Writing answers:

How are our answers understood by readers?

What are the expectations?

How can texts be improved?

Olav Spigset
Senior consultant, professor
Department of Clinical Pharmacology
St Olav University Hospital
Norwegian University of Science and Technology
Trondheim, Norway

CrossMark

PHARMACOEPIDEMIOLOGY AND PRESCRIPTION

Quality assessment of structure and language elements of written responses given by seven Scandinavian drug information centres

Linda Amundstuen Reppe[1,2,3] · Olav Spigset[2,3] · Jens Peter Kampmann[4] ·
Per Damkier[5] · Hanne Rolighed Christensen[4] · Ylva Böttiger[6] · Jan Schjøtt[7,8,9]

… and thanks to everyone else involved!

Relationship Between Time Consumption and Quality of Responses to Drug-related Queries: A Study From Seven Drug Information Centers in Scandinavia

Linda Amundstuen Reppe, MSc.Pharm[1,2,3]; Stian Lydersen, PhD[4];
Jan Schjøtt, MD, PhD[5,6,7]; Per Damkier, MD, PhD[8];
Hanne Rolighed Christensen, MD, PhD[9]; Jens Peter Kampmann, MD, Dr.Med.Sci[9];
Ylva Böttiger, MD, PhD[10]; and Olav Spigset, MD, PhD[2,3]

http://www.clinicaltherapeutics.com/
article/S0149-2918(16)30395-2/pdf

Method

- Six fictitious queries
  - Five patient-specific and one general
  - Originally in Norwegian
- Principles
  - "Typical" questions
  - Most common query catogories
  - No clear-cut answers
- Translated to Swedish and Danish

Query 1-2

- A female patient presents with deep, infected pockets close to the jaw bone, and needs to have these rinsed every 4th to 5th week for the next six months. The patient uses alendronate 70 mg once weekly. Should alendronate be discontinued during treatment?

- A male patient with formerly performed gastric bypass needs treatment for *Helicobacter pylori* infection due to symptomatic ulcus. The GP wants to start treatment with pantoprazole 40 mg once daily, metronidazole 400 mg twice daily and amoxicillin 750 mg twice daily. Will the absorption of the drugs be reduced, and are dosage adjustments necessary?

Query 3-4

- A pregnant women manifests with moderate depression (MADRS score 29), and there is indication for treatment with an antidepressant. What antidepressant is the first choice of drug during pregnancy?

- A GP has registered an increasing use of Ginkgo biloba in nursing care homes and home nursing services. He (she) has also registered that ginkgo might increase bleeding time. What documentation exists on this topic, and what is the relevance for concomitant use of e.g. warfarin, acetylsalicylic acid, clopidogrel and enoxaparin?

5.

A male patient, 75 years old, has gradually developed impaired cognition for the last 5-6 months (MMS score 18 at examination). He has essential hypertension treated with atenolol 100 mg once daily and losartan/hydrochlorothiazide 100/12.5 mg once daily. His blood pressure was 130/90 mmHg at the latest appointment. He also uses simvastatin 40 mg and acetylsalisylic acid 160 mg once daily. He uses paroxetine 40 mg in the morning for anxiety/depression, diazepam 5 mg as needed and promethazine/propiomazine*50 mg at night for sleep. He also uses tolterodine 2.8/4 mg**once daily for overactive bladder (dosage increased from 1.4/2 mg** three months ago). The patient does not smoke. Can any of these drugs, or drug interactions, increase the risk of impaired cognition?

*Promethazine is marketed in Norway and Denmark, but not in Sweden. For the Swedish query, we used the drug propiomazine. Both drugs are derivatives of phenthiazines with antihistaminic and anticholinergic effects.
** Tolterodine is marketed as 2 and 4 mg depot capsules in Norway and Sweden, and as 1.4 and 2.8 mg depot capsules in Denmark.

Query 6

- A female patient, 13 weeks post partum, presents with active ulcerative colitis. She has earlier been treated with sulfasalazine 500 mg x 3 (discontinued during pregnancy). Can she use sulfasalazine while breast-feeding?

Method

- Eleven DICs invited
  - Seven DICs included
    - 4 NO + 1 SE + 2 DK
- All centers informed about the study
  - Blinded to which questions were the fictitious ones
- 22 general practitioners recruited
  - 3-4 per DIC
- Carried out druing an 8-week period

Responses evaluated by

- Seven clinical pharmacologists
  - (2 NO, 2 SE, 3 DK)
- Six general practitioners (GPs)
  - (2 NO, 2 SE, 2 DK)
- One language expert
  - Master of Arts in Rhetoric
  - Involved in "Clear language inNorway's Civil Service"
  - Trained in all Scandinavian languages

Assessment of language quality (I)

A pilot study defined 8 quality criteria – scored from 0 to 4 for for each

No.	Criteria	Mean score (min to max)	Median score
1	Does the response have a distinct structure?	3.1 (2 to 4)	3
2	Are words and concepts used in the query, also repeated in the response?	2.5 (1 to 4)	3
3	Are words and concepts explained or defined? Are other words used to explain them?	2.5 (1 to 4)	3
4	Is the response written in an abstract or concrete style?	2.4 (1 to 3)	2.5

Assessment of language quality (II)

No.	Criteria	Mean score (min to max)	Median score
5	Does the use of pronouns, verbs and dependent clauses help making the text less compressed?	2.5 (0 to 4)	3
6	Is it easy for the reader to understand who should perform the described actions?	2.4 (0 to 3)	2
7	Are answers to the query given scattered or as one common conclusion?	2.8 (1 to 4)	3
8	Is it easy to find a concrete answer to the query?	2.9 (1 to 4)	3

Total language score vs. number of responses
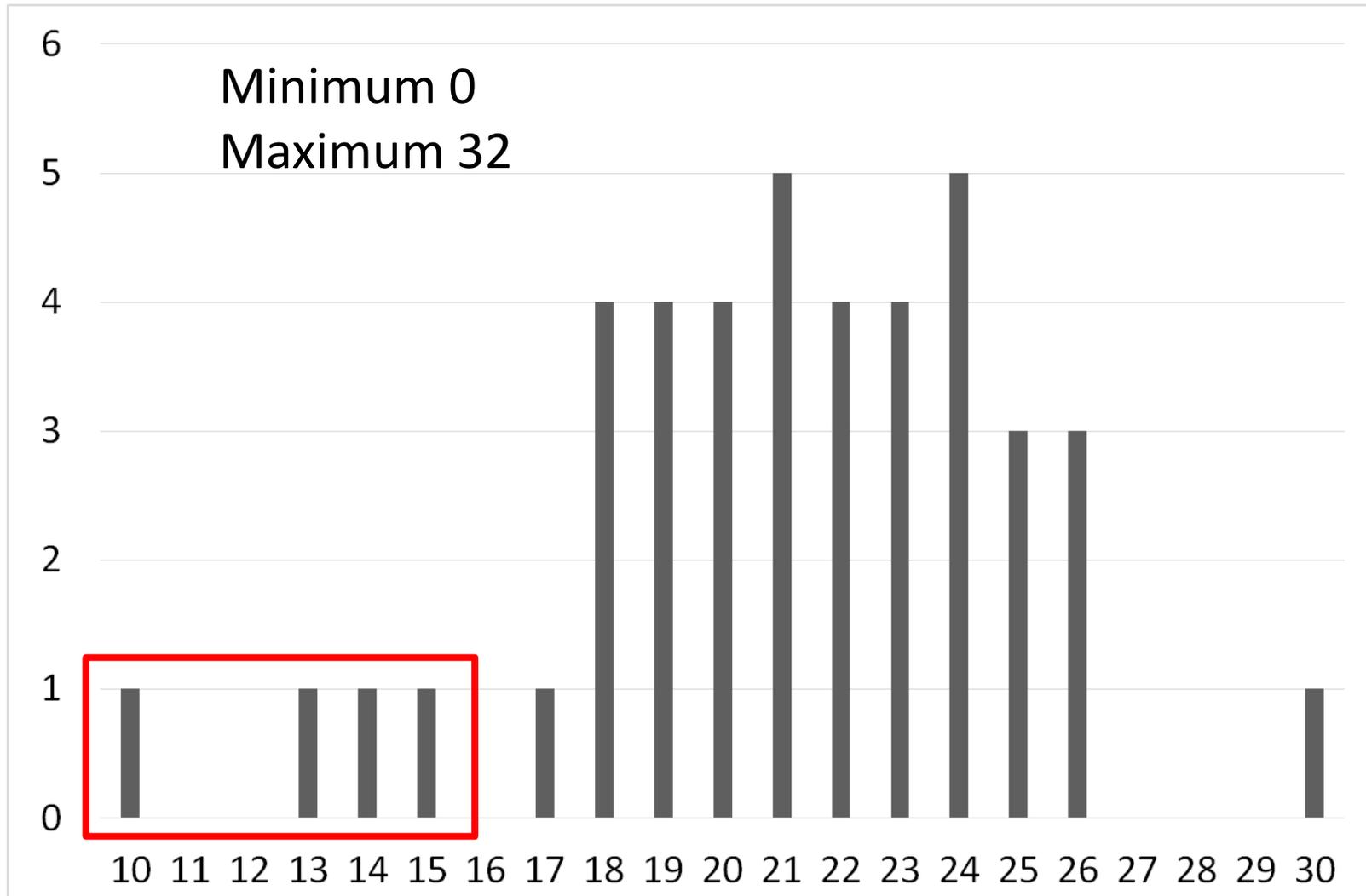
Language elements

- All responses fulfilled their intentions
- However, room for imporvement
- It was assumed that the medical language is known for both parties
- Avoid oxymorons!
- Complicated terms and unexplained abbreviations that should have been clarified
  - Double-dummy; drug holiday; weight-adjusted dosage; stereochemistry; rapid acetylator; confounding by indication
  - SNRI; PPHN; AUC; C_{max}; t_{max}

The clinical experts…

- Evaluated the 42 answers one by one by means of a questionnaire
- 16 criteria in total
- Scores from 0 to 4 (or 4 to 0) – or dichotomous
- Plus comments in free text

Quality Criteria	Cl. Ph.	GP
Overall impression	X	X
Relevance	X	X
Conclusions/advice coherent	X	X
Length	X	X
Unnecessary information		X
Readability		X
Logical structure		X
Clinical usefulness		X
Would you feel safe in making a decision?		X

Quality Criteria	Cl. Ph.	GP
Easy to understand?		X
Staff member provided extra value?	X	
All relevant aspects covered	X	
Advice given (Y/N)	X	X
No incorrect statements	X	
No statements that could not be understood (or could be misunderstood)	X	
No expressions that should have been explained	X	

Quality scores

KIND   **STRICT**

Table IV. Summarized quality scores for internal (clinical pharmacologists) and external (general practitioners) experts reviewing responses to six drug-related queries posed to seven Scandinavian drug information centers (DICs).[*]

Query No.	**0–48** Internal Experts' Sum Score[†]			**0–40** External Experts' Sum Score[‡]		
	Mean (%)	Minimum (%)	Maximum (%)	Mean (%)	Minimum (%)	Maximum (%)
I	37.0 (78)	33.3 (69)	40.9 (85)	26.7 (67)	21.3 (53)	33.0 (83)
II	37.0 (78)	28.2 (59)	41.2 (86)	29.9 (75)	17.2 (43)	35.2 (88)
III	38.1 (79)	35.4 (74)	40.8 (85)	32.0 (80)	27.8 (70)	36.2 (91)
IV	36.2 (75)	30.8 (64)	40.3 (84)	27.7 (69)	20.3 (51)	30.8 (77)
V	35.8 (75)	30.0 (63)	39.6 (83)	28.4 (71)	22.0 (55)	35.0 (88)
VI	38.7 (81)	35.1 (73)	40.3 (84)	32.5 (81)	27.2 (68)	36.8 (92)
All queries	37.1 (77)	28.2 (59)	41.2 (86)	29.5 (74)	17.2 (43)	36.8 (92)

[*]The possible range of sum score was 0 to 40 for the external experts' score and 0 to 48 for the internal experts' score. The percentages represent the specific score in percentage of the maximum possible score in the same category.
[†]Criteria 2 to 7 and 11 to 16 (Table II) were included in the internal experts sum score.
[‡]Criteria 2 to 10 and 13 (Table II) were included in the external experts sum score.

Time consumption

Table III. Time consumption for seven Scandinavian drug information centers (DICs) answering six specific drug-related queries.

Query No.	Requested Time Frame	Time Consumption per Response (h:min)				Mean Response Time (d)
		Mean	Median	Minimum	Maximum	
I	Within a week	05:16	03:35	01:35	12:36	3.8
II	Within 2 days	04:29	03:44	02:00	10:39	1.4
III	Within the next day	01:38	01:09	00:17	03:30	0.1
IV	None	03:59	02:05	00:42	16:04	4.0
V	None	04:50	04:08	02:06	12:25	4.1
VI	None	06:38	02:50	00:35	24:17	1.4
All queries	–	04:26	03:08	00:17	24:17	2.6

Four centers: Mean of 2-3 hours per query
Two centers: Mean of 3-4 hours per query
One center: Mean of >13 hours per query

Regression analysis

Table V. Results from mixed model linear regression analyses exploring whether time consumption predicts different measures of quality of written answers to six drug-related queries posed to seven Scandinavian drug information centers (DICs) (n = 40; data from 2 cases are unavailable).[*]

Dependent Variable	Possible Score	Estimated Change in Quality Score per Extra Hour Spent on Response, (95% CI)	P
Internal experts' common quality score[†]	0–28	–0.04 (–0.17 to 0.09)	0.52
External experts' common quality score[†]	0–28	–0.20 (–0.33 to –0.06)	0.004
Internal experts' sum score[‡]	0–48	–0.06 (–0.24 to 0.13)	0.56
External experts' sum score[§]	0–40	–0.21 (–0.47 to 0.04)	0.10
Internal experts' overall impression of response	0–4	0.01 (–0.02 to 0.04)	0.47
External experts' overall impression of response	0–4	–0.05 (–0.08 to –0.01)	0.005

[*]Sums of quality scores were calculated using different combinations of 16 individual quality criteria assessed by internal experts (clinical pharmacologists) and external experts (general practitioners).
[†]Criteria 2 to 7 and 13 (Table II) was included in the score.
[‡]Criteria 2 to 7 and 11 to 16 (Table II) was included in the score.
[§]Criteria 2 to 10 and 13 (Table II) was included in the score.

Why?

- Sensitivity analyses did not change the results
- Inclusion of covariates (such as experience) did not change the results

- Can we rely on the result?
  - Small material
  - Cannot be extrapolated outside the time intervals in the study

Interesting?

- Score reduced by 0.2 per hour
- I.e. a reduction of 1.0 per 5 hours
- On a scale with a possible maximum of 28

- No linear effect
- Extra time used for adding (uninteresting) details?
- Extra time used for increasing length

The clinical experts...

- Evaluated the 42 answers one by one
- Comments in free text were encouraged
  - In total, 334 comments were received
- In addition, one overall assessment for each of the six queries:
  - The degree of concordance in the conclusions
    - Speficially, did the documentation cited and/or the advice given differ between responses?

Clin. ph. – aspects *increasing* quality

- Enclosed articles should be "to the point", highly relevant and not too demanding for the enquirer to read

- Systematic and well-documented response that results in specific advice

- Very thorough and deepened response with a review of primary literature. It depends on your preferences whether this is necessary, but the conclusion is precise

- The best responses give specific and practically oriented answers, while the poorest ones give more floating, indirect answers like «studies have not shown etc.»

Clin. ph. – aspects *decreasing* quality

- Generally, I feel it is «bad service» to write very short and then refer to many enclosed articles in an un-prioritized order. Then the enquirer has to find the answer himself, especially when the enclosed articles have different conclusions
- One response distinguishes itself from the others with inadequate background information, even though it is possible that the response is based on literature and sources that are not referred to
- I miss a summary; this makes it difficult to get an overview of the response
- This response does not contain anything that resembles a specific advice. The sentence "should be assessed based on a clinical risk-benefit assessment" is a triviality. Isn't that always true??

GPs – aspects *increasing* quality

- The response is concise. It does not contain unnecessary details, and is presented with a clear conclusion that is logical and practical and useful
- Well-done and thorough review of all drugs involved. A clear, sensible and practically useful conclusion too. The response witnesses considerable clinical insight. Good argumentation and some background information that support the given advice
- Since the responses are supposed to be relevant for clinicians, it is important to have good and transparent routines for collection of evidence.
- To me, it is crucial that they dare making a specific decision

GPs – aspects *decreasing* quality

- All responses conclude that the evidence is sparse, but not all of them dare to draw a useful conclusion, and that is what you are lacking as a physician, whom is supposed to use the answer. The physician has to make a decision!
- If you enclose an article, it is important to comment WHETHER and WHAT the enquirer should read that gives him/her supplementary information to the response, otherwise you should not enclose it
- Information on possible measures of action are given, but no prioritization or specific order is suggested. That might be correct, because no "perfect" answer exist. However, I would prefer that they were more specific, e.g. suggested one or two measures to be tried first

A few specific (bad) examples

- The solubility in water and ethanol was mentioned without commenting the significance of fat solubility on the degree of absorption
- How du you know whether the mother of the baby is a "rapid acetylator"?
- Not: "Consider discontinuing diazepam" – write "We recommend discontinuing diazepam"
- The conclusion is "be cautious". This is your starting point when you ask for advice!
- "Perform a risk-benefit assessment"

What is a correct answer?

- Paroxetine + atenolol
  - Some experts classified it as clincially significant
    - Rewarding the response
  - Other experts classified it as non-significant and uninteresting
    - Discrediting the response

Use of references

- Primary, secondary or tertiary literature?
- Previous DIC responses to queries?
- Which sources were checked vs. which were used?
  - Unreliable data regarding sources consulted

- Complicated
  - No concensus; no clear-cut answers
  - But some answers with "thin" or "weak" reference lists were identified

Co-signatures
and discussion with colleagues

- 27/42 responses (64 %) had co-signatures
- 24/42 (57 %) discussion with colleagues
  – 16/42 (38 %) both methods

- Of the four responses with lowest scores
  – None had a co-signature
  – Three were discussed with colleagues

Staff members at drug information centres should *dare to make a choice about*

Data presented

Advice and conclusions

References/enclosed articles

Thank you!