Jaggi, A., Drake, M., Siddiqui, E., Nazir, J., Giagos, V., & Fatoye, F. (2018). A critical appraisal of the principal guidelines for neurogenic lower urinary tract dysfunction using the AGREE II instrument6. *Neurourology and Urodynamics*. https://doi.org/10.1002/nau.23775

A Critical Appraisal of the Principal Guidelines for Neurogenic Lower Urinary Tract Dysfunction Using the AGREE II Instrument.

ABSTRACT

AIMS: The process of identifying research questions, synthesising and interpreting evidence, and weight given to health economics differs between the clinical guidelines (CGs) for neurogenic lower urinary tract dysfunction (NLUTD). Consequently, the quality also varies which can have implications for clinical practice.

METHODS: We used the Appraisal of Guidelines for Research and Evaluation (AGREE) instrument to assess the quality of the National Institute for Health and Care Excellent (NICE), European Association of Urology (EAU) and the International Consultations on Incontinence (ICI) CGs on neurogenic bladder.

RESULTS: The NICE CGs were deemed to be of the highest quality (overall score of 92%). NICE were the only guidelines to systematically incorporate cost-effectiveness research into their recommendations. The EAU CGs received an overall score of 83% and the ICI CGs achieved the lowest overall score (75%). The highest scoring domain amongst all the CGs was scope purpose (86%) and the lowest scoring domain was applicability (69%). All guidelines were recommended for use (mostly with some modifications).

CONCLUSIONS: All CGs had their inherent advantages and disadvantages, though all were still deemed to be of high quality. Incorporating cost-effectiveness research would be near impossible for guidelines with a broad-country remit. Incorporating the AGREE II instrument in the development of CGs and better collaboration between the ICI, NICE and EAU could improve the quality, and consistency between NLUTD CGs and ultimately improve health outcomes for this important patient group.

INTRODUCTION Neurogenic lower urinary tract dysfunction (NLUTD) is a urological dysfunction that occurs as a consequence of neurologic disease. It affects approximately 27-85% of patients with Parkinson's disease (PD), 70-84% with spinal cord injuries (SCI), up to 70% of those with stroke, and 40–90% of persons with multiple sclerosis (MS) [1-3]. Individuals with NLUTD may experience neurogenic detrusor overactivity (NDO), which is characterised by increased frequency of micturition, urinary urgency (if sensation is unaffected by the underlying condition) and urinary incontinence. Alternatively, patients may have problems in voiding, with symptoms including hesitancy, a slow urinary stream, the need to strain and urinary retention. NLUTD has a substantial impact on patients' health related quality of life (HRQoL) and use of healthcare resources due to bladder symptoms and associated sequela [4].

The National Academy of Medicine (NAM) (formerly known as the Institute of Medicine (IOM)) **was founded in 1970, under the charter of the National Academy of Sciences. The organisation comprises of 80 prominent members in the field of medicine and beyond[5]. The NAM** define clinical guidelines (CGs) as "statements that include recommendations intended to optimize patient care that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options".[6] CGs are an important tool in establishing evidence-based medicine (EBM) in clinical practice, and adequate implementation can improve patient outcomes, as well as inefficiencies and inequity across care institutions [7]. The three most prominent organisations that produce CGs for the management of NLUTD are the National Institute for Health and Care Excellent

(NICE), the European Association of Urology (EAU) and the International Consultations on Incontinence (ICI). [8-10]

The process of identifying research questions, synthesising and interpreting evidence differs between CGs. These differences are often the result of differing goals, financial resources and membership of organisations. Consequently, the quality also varies, which can have implications for clinical practice. For example, some developers employ rigorous systematic reviewing techniques whilst other CGs weigh more heavily on expert opinion. Another key differentiating factor is the weight given to health economic evidence. Whereas some CGs include well-integrated economic analysis to determine the most cost-effective management strategies, others focus solely on clinical outcomes.

The Appraisal of Guidelines for Research and Evaluation (AGREE) collaboration defines good quality CGs as "the confidence that the potential biases of guideline development have been addressed adequately and that the recommendations are both internally and externally valid, and are feasible for practice". [11] They developed the AGREE II instrument to critically appraise the transparency and methodological rigour of CG development. [11] The instrument was utilised in the current study to determine the quality of available NLUTD CGs, and identify where potential improvements could be made.

MATERIALS AND METHODS

AGREE II Instrument

Two appraisers (AJ, ES) independently assessed the quality of the NLUTD CGs using the AGREE II instrument. The instrument consists of 23 items, grouped into six domains: (1) scope and purpose (items 1-3), (2) stakeholder involvement (items 4-6), (3) rigor of development (items 7-14), (4) clarity and presentation (items 15-17), (5) applicability (items 18-21), and (6) editorial independence (items 22-23) (Table 1). Each item is rated on a seven-point Likert scale, where seven correlates to strongly agree. It is important to note that a score of 1 does not necessarily mean that the item criterion was not fulfilled, instead this could represent a lack of relevant information available to the appraiser to assign an appropriate score. The instrument also asks appraisers to make two assessments; on the overall guideline quality, and whether they would recommend the CGs for use.

Table 1 - Description of the AGREE II instrument

Data analysis

Descriptive statistics were used to summarise the domain and overall scores. A standardised score for the six domains and overall score was calculated by summing the scores of the individual items within each domain to achieve a percentage of the maximum possible score.

The IBM SPSS Statistics version 24 package was used to calculate the agreement between the appraisers using intraclass correlation (ICC), which demonstrates the level of agreement between appraisers. A single measures, two-way random effects model was utilised. The range of ICC is between 0 and 1, where the closer to 1 a score is the smaller the variation between scores of raters on each item [12].

RESULTS

Clinical guidelines

The three CGs included in this study had notably different intentions of use. In contrast to NICE and the EAU, ICI is not intended to be applied directly in clinical practice. Table 2 describes the characteristics of the CGs.

Table 2: Description of Neurogenic bladder (NLUTD) clinical guidelines

Scaled domain scores are presented in Table 3. The NICE CGs were deemed to be of the highest quality (overall score of 92%), they scored highest in stakeholder involvement domain (94%), and the lowest scoring domains were clarity of presentation and scope and purpose (86% in both domains). The EAU CGs received an overall score of 83%, the highest scoring domain was clarity of presentation (89%) and the lowest scoring domain was the applicability domain (63%). The ICI CGs achieved the lowest overall score amongst the CGs (75%). The highest scoring domain in this CG was clarity of presentation (94%) and the lowest scoring domain was applicability (54%). The ICC varied from low to excellent reliability (0.3-1); however, confidence intervals were insignificant in some domains (Table 4).

Table 3: Scaled domain percentages for AGREE II domains in the appraisal of neurogenic lower urinary tract dysfunction (NLUTD) guidelines

Table 4: Intraclass correlation between two appraisers of neurogenic bladder guidelines 95% CI not presented = CI crossed 0, therefore not significant ICC = <0.5 poor reliability, 0.5-<0.75 moderate reliability, 0.75-0.9 good reliability, >0.9 excellent reliability.

DISCUSSION

To the best of the authors' knowledge, this is the first study that assesses the quality of the NLUTD guidelines by using the AGREE II scores. Quality varied moderately across the AGREE II domains as well as between the NLUTD CGs. Amongst all CGs, the highest scoring domain was clarity of presentation and the lowest scoring was applicability. NICE achieved the highest overall score and the ICI achieved the lowest overall score, however all CGs were deemed to be of high quality, and were recommended for use in clinical practice (mostly with some modifications).

The stakeholder representation domain evaluates the extent to which CGs have accurate representation from all relevant intended users, including professional groups and patients. Involving a broad range of stakeholders allows the integration of several unique perspectives on optimal healthcare, aids in the prioritisation of important topics, and minimises bias towards certain treatment options caused by conflicts of interest[13]. NICE scored exceptionally high in the stakeholder involvement domain (94%). The NICE CGs are developed not only by urological experts working in the field but also by a rigorous process of cross-collaboration with specialist and/or general physicians, HEOR specialists and patient groups. In contrast, the development group for both the EAU and ICI NLUTD CGs are made up almost exclusively of neuro-urological experts; they achieved 78% and 67% respectively. The EAU is slowly integrating patient perspective into their development process by engaging patient organisations, whereas the ICI acknowledge that increased efforts to incorporate the

patient voice into their CGs is necessary. The transparency with which the stakeholders' comments are incorporated into recommendations is an aspect all CGs need to improve upon.

The most vital aspect in the formation of evidence-based recommendations is a comprehensive systematic review of available evidence [14]. Recommendations in all three NLUTD CGs occasionally relied on expert opinion. Unfortunately, as the evidence base underlying NLUTD is composed of mainly observational studies, and trials with relatively small patient numbers and perceived weak methodological design, this cannot be avoided. The NICE systematic review process was deemed the most superior by the appraisers, thus achieved the highest score in the rigour of development domain (score 89%). The EAU previously employed a condensed process of evidence review; however, they recently announced a gradual implementation of the Cochrane methodology across their guideline panels. The 2017 version of the NLUTD CGs contained three new systematic reviews using this methodology. All three CGs used a validated grading system to describe the strengths and limitations of the underlying body of evidence.

All CGs scored highly in the clarity of presentation domain, as the recommendations were easily identifiable, specific and unambiguous. It is important that all management options are presented, so end users can make fully informed clinical decisions. Although the ICI do not promote their CGs to be used directly in clinical practice, in reality they may be interpreted to be used in this way. Instead, the ICI GCs are endorsed as the reference work for the condition of interest, thus they consider an exhaustive number of management strategies compared to the other CGs[15]. This helped achieve the highest score in this domain (94%). The EAU lost points in this domain, as despite providing a thorough discussion on behavioural techniques, no graded recommendations were made for certain forms of management.

It has been demonstrated that improvement in health outcomes is related to adherence to CGs[16]; however, due to multifaceted barriers to implementation, uptake of CGs has remained notoriously low.[17] A 2007 survey sent out to Dutch urologists revealed that the EAU CGs for NLUTD were not systematically employed in clinical practice [18]. The applicability domain measures the steps taken by the developers to improve uptake of the CGs and to what extent the resource implications of application have been considered. In light of the international scope of the ICI, the CGs achieved a low score for the applicability domain (54%). NICE and EAU have designated implementation teams with the aim of promoting uptake of CGs and overcoming barriers to implementation. They scored 90% and 60% respectively. Due to their national scope (UK only); it is easier for NICE to introduce strategies at a local level, including promoting a wide range of resources (e.g. educational presentations and patient leaflets), and engaging multiple organisations. For the same reason, NICE were able to consider the cost-effectiveness of treatments. Integrating economic evaluation into CGs is imperative given the ever increasing healthcare costs and the introduction of costly innovative products [19]. The EAU was unique from the other CGs in that it has a designated team named the 'Social Media (SoMe) working group', who are responsible for promoting the guidelines on Facebook and Twitter. This is particularly important in an age where SoMe has become a frequent vehicle to disseminate medical information.

The editorial independence domain reviews whether the funding body may have influenced the guideline content, and asks whether potential conflicts of interests (COI) have been adequately recorded and addressed. None of the CGs were pharmaceutical industry funded; however, some

development members in all CGs declared financial relationships with industry. The NICE and EAU guidelines have specific policies on how to manage COI, thus scored a higher percentage in this domain (88% in both CGs). Both CGs employ cautionary measures such as excluding development members from voting or in the development of recommendations related to their area of COI[20,21]. A qualitative study into the NICE COI process determined that it was effective and transparent; however, as expected, it relied upon a process of self-reporting, which runs the risk of important omissions being made[22]. Some alternative opinion suggests that financial relationships with industry could provide unique and important expertise into the input of guideline development[23].

There are some limitations in this study that should be discussed. The AGREE II developers do not provide thresholds for what should be considered 'low quality' and 'high quality' CGs, thus interpretation of the resulting scores was ultimately a subjective exercise. Although the number of appraisers in this study was in line with the recommendations from the AGREE II collaborators, increasing this number could have improved the inter-rater reliability. One of the authors (MJD) was involved in the development of the ICI CGs, which could have introduced an element of bias; for this reason, MJD was not involved in the appraisal of any of the CGs for the current study. In addition, two authors work in Urology Research and Development based roles for a pharmaceutical company (ES & JN) and all authors based in the UK, which could affect the reliability of conclusions.

CONCLUSIONS

All CGs had their inherent advantages and disadvantages, although all were still deemed to be of high quality. The lower score overall for the ICI guidelines could partly be attributed to the contrasting purpose of development and intention of use as an international guidance document. NICE CGs were deemed to be of the highest quality due to attributes such as the involvement of multiple stakeholders and economic evaluation of treatment options. The EAU has some promising initiatives that will elevate the quality of their CGs in coming years. Incorporating the AGREE II instrument in the development of CGs and better collaboration between the ICI, NICE and EAU could improve the quality of NLUTD CGs and ultimately improve health outcomes for this important patient group. Institutions will have to overcome barriers such as ensuring the clinical and economic applicability of recommendations to a diverse range of healthcare systems across the globe.

REFERENCES

(1) McDonald, C.; Winge, K.; Burn, D. J. Lower urinary tract symptoms in Parkinson's disease: Prevalence, aetiology and management. *Parkinsonism Relat Disord* **2017**, *35*, 8-16.

(2) Ginsberg, D. The epidemiology and pathophysiology of neurogenic bladder. *Am J Manag Care* **2013**, *19*, s191-196.

(3) Game, X., Fowler, C, J., Panicker, J. Neuropathic Bladder Dysfunction. *Trends in Urol, Gynae & Sexual Health* **2010**, *15*, 23-28.

(4) Tapia, C. I.; Khalaf, K.; Berenson, K.; Globe, D.; Chancellor, M.; Carr, L. K. Health-related quality of life and economic impact of urinary incontinence due to detrusor overactivity associated with a neurologic condition: a systematic review. *Health Qual Life Outcomes* **2013**, *11*, 13.

(5) About the National Academy of Medicine. https://nam.edu/about-the-nam/ (accessed 12th May 2018).

(6) Graham R, M. M., Wolman DM, Greenfield S, Steinberg E.: *Clinical Practice Guidelines We Can Trust*; The National Acadamies Press: Washington DC, 2011; Vol. 290.

(7) Graham, I. D.; Harrison, M. B. Evaluation and adaptation of clinical practice guidelines. *Evidence Based Nursing* **2005**, *8*, 68-72.

(8) Apostolidis A., Drake, M, J., A, Emmanuel., et al.: Neurologic Urinary and Fecal Incontinence In *Incontinence*; International Consultation on Urological Diseases Tokyo, 2017.

(9) Bloc, B., Pannek, J., Castro-Diaz, D., Popolo, G., Groen, J., Hamid, R., Karsenty, G., Kessler, T, M., Bossier, R., Ecclesone, H., Padilla Fernandez, B., Gross, T., t' Hoen, L., Musco, S., Phe, V., Schneider, M, P. *EAU Guidelines on Neuro-Urology*: The Netherlands2017.

(10) NICE. *Urinary incontinence in neurological disease: management of lower urinary tract dysfunction in neurological disease*2012.

(11) Brouwers, M. C.; Kho, M. E.; Browman, G. P.; Burgers, J. S.; Cluzeau, F.; Feder, G.; Fervers, B.; Graham, I. D.; Grimshaw, J.; Hanna, S. E.; Littlejohns, P.; Makarski, J.; Zitzelsberger, L.; Consortium, A. N. S. AGREE II: advancing guideline development, reporting and evaluation in health care. *CMAJ* **2010**, *182*, E839-842.

(12) Koo, T. K.; Li, M. Y. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med* **2016**, *15*, 155-163.

(13) Rosenfeld, R. M.; Shiffman, R. N. Clinical practice guideline development manual: a quality-driven approach for translating evidence into action. *Otolaryngol Head Neck Surg* **2009**, *140*, S1-43.

(14) Jiang, M.; Liao, L. Y.; Liu, X. Q.; He, W. Q.; Guan, W. J.; Chen, H.; Li, Y. M. Quality Assessment of Clinical Practice Guidelines for Respiratory Diseases in China: A Systematic Appraisal. *Chest* **2015**, *148*, 759-766.

(15) Publication Information. http://www.icud.info/index.html (accessed 2nd Jan 2018).

(16) Woolf, S. H.; Grol, R.; Hutchinson, A.; Eccles, M.; Grimshaw, J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. *BMJ* **1999**, *318*, 527-530.

(17) Jha, A. K.; Orav, E. J.; Zheng, J.; Epstein, A. M. Patients' perception of hospital care in the United States. *N Engl J Med* **2008**, *359*, 1921-1931.

(18) Rikken, B.; Blok, B. F. Management of neurogenic bladder patients in The Netherlands: do urologists follow guidelines? *Neurourol Urodyn* **2008**, *27*, 758-762.

(19) Ramsey, S. D. Economic analyses and clinical practice guidelines: why not a match made in heaven? *J Gen Intern Med* **2002**, *17*, 235-237.

(20) EAU. *Managing conflicts of interest in the EAU Guidelines GO Policy*: The Netherlands2017.

(21) NICE. *Developing NICE guidelines: the manual*: Manchester2014.

(22) Graham, T.; Alderson, P.; Stokes, T. Managing conflicts of interest in the UK National Institute for Health and Care Excellence (NICE) clinical guidelines programme: qualitative study. *PLoS One* **2015**, *10*, e0122313.

(23) IOM: Conflicts of Interest and Development of Clinical Practice Guidelines. In *Conflict of Interest in Medical Research, Education, and Practice*; Lo, B., Field, M, J., Ed.; National Academies Press: Washington (DC), 2009.