵.

The SCT presents a clinical case (Fig. 1), defines the hypotheses that the professional can consider based on the presenting signs and symptoms of the patient, and lastly provides additional information. When professionals receive this information, they have to assess how it fits in the clinical case originally described and how it affects the proposed hypothesis (confirmation, rejection or no effect) on a 5-point Likert scale⁵.

Several studies that have used SCTs have evinced that professional experience, measured merely as the number of years working in the field, is not associated with greater clinical reasoning skills if this experience is not associated with increasing involvement in the type of clinical practice that is under evaluation (for instance, the number of surgeries performed, in the case of gynaecology)⁶ or ongoing updating of knowledge and skills in the field (continuing education courses, in the case of paediatric cardiology)⁵. This suggests that clinical reasoning skills may decrease with professional experience if the latter is not associated with these two elements.

In addition, there is evidence that knowledge wanes over time in a so-called forgetting curve, which shows a logarithmic reduction in memory retention as a function of the time elapsed⁷. At present, many training programmes focus on the need to curb knowledge loss, improving long-term retention through continuing education of clinicians⁸. Constant updating of knowledge, which is the basis of clinical reasoning, is a way to fight against the forgetting curve, and thus can indirectly have a positive impact on clinical reasoning.

To date, SCTs have been applied in several specialities to assess clinical reasoning, but few studies have been conducted in the field of paediatrics, and none have been conducted in Spain with a focus on contents in the scope of PCP, despite the fact that PC is the setting where 90% of the presenting complaints of paediatric patients are resolved and where nearly 80% of residents that complete their MIR training end up working².

For this reason, we thought it would be relevant to design a SCT for PCP focusing on meeting the content and criterion-related validity, reliability and acceptability criteria described in the literature. The aim of our study was to assess:

- Whether residents have adequate enough clinical reasoning skills in PCP to engage in clinical practice,
- Whether the residency year in the MIR programme had an impact in the SCT score,
- Whether the years of experience in PCP of experts had an impact in the SCT score,
- Whether there is an association between the participation and duration of the rotation in PCP and the SCT score,
- Whether there is an association between the sex and age of residents and the SCT score,
- Whether participation in a PCP course is sufficient to improve the SCT score.

Methods

The lead author developed a SCT comprised of 115 items, adhering to the guidelines for the construction of these tests⁹ and using frequent reasons for consultation in PCP and the literature on PCP training and education¹⁰.

The validation of the questionnaire, a summary of which is presented in Fig. 2, started with the revision by 5 experts in PCP (with 6 or more years of experience), who edited the items and selected for inclusion those approved by 80% of the raters (expert validity criterion¹¹), thus reducing the length of the questionnaire to 105 items.

To form the panel of experts, we submitted the questionnaire online through the Google Form platform (available at <https://www.google.es/intl/es/forms/about/>) to several paediatrics societies and forums. We received 45 completed questionnaires, thus amply exceeding the response rate in previous studies in the field of paediatrics^{12,13}.

We also collected data on the sex, age and years of experience in PCP. The scoring key was created based on the answers of the panel of experts, eliminating items with low discrimination (extreme variance, unanimous responses or uniform distribution of responses), so that the final questionnaire was reduced to 67 items.

Lastly, the definitive SCT was sent to 1750 residents in paediatrics teaching units in Spain, and we received 102 completed questionnaires (response rate of 5.8%).

The passing score of the questionnaire was determined based on the criterion proposed by Duggan of classifying subjects based on their score transformed to standard deviations (SDs) from the expert mean, setting 4 SDs as the cut-off point¹⁴.

The questionnaire collected the following data on respondents: residency year, age, sex, rotation or lack of rotation in PCP, duration of this rotation (<1 month, 1–3 months, >3 months), and participation or lack of participation in a PCP course.

We used the Student *t*-test to compare 2 means and one-way analysis of variance (ANOVA) to compare more than 2 means. We analysed the association between quantitative variables graphically through linear regression. Statistical significance was defined as a two-tailed *P* value of 0.05 or less, and all tests were performed with the Stata© software, version 14.2.

The psychometric properties of the questionnaire assessed in the study were its validity, reliability (using the Cronbach alfa¹⁵) and acceptability.

Results

When it came to validity, the 67 questions of the questionnaire fulfilled the expert validity criterion (more than 80% of agreement) and allowed confirmation of the hypothesis regarding the level of training: experts scored higher than residents, and therefore the questionnaire could discriminate between them.

As concerns reliability, the Cronbach alfa for medical residents was 0.74.

Five residents (4.9%) failed the test based on the Duggan criterion Duggan (4 SD cutoff). Two experts were excluded

Child aged 8 years whose presenting complaint is an episode of syncope in the context of physical exertion		
	If you were thinking of:	And then you were to find:
9	A diagnosis of long QT syndrome	The patient reports having experienced this several times in the past
10	A diagnosis of long QT syndrome	The patient was swimming in the school pool when the episode took place

	1 (completely rules out the initial hypothesis)	2 (makes the hypothesis less likely but does not rule it out completely)	3 (Does not affect the initial hypothesis)	4 (Makes the hypothesis more likely but does not fully confirm it)	5 (Fully confirms the hypothesis)
Item 9	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Item 10	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figure 1 Example of a clinical case in a SCT⁵.

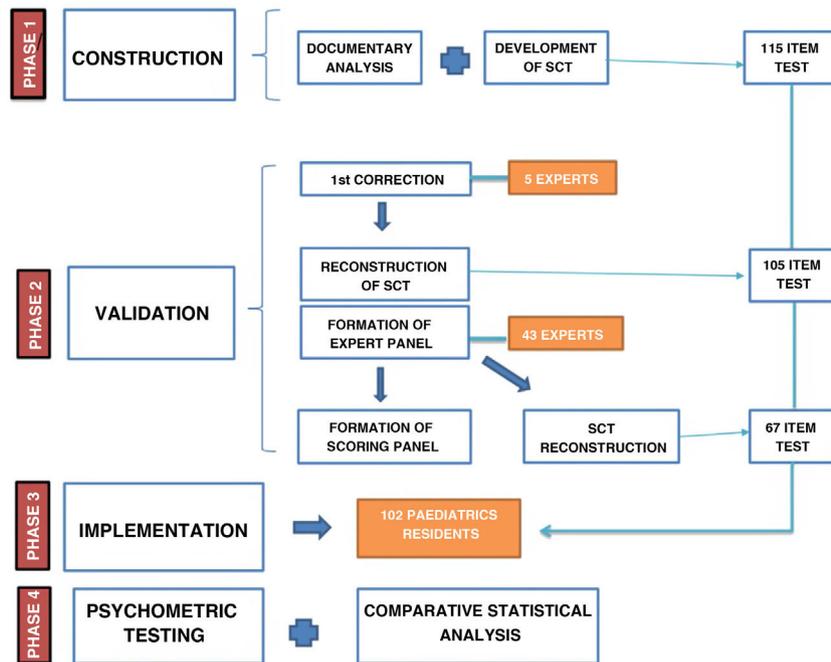


Figure 2 Flowchart of the phases of construction, validation, implementation and psychometric analysis of the SCT.

from the analysis because their scores were below 50 (49.4 and 46.1).

The mean score of the experts ($N=43$) was 77.3 (SD, 5.4), greater than the mean of 71.7 (SD, 7.8) in the 102 residents ($P<0.00005$). This significant difference compared to the experts was observed in year 1 residents (R1) ($P<0.00005$), year 3 residents (R3) ($P<0.001$) and year 4 residents (R4) ($P<0.011$), but not in year 2 residents (R2) ($P=0.078$). Year 2 residents had the highest mean score (Table 1 and Fig. 3), 4.5 points above R1s ($P=0.0287$; 95% confidence interval [CI], 0.5–8.5). The differences relative to R3s and R4s were not statistically significant.

As can be seen in Fig. 4, we found no evidence of an association between the years of experience in PCP of the experts and their score in the SCT ($P=0.1$).

We did not find statistically significant differences in the scores of male vs female residents ($P=0.331$). The association between age and the SCT score was also not significant ($P=0.5753$) (Fig. 5).

As regards rotations in PCP (Fig. 6), we did not find significantly better scores in residents that did rotations ($P=0.058$). We also found no significant differences comparing the 3 possible durations of rotations in PCP ($P=0.4089$).

When we compared the mean scores of residents based on whether or not they had taken a course in PCP (Fig. 7),

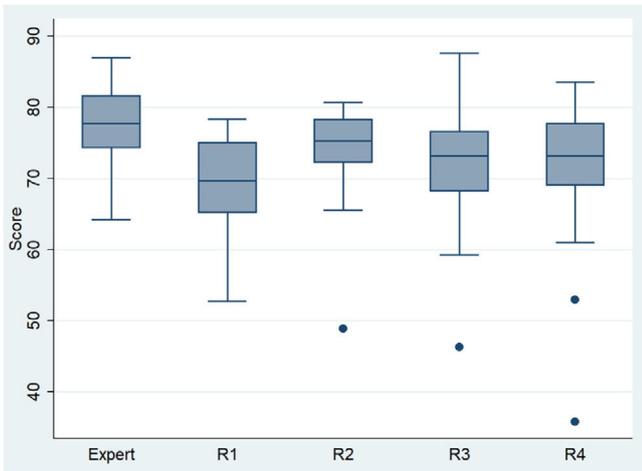


Figure 3 Boxplot of scores in the group of experts and residents in each year of the residency: (R1 = year 1 residents; R2, year 2 residents; R3, year 3 residents; R4, year 4 residents).

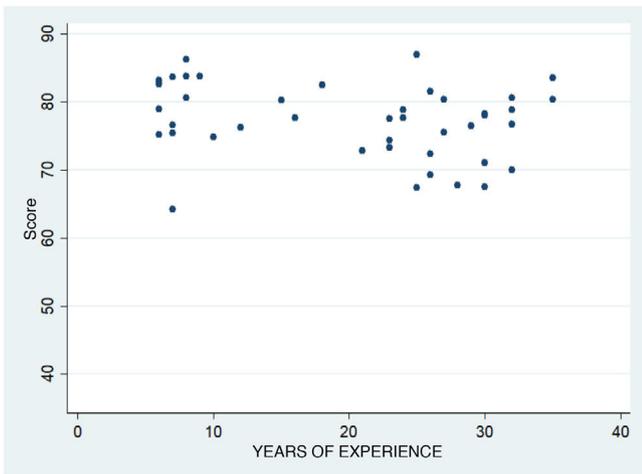


Figure 4 Scatterplot showing the distribution of assessed experts by SCT score and years of experience in PCP.

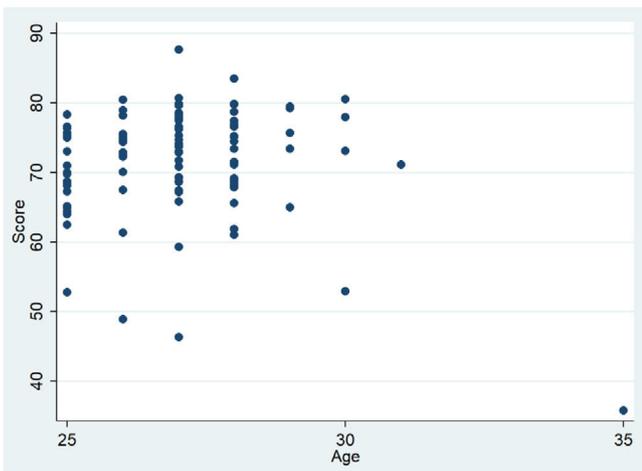


Figure 5 Scatter plot showing the distribution of the assessed residents by SCT score and age.

Table 1 Measures of central tendency and dispersion of the scores of the residents in the script concordance test.

	<i>n</i>	Mean	SD
<i>Year of residency</i>			
R1	25	69.6	6.1
R2	21	74.1	7.2
R3	32	71.9	7.5
R4	24	71.4	10
<i>Sex</i>			
Male	24	70.1	9.2
Female	78	72.1	7.3
<i>Rotation in PCP</i>			
Yes	78	72.4	8.1
No	24	69.3	6.3
<i>Duration of rotation</i>			
1 month	15	73.8	8.2
1-3 months	56	72.5	7
>3 months	7	68.8	15
<i>PCP course</i>			
Yes	55	71.7	9.5
No	47	71.7	5.4

PCP, primary care paediatrics; SD, standard deviation.

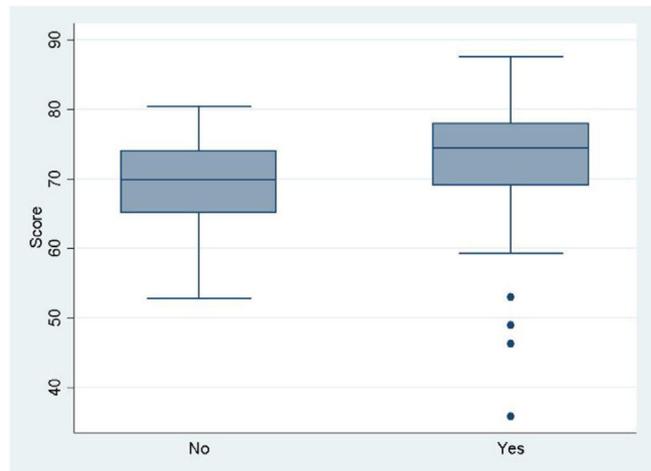


Figure 6 Boxplot showing the measures of central tendency and dispersion of the scores achieved by paediatrics residents based on whether or not they participated in a rotation in PCP.

we did not find significantly better scores in those that had taken one ($P=0.995$).

Discussion

We created a SCT questionnaire that comprising 67 items that met validity criteria, as clinical reasoning was better in the experts compared to the overall group of residents, as observed in previous works^{13,16–21}.

As regards reliability, we estimated an internal consistency of 0.74, which would be adequate, given that it exceeded the previously established minimum threshold of 0.7²².

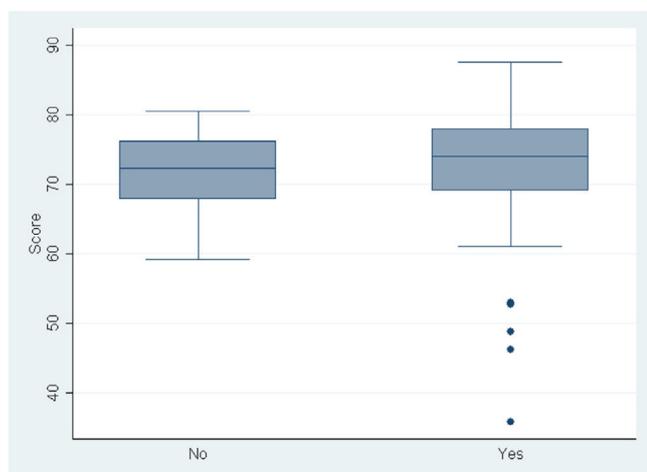


Figure 7 Boxplot showing the measures of central tendency and dispersion of the scores achieved by paediatrics residents based on whether or not they participated in a PCP course.

We believe that reliability could have improved by increasing the length of the questionnaire, but that this would have had a negative impact on the sample size. In spite of this, the sample of 102 responses allowed us to achieve the validity objectives set for the questionnaire, and has only been surpassed by a study that included 268 responses, although all were from residents in the same year of training²³.

The acceptability (construction, scoring, correction, codification and interpretation of the questionnaire) corresponded to what has been described in the literature.

The percentage of residents that failed the test (4.9%) was very similar to the one described by Duggan (3%)¹⁴ and most other studies^{6,24-27}, which suggests that we used an adequate threshold to determine whether residents passed the test and correctly identify those with clinical reasoning deficits.

As expected, in association with the level of training of respondents, scores were higher in the group of experts compared to the residents. These differences were consistent with multiple previous studies^{12-14,28-34}, which confirms that the questionnaire was constructed correctly.

However, we did not find an increase in the score based on advancing through the residency years. This was consistent with the findings of Power et al.¹⁸ and Steinberg et al.³⁵, whose studies did not find an increase in the score based on increasing year of residency, and contrary to the results of 2 studies that also assessed residents in paediatrics^{12,13}.

Although it would have been reasonable to expect higher scores in year 4 residents, who in theory would be better trained, it was year 2 residents that had the highest scores.

We can think of two factors that may explain the higher scores in second-year residents: first, having completed the first year of residency, which sets the foundation of the training, and second, that rotations completed in the first year of residency are in settings that require knowledge and skills that are more pertinent to clinical reasoning in PCP, such as the emergency department and neonatal care.

The lack of significant differences in the comparison of R2s, R3s and R4s was also consistent with the studies

of Power et al.¹⁸ and Steinberg et al.³⁵, also supporting the notion that this tool would not be useful to discriminate between residency levels, but instead to discriminate between fully trained physicians and residents in training.

Contrary to the studies conducted by Rajapreyar et al.³¹ and Kazour et al.³⁶, in which scores decreased with increasing professional experience or with the lack of continuing education⁵, we did not find an association between the years of experience in PCP of the experts and the score in the SCT.

We consider that, given the broad scope of the field, it would be unusual for assessed experts to not have taken at least on course on PCP during their careers. Especially given that we recruited the sample of experts from the membership of several paediatrics societies and forums, both of which promote continuing education. This could have been a source of selection bias and explain why the score did not decrease with age, as older participants could have taken more courses and then have more updated training, although we did not explore this variable in our study.

We did not find differences in the SCT score based on sex, and our sample, in which 76.4% of the residents were female, had a sex distribution consistent with the descriptions in the current literature¹.

In relation to rotations in PCP, we did not find statistically significant differences in the score. This was consistent with the findings of Peyrony et al.³⁷ However, other authors have found statistically significant differences in SCT scores following rotations^{36,38,39}.

We also compared the scores of experts and residents, between residents by residency year, and by duration of PCP rotation using the Mann-Whitney *U* test and Kruskal-Wallis test, and had the same results.

In light of the findings of these works, we believe that a variable that may have affected our results is the current 3-month duration of PCP rotations, which may be insufficient to improve clinical reasoning, as longer rotations may be required, as advocated by paediatricians², to improve outcomes and achieve significant differences, an argument that is supported by previous evidence^{5,36,38}.

The absence of significant differences in the SCT score in favour of those who had taken courses in PC contradicted the findings of several studies in the field of family medicine, paediatrics and pharmacy^{5,40,41}.

One of the possible explanations for this finding could be that we did not take into account the number of continuing education courses that had been taken, but only whether respondents had participated in any. We think that this may have made it difficult to discriminate between clinicians that had been updating their knowledge and skills versus those who had not, which could account for this result.

The hypothesis that several courses need to be taken to improve clinical reasoning and, therefore, the score in the SCT, reinforces the importance of continuing education, as opposed to the accumulation of years of paediatric clinical practice^{5,31,42}, as we mentioned in the introduction. In addition, the development of continuing education programmes focused on the educational deficiencies of clinicians, would ensure the quality of care delivery and allow meeting the health care needs of the population more efficiently. Although few studies have analysed this aspect, there is evidence of a substantial increase in scores after completion of educational programmes^{5,40,41}.

Chief among the limitations of our study is the inability to cover all the contents of PCP in the SCT, as the excessive length of the questionnaire would have resulted in an even smaller sample size. We also think that the low response rate may be due to a reluctance on the part of some residents to answer questions due to their limited clinical reasoning skills, leading them to refuse participation, which would have limited the sample size but also given rise to a sample that was not uniformly representative of residents, so that its reliability, while adequate, could not be higher. We believe that administration of the questionnaire under supervision and placing a limit on the response time could help achieve a larger sample size.

Thus, we conclude that this SCT could be useful for evaluation and in the future continuing education of residents. In developing this tool, our intent is not to substitute but to complement the current assessment of residents, assisting resident supervisors and the individuals in the institution responsible for developing the MIR programme in assessing residents comprehensively and objectively, so we propose its incorporation by health care administrations in the evaluations conducted at the different stages of paediatrics specialty training.

We believe that if the SCT was administered in person, under supervision and as a compulsory step in residency training, limiting the time given for its administration and optimising its contents and length to increase the sample size and internal consistency, it could sensibly improve paediatric care delivery.

Conclusions

We built a SCT with 67 items that was approved by a panel of experts, could discriminate clinicians based on their level of training, allowed assessment of whether clinicians had adequate clinical reasoning skills in PCP and whether there were differences between fully trained physicians on staff and residents.

Except for year 2 of the residency, we did not find statistically significant differences based on the year of residency, age, sex, participation in a rotation in PC and its duration and participation in a PCP course.

Having analysed these parameters, we may conclude that this SCT could be a useful tool to assess the adequacy of clinical reasoning skills in PCP in medical residents.

Conflicts of interest

The authors have no conflicts of interest to declare.

References

- González de Dios J, Polanco Allue I, Díaz Vázquez CA. De las facultades de Medicina a la residencia de Pediatría, pasando por el examen de médico interno residente: ¿algo debe cambiar? Resultados de una encuesta en residentes de Pediatría de cuarto año. *An Pediatr (Barc)*. 2009;70:467–76.
- García Puga JM, Villazán Pérez C, Domínguez Aurrecochea B, Ugarte Libano R. ¿Qué opinan pediatras y residentes de Pediatría sobre la rotación por Atención Primaria? *An Pediatría (Barc)*. 2009;70:460–6.
- Domínguez Aurrecochea B, Valdivia Jiménez C. La pediatría de atención primaria en el sistema público de salud del siglo XXI. *Informe SESPAS. Gac Sanit*. 2012;26:82–7.
- Gorrotxategi Gorrotxategui P, García Vera C, Graffigna Lojendio A, Sánchez Pina C, Palomino Urda N, Rosa Rodríguez Fernández-Oliva C. Situación de la Pediatría de Atención Primaria en España en 2018. *Rev Pediatr Aten Primaria*. 2018;20:89–104.
- Iglesias Gómez C, González Sequeros O, Sarquella Brugada G, Padilla del Rey ML, Salmerón Martínez D. Usefulness of SCT in detecting clinical reasoning deficits among pediatric professionals. *Prog Pediatr Cardiol*. 2021;61:101340.
- Kow N, Walters MD, Karram MM, Sarsotti CJ, Jelovsek JE. Assessing intraoperative judgment using script concordance testing through the gynecology continuum of practice. *Med Teach*. 2014;36:724–9.
- Murre JMJ, Dros J. Replication and analysis of Ebbinghaus' forgetting curve. *PLoS One*. 2015;10:e0120644.
- Kerfoot BP, Fu Y, Baker H, Connelly D, Ritchey ML, Genega EM. Online spaced education generates transfer and improves long-term retention of diagnostic skills: a randomized controlled trial. *J Am Coll Surg*. 2010;211:331–7.e1.
- Fournier JP, Demeester A, Charlin B. Script concordance tests: guidelines for construction. *BMC Med Inform Decis Mak*. 2008;8:18.
- Grupo Docencia Mir, Available in: <https://www.aepap.org/grupos/grupo-de-docencia/biblioteca/programas-formativos-0>, 2010.
- Escurre Mayaute LM. Cuantificación de la validez de contenido por criterio de jueces. *Rev Psicol*. 1988;6:103–11.
- Cooke S, Lemay J-F, Beran T, Sandhu A, Amin H. Development of a method to measure clinical reasoning in pediatric residents: the pediatric script concordance test. *Creat Educ*. 2016;07:814–23.
- Carrière B, Gagnon R, Charlin B, Downing S, Bordage G. Assessing clinical reasoning in pediatric emergency medicine: validity evidence for a script concordance test. *Ann Emerg Med*. 2009;53:647–52.
- Duggan P, Charlin B. Summative assessment of 5th year medical students' clinical reasoning by script concordance test: requirements and challenges. *BMC Med Educ*. 2012;12:29.
- Friás Navarro D. Análisis de fiabilidad de las puntuaciones de un instrumento de medida. Alfa de Cronbach: un coeficiente de fiabilidad. In: *Apuntes de SPSS*. Universidad de Valencia; 2014. p. 10.
- Brailovsky C, Charlin B, Beausoleil S, Cote S, Van der Vleuten C. Measurement of clinical reflective capacity early in training as a predictor of clinical reasoning performance at the end of residency: an experimental study on the script concordance test. *Med Educ*. 2001;35:430–6.
- Brazeau-Lamontagne L, Charlin B, Gagnon R, Samson L. Measurement of perception and interpretation skills during radiology training: utility of the script concordance approach. *Med Teach*. 2004;26:326–32.
- Power A, Lemay J-F, Cooke S. Justify your answer: the role of written think aloud in script concordance testing. *Teach Learn Med*. 2017;29:59–67.
- Park AJ, Barber MD, Bent AE, Dooley YT, Dancz C, Sutkin G, et al. Assessment of intraoperative judgment during gynecologic surgery using the Script Concordance Test. *Am J Obstet Gynecol*. 2010;203:1–6.
- Ruiz JG, Tunuguntla R, Charlin B, Ouslander JG, Symes SN, Gagnon R, et al. The script concordance test as a measure of clinical reasoning skills in geriatric urinary incontinence. *Am Geriatr Soc*. 2010;58:2178–84.
- Sibert L, Darmoni SJ, Dahamna B, Hellot F, Weber J, Charlin B. On line clinical reasoning assessment with Script Concordance

- test in urology: results of a French pilot study. *BMC Med Educ.* 2006;6:45.
22. Lubarsky S, Charlin B, Cook DA, Chalk C, van der Vleuten CPM. Script concordance testing: a review of published validity evidence: validity evidence for script concordance tests. *Med Educ.* 2011;45:329–38.
 23. Hamui M, Ferreira JP, Torrents M, Torres F, Ibarra M, Ossorio MF, et al. Prueba de concordancia de script: primera experiencia en pediatría de alcance nacional. *Arch Argent Pediatr [Internet]*. 2018;116 [Accessed 26 June 2021]. Available in: <http://www.sap.org.ar/docs/publicaciones/archivosarg/2018/v116n1a43.pdf>
 24. Nseir S, Elkalioubie A, Deruelle P, Lacroix D, Gosset D. Accuracy of script concordance tests in fourth-year medical students. *Int J Med Educ.* 2017;8:63–9.
 25. Goulet F, Jacques A, Gagnon R, Charlin B, Shabah A. Poorly performing physicians: does the script concordance test detect bad clinical reasoning? *J Contin Educ Health Prof.* 2010;30:161–6.
 26. Irfannuddin I. Knowledge and critical thinking skills increase clinical reasoning ability in urogenital disorders: a Universitas Sriwijaya Medical Faculty experience. *Med J Indones.* 2009;18:53.
 27. Piovezan RD, Batista NA, Lubarsky S, Charlin B. Assessment of undergraduate clinical reasoning in geriatric medicine: application of a script concordance test. *J Am Geriatr Soc.* 2012;60:5.
 28. Lubarsky S, Durning S, Charlin B. AM last page: the script concordance test: a tool for assessing clinical data interpretation under conditions of uncertainty. *Acad Med.* 2014;89:1.
 29. Faucher C, Dufour-guindon M, Lapointe G, Gagnon R, Charlin B. Assessing clinical reasoning in optometry using the script concordance test. *Clin Exp Optom.* 2016;99:280–6.
 30. Humbert AJ, Johnson MT, Miech E, Friedberg F, Grackin JA, Seidman PA. Assessment of clinical reasoning: A Script Concordance test designed for pre-clinical medical students. *Med Teach.* 2011;33:472–7.
 31. Rajapreyar P, Marcdante K, Zhang L, Simpson P, Meyer MT. Decision-making in pediatric transport team dispatch using script concordance testing. *Pediatr Crit Care Med.* 2017;18:e530–5.
 32. Chime NO, Pusic MV, Auerbach M, Mehta R, Scherzer DJ, Ittersum WV, et al. Script concordance testing to determine infant lumbar puncture practice variation. *Pediatr Emerg Care.* 2018;34:84–92.
 33. Charlin B, Gagnon R, Lubarsky S, Lambert C, Meterissian S, Chalk C, et al. Assessment in the context of uncertainty using the script concordance test: more meaning for scores. *Teach Learn Med.* 2010;22:180–6.
 34. Talvard M, Olives J-P, Mas E. Évaluation des étudiants en médecine lors de leur stage en gastro-entérologie pédiatrique par un test de concordance de script. *Arch Pédiatrie.* 2014;21:372–6.
 35. Steinberg E, Cowan E, Lin MP, Sielicki A, Warrington S. Assessment of emergency medicine residents' clinical reasoning: validation of a script concordance test. *West J Emerg Med.* 2020;21:7.
 36. Kazour F, Richa S, Zoghbi M, El-Hage W, Haddad FG. Using the script concordance test to evaluate clinical reasoning skills in psychiatry. *Acad Psychiatry.* 2017;41:86–90.
 37. Peyrony O, Hutin A, Truchot J, Borie R, Calvet D, Albaladejo A, et al. Impact of panelists' experience on script concordance test scores of medical students. *BMC Med Educ.* 2020;20:313.
 38. Lambert C, Gagnon R, Nguyen D, Charlin B. The script concordance test in radiation oncology: validation study of a new tool to assess clinical reasoning. *Radiat Oncol.* 2009;4:7.
 39. Piquette D, Goffi A, Lee C, Brydges R, Walsh CM, Mema B, et al. Resident competencies before and after short intensive care unit rotations: a multicentre pilot observational study. *Can Anesth.* 2021;68:235–44.
 40. Olszynski P, Anderson J, Trinder K, Domes T. Point-of-care ultrasound in undergraduate urology education: a prospective control-intervention study: point-of-care ultrasound in urology education. *J Ultrasound Med.* 2018;37:2209–13.
 41. Atayee RS, Lockman K, Brock C, Abazia DT, Brooks TL, Pawasauskas J, et al. Multicentered study evaluating pharmacy students' perception of palliative care and clinical reasoning using script concordance testing. *Am J Hosp Palliat Med.* 2018;35:1394–401.
 42. Dufour S, Latour S, Chicoine Y, Fecteau G, Forget S, Moreau J. Use of the script concordance approach to evaluate clinical reasoning in food-ruminant practitioners. *J Vet Med Educ.* 2012;39:267–75.