Does the internet provide accurate and valid health information regarding disorders of sex development?

T Candler[1], A Hough[2], A Hamilton-Shield[3], J Alderson[4], E Crowne[4]
[1]MRC The Gambia at LSHTM [2]University of Exeter Medical School[3]University of Oxford Medical School [4]Bristol Royal Hospital for Children

Background

- Understanding disorders/differences of sex development (DSD), can be difficult for patients and their families due to their complexity and low prevalence.
- Increasingly, families are turning to the internet to access health information including for DSD.
- However the quality, validity and accuracy of the information available online regarding DSD has not been formally assessed before.

Aim: To assess the quality, validity and accuracy of website health information concerning commonly searched terms related to DSD

Methods

- Families of children with DSD were consulted to generate 5 search terms: “Disorders of Sex Development OR Differences of Sex Development”, “Congenital Adrenal Hyperplasia” (CAH), “Ambiguous Genitalia OR Atypical Genitalia”, “Cliteromegaly OR Clitoromegaly” and “Micropenis”.
- Top 20 Google search results were scored by two independent reviewers using the validated QUality Evaluation Scoring Tool (QUEST)[1]
- The tool scored 6 domains (authorship attribution, conflict of interest, currency, complementarity and tone), with a maximum score of 28 (figure 1).
- Website inclusion criteria: article/information-like leaflet format, in English, no payment/login required, and articles considering aetiology/diagnosis/treatment of disorder

QUEST Tool Assessment Criteria

Authorship (Score x 1)
0 – No indication of authorship or username
1 – All other indications of authorship
2 – Author’s name and qualification clearly stated

Attribution (Score x 3)
0 – No sources
1 – Mention of expert source, research findings (though with insufficient information to identify the specific studies), links to various sites, advocacy body, or other
2 – Reference to at least one identifiable scientific study, regardless of format (e.g., information in text, reference list)
3 – Reference to mainly identifiable scientific studies, regardless of format (in >50% of claims)

For all articles scoring 2 or 3 on Attribution: (Score x 1)
Type of study
0 – In vitro, animal models, or editorials
1 – All observational work
2 – Meta-analyses, randomized controlled trials, clinical studies

Conflict of interest (Score x 3)
0 – Endorsement or promotion of intervention designed to prevent or treat condition (e.g., supplements, brain training games, foods) within the article
1 – Endorsement or promotion of educational products & services (e.g., books, care home services)
2 – Unbiased information

Currency (Score x 1)
0 – No date present
1 – Article is dated but 5 years or older
2 – Article is dated within the last 5 years

Complementarity (Score x 1)
0 – No support of the patient-physician relationship
1 – Support of the patient-physician relationship

Tone (includes title) (Score x 3)
0 – Fully supported (authors fully and unequivocally support the claims, strong vocabulary such as “cure”, “guarantee”, and “easy”, mostly use of non-conditional verb tenses (“can”, “will”), no discussion of limitations)
1 – Mainly supported (authors mainly support their claims but with more cautious vocabulary such as “can reduce your risk” or “may help prevent”, no discussion of limitations)
2 – Balanced/cautious support (authors’ claims are balanced by caution, includes statements of limitations and/or contrasting findings)

Figure 1. Scoring criteria for QUEST Tool[1]

Results

- Thirty per cent of Google search results did not satisfy inclusion criteria, leaving a total 70 webpages for analysis.

Type of website	Micropenis (%)	Cliteromegaly (%)	Disorders of sexual development (%)	Congenital adrenal hyperplasia (%)	Ambiguous genitalia (%)	Total (%)
Hospital	4 (20)	0 (0)	4 (20)	3 (15)	10 (50)	21 (21)
Charity	1 (5)	1 (5)	1 (5)	3 (15)	0 (0)	6 (6)
General information	1 (5)	1 (5)	1 (5)	1 (5)	0 (0)	4 (4)
Health information	5 (25)	1 (5)	1 (5)	7 (35)	3 (15)	17 (17)
Publications	1 (5)	10 (50)	7 (35)	0 (0)	2 (10)	20 (20)
Tabloid*	6 (30)	0 (0)	0 (0)	0 (0)	0 (0)	6 (6)
Inappropriate format*	1 (5)	3 (15)	1 (5)	3 (15)	2 (10)	10 (10)
Paywall*	1 (5)	4 (20)	5 (25)	3 (15)	3 (15)	16 (16)

Table 1. Category of website per search term in Top 20 Google hits *=reason for exclusion

- There was substantial inter-rater agreement across all domains, except ‘Tone’ where there was moderate agreement.

	Authorship	Attribution	Type of study	Conflict of interest	Currency	Complementarity	Tone
Observed kappa	0.83	0.72	0.75	0.71	0.95	0.71	0.55
SE	0.09	0.08	0.09	0.11	0.09	0.12	0.11

Table 2. Assessment of inter-rater reliability across all 6 QUEST domains.

- There was no evidence that average QUEST score varied between chosen search terms, or google rank.
- ‘Micropenis’ demonstrated the most variable results (SD 7.4), ‘CAH’ had the least variable results (SD 3.4)

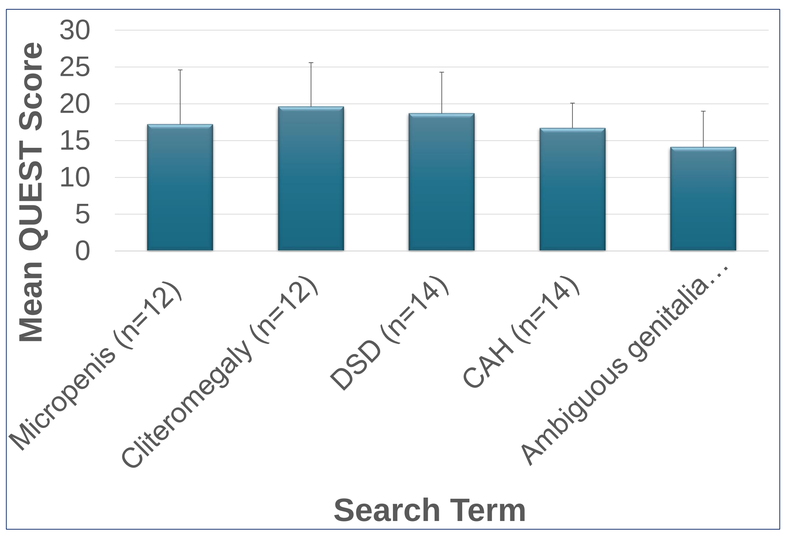

Figure 2. Mean QUEST score according to search term. Data are mean ± Standard deviation (SD)

- There was strong evidence that average QUEST score was related to category of website ($p<0.001$), with hospital websites the lowest scoring category.

Results

Website type	Average score						
	Authorship (2)	Attribution (3)	Conflict of interest (2)	Currency (2)	Complementarity (1)	Tone (2)	Overall score (28)
Hospital	0.4	0.5	1.2	0.7	0.8	1.6	11.8
Charity	0.7	1	1.8	1	0.4	1.5	15.2
General information	0	1.9	2	2	0	1.9	21
Health information	1.1	1.2	1.9	1.4	0.9	1.7	18
Peer reviewed publication	1.3	2.1	2	1.5	0.4	1.8	22.1

Table 3. Average score across QUEST Domains by category of website

Website type	Average difference (p value)
Hospital	Baseline, average score of 11.8
Charity	+3.37 (0.6)
General information	+9.17 (0.003)
Health information	+6.17 (0.001)
Peer reviewed publication	+10.25 (<0.001)

Table 4. Scheffé multiple comparison test with average difference in mean scores compared to Hospital category

Conclusions

- A high proportion of articles in Google searches are either not accessible or are from tabloid sources
- More colloquial terms e.g. micropenis have more variability in information quality
- This study provides further validation of the QUEST score with near perfect inter-rater agreement across nearly all categories
- The lowest quality information comes from hospital websites – often due to lack of clarity about who the author was, where the information comes from or they were promoting their own healthcare services (especially US sites)
- The highest quality information comes from peer reviewed publications
- The main limitation of QUEST is there is no score of *accessibility* nor is there a clear cut-off of what score would be deemed acceptable or indeed ‘good’.

We would recommend professionals consider the quality criteria in the QUEST tool when designing health information websites for DSD.

Acknowledgements

We would like to thank the parents of patients at Bristol Royal Hospital for Children who helped with this project

References:

[1]Robillard JM, Jun JH, Lai JA, Feng TL. The QUEST for quality online health information: validation of a short quantitative tool. *BMC Med Inform Decis Mak*. 2018;18(1):87. Published 2018 Oct 19. doi:10.1186/s12911-018-0668-9